HHS Public Access

Author manuscript

Gynecol Oncol. Author manuscript; available in PMC 2020 August 11.

Published in final edited form as:

Gynecol Oncol. 2019 July ; 154(1): 72–76. doi:10.1016/j.ygyno.2019.04.011.

Models to predict outcomes after primary debulking surgery: Independent validation of models to predict suboptimal cytoreduction and gross residual disease

Amanika Kumar, M.D.[1], **Shannon Sheedy, M.D.**[2], **Bohyun Kim, M.D., Ph.D.**[2], **Rudy Suidan, M.D.**[3], **Debra M. Sarasohn, M.D.**[4], **Ines Nikolovski, M.B.B.S.**[4], **Yulia Lakhman, M.D.**[4], **Michaela E. McGree, B.S.**[5], **Amy L. Weaver, M.S.**[5], **Dennis Chi, M.D.**[3], **William A. Cliby, M.D.**[1]

[1]Division of Gynecologic Surgery, Mayo Clinic, Rochester, MN

[2]Department of Radiology, Mayo Clinic, Rochester, MN

[3]Department of Gynecologic Oncology, Memorial Sloan Kettering Cancer Center, NYC, NY

[4]Department of Radiology, Memorial Sloan Kettering Cancer Center, NYC, NY

[5]Division of Biomedical Statistics and Informatics, Mayo Clinic, Rochester, MN

Abstract

Objective: Treatment planning requires accurate estimation of surgical complexity (SC) and residual disease (RD) at primary debulking surgery (PDS) for advanced ovarian cancer (OC). We sought to independently validate two published computed tomography (CT) prediction models.

Methods: We included stage IIIC/IV OC patients who underwent PDS from 2003–2011. Two prediction models which included imaging and clinical variables to predict RD>1 and any gross RD, respectively, were applied to our cohort. Two radiologists scored CTs. Discrimination was estimated using the c-index and calibration were assessed by comparing the observed and predicted estimates.

Results: The validation cohort consisted of 276 patients; median age of the cohort was 64 years old and majority had serous histology. The validation and model development cohorts were similar in terms of baseline characteristics, however the RD rates differed between cohorts (9.4% vs 25.4% had RD >1 cm; 50.7% vs. 66.6% had gross RD). Model 1, the model to predict RD >1 cm,

Address Correspondence to: Amanika Kumar, MD, Eisenberg Lobby 71, Mayo Clinic, 200 First Street SW, Rochester, MN 55905, 507-566-8400 (phone), 507-266-9300 (fax), kumar.amanika@mayo.edu.

AUTHORS CONTRIBUTIONS

Study conception and design: AK, RS, DS, IN, YL, DC, WC

Acquisition of data: AK, SS, BK, MM, AW

Analysis and interpretation of data: AK, SS, BK, MM, AW

Drafting of manuscript: AK, MM, AW, WC

Critical revision: AK, RS, DS, IN, YL, DC, WC

The authors have no relevant conflicts of interest.

did not validate well. The c-index of 0.653 for the validation cohort was lower than reported in the development cohort (0.758) and the model over-predicted the proportion with RD >1 cm. The second model to predict gross RD had excellent discrimination with a c-index of 0.762.

Conclusions: We are able to validate a CT model to predict presence of gross RD in an independent center; the separate model to predict RD >1 cm did not validate. Application of the model to predict gross RD can help with clinical decision making in advanced ovarian cancer.

Introduction

Epithelial ovarian cancer (EOC) most frequently presents at the advanced stage (stage IIIC/IV). In the fit patient, ideal treatment for advanced EOC includes a combination of primary cytoreductive surgery followed by adjuvant chemotherapy when cytoreduction to either no visible disease (RD0) or to 1 cm or less (RD1) can be accomplished [1]. Therefore predicting surgical residual disease after primary surgery is clinically useful and would aid in planning and counseling.

The goals for any method to predict residual disease are two-fold: 1) assessment of resectability to RD0 or RD1, and 2) evaluation of surgical procedures needed to accomplish RD0 or RD1 resection. Several methods are available with decreasing degrees of invasiveness, including mini-laparotomy with exploration, laparoscopy, and preoperative radiologic assessment. Laparoscopic approaches to assess extent of disease have been used in multiple centers [2–4]. A predictive score can be calculated according to presence of the specific elements, for example omental cake, peritoneal and diaphragmatic extensive carcinomatosis, mesenteric retraction, bowel and stomach infiltration, and spleen and/or liver superficial metastasis[5]. The derived score correlates with the predicted final residual disease; if near complete resection cannot be reached, primary cytoreductive surgery is abandoned for neoadjuvant chemotherapy. There are disadvantages, foremost being differences in the reliability of the scoring system dependent upon surgeon and center rates of successful cytoreduction and complexity of surgery used; this underlying issues would apply to most predictive models when applied widely. Additional disadvantages to laparoscopic scoring include the need for a major surgical procedure, need for experience and training in the scoring systems, and potential delay of chemotherapy.

To avoid the negatives of a surgical intervention to assess resectability, computed tomography (CT) imaging based scoring systems have been proposed. Several early models have been published based on retrospective single institution cohorts [6,7], but recently two prospective models were derived from a multi-institutional study: one to predict suboptimal cytoreduction (RD >1 cm) and one to predict any gross residual disease[8,9]. In these two models, the authors demonstrate high levels of sensitivity and specificity by using a combination of clinical and radiologic factors but these models have yet to be tested in an external cohort. The first model predicts RD >1 cm and includes three clinical criteria (age, CA-125, ASA score) and 6 radiologic criteria (suprarenal lymph nodes, diffuse small bowel disease, small bowel mesenteric disease, root of superior mesenteric artery (SMA) disease, perisplenic disease, and lesser sac disease) [8]. The second model predicts any residual disease and includes the same 3 clinical criteria and 8 radiologic criteria (root of SMA

disease, splenic hilum disease, lesser sac disease, gastrohepatic or porta hepatis disease, gallbladder fossa disease or liver intersegmental fissure lesion, retroperitoneal lymph nodes above the renal hilum, small bowel disease, and moderate to severe ascites) [9]. Our goal was to validate these two models in an independent patient cohort.

Methods

This is a single institution retrospective cohort study approved by the Mayo Clinic Institutional Review Board. Patients were identified from a prospectively maintained ovarian cancer surgical database. Patients that underwent primary cytoreductive surgery for stage IIIC or IV ovarian, fallopian tube, or primary peritoneal cancer from 1/2/2003–12/30/2011 were included. Patients who underwent surgery with palliative intent or received neoadjuvant chemotherapy, or who had denied access to their medical records for research were excluded. For inclusion, patients were required to have a digitally available contrast-enhanced CT scans of the abdomen and pelvis.

Patient demographics were abstracted from the medical record. Intraoperative and postoperative information on residual disease, stage, and surgical complexity were also abstracted. Residual disease was categorized into three groups: RD0 defined as no gross residual disease, RD1 defined as residual disease 0.1–1 cm, and RD >1 cm. CT scans were assessed for 20 radiologic findings by one of two abdominal radiologists (BK and SS) who reviewed the CT criteria definitions with radiologists from the authors of the published models [8,9]. Each visualized lesion was measured bi-dimensionally. All lesions were given a qualitative score of 1–5 according to degree of radiologic certainty that a lesion identified on CT was a metastatic lesion: 1=definitely normal, 2=probably normal, 3=indeterminate, 4=probably metastatic, 5=definitely metastatic. After scoring each radiologic finding, malignancy was determined to be present or absent as defined by the published methods from the cohorts we were validating [8,9]. Specifically, for the suboptimal cytoreduction prediction model (RD >1 cm) the following definitions were used: present if the qualitative score was 4 or 5 AND quantitative measurement was >1 cm. For the gross RD predication model the following definitions were utilized: present if qualitative score was 4 or 5 with the exception of lesser sac lesion which was considered present if the qualitative score was 4 or 5 AND quantitative measurement was >1 cm. Anything not meeting the definition of present was defined as absent.

Statistical analysis was performed using the SAS version 9.4 software package (SAS Institute, Inc.; Cary, NC). Patient characteristics were descriptively summarized using median and range for age and preoperative CA-125 and using frequency count and percentage for all categorical variables. We assessed the discrimination and calibration of the two prediction models, separately, using the total summated predictive scores previously determined for each model. The first model was for the outcome of suboptimal disease (RD >1 cm) and the second model was for the outcome of gross RD. Discrimination refers to the model's ability to correctly discriminate between patients with and without the outcome of interest. Discrimination was quantified using the concordance index (c-index) estimated from a logistic regression model fit using the outcome as the binary variable and the total summated predictive score as the independent variable. The c-index is identical to the area-

under-the-curve for a Receiver Operating Characteristic (ROC) curve. A 95% confidence interval for the AUC was generated using 1000 bootstrapped samples with replacement. Calibration refers to the agreement between the observed outcomes and the predicted outcomes, and was assessed within the collapsed categories of the predictive scores used in the previously published papers.

Results

A total of 276 patients were eligible from January 2003-December 2011. Patients without available digital CT images for review were excluded. Patient characteristics for the model development and validation cohorts are presented in Table 1. Characteristics were generally similar between the two cohorts, in terms of age, ASA score, FIGO grade, and serous histology. Stage distribution of disease was similar in that 76.1% versus 70.9% of patients were stage IIIC and 23.9% versus 26.9% of patients were stage IV, however the model development cohort did include 8 patients (2.3%) that were stage IIIA-B. In the validation cohort, complete cytoreduction (RD0) was achieved in 136 patients (49.3%) while 114 patients (41.3%) had RD1 cytoreduction and 26 patients (9.4%) had RD >1 cm cytoreduction. In comparison, there was a lower rate of RD0 and higher rate of RD >1 cm cytoreduction in the published model development cohort (RD0, RD1, and RD >1 cm in 33.4%, 41.1% and 25.4% of cases, respectively) [8,9].

The presence of each radiologic criterion using the definitions created for the RD >1 cm model was similar in our cohort and the published cohort (Table 2). Similar rates were found for retroperitoneal lymph nodes above the renal hilum > 1 cm (21.4% versus 20.6%), lesser sac lesion (10.9% versus 10.0%), small bowel mesentery lesion >1 cm (21.0% versus 17.4%), and perisplenic lesion >1 cm (19.9% versus 16.9%). There were higher rates of some findings in our cohort, specifically gallbladder fossa lesion >1 cm (14.9% versus 7.1%) and diffuse small bowel adhesions/thickening (19.2% versus 6.9%).

We first evaluated the published model to predict RD >1 cm resection. The published model, with the assigned points reported in brackets, includes 3 clinical criteria (age>60 years [1], preoperative CA-125 ≥500 U/mL [1], ASA score of 3–4 [3]) and 6 radiologic criteria (retroperitoneal lymph nodes above the renal hilum >1 cm [1], diffuse small bowel adhesions/thickening [1], perisplenic lesion >1 cm [2], small bowel mesentery lesion >1 cm [2], root of SMA lesion >1 cm [2], lesser sac lesion >1 cm [4]) [8]. It is notable that only 9% of patients (n=26) in our cohort had RD >1 cm resection compared to 25% (89/350) of the patients in the published cohort. When applied to our cohort, the model performed poorly to predict RD >1 cm. The c-index calculated for this model when applied to our data was 0.653 (95% CI, 0.532–0.773), which is considerably lower than the c-index of 0.758 reported in the original paper. Calibration estimates are presented in Table 4 using the collapsed predictive score categories reported in the original paper. The model provided accurate predictions for patients with scores of 0 and 1–2. Specifically, 6.8% (95% CI, 2.2–15.1%) of the patients in our cohort with a score of 1–2 had RD >1 cm which is consistent with the predicted rate of 10% for this category. However, for patients with scores of 3–4, 5–6, 7–8, and ≥9 the model over-predicted the proportion with RD >1 cm in each category. In

particular, just 17.2% (95% CI, 5.9–35.8%) of the patients in our cohort with a predictive score ≥9 had RD >1 cm compared to 74% predicted by the model.

We next evaluated the published model to predict any gross residual disease. Half (50.7%) of the patients in our cohort had gross residual disease compared to 66.6% (233/350) of the patients in the published cohort. The published model, with the assigned points reported in brackets, includes 3 clinical criteria and 8 radiologic criteria with a similar point-based system. The clinical criteria include: age ≥60 years old [1], preoperative CA-125 ≥600 U/mL [1], ASA score 3–4 [1], and the radiologic criteria include: lesion in the splenic hilum/ligaments [1], gastrohepatic ligament/porta hepatis lesion [1], retroperitoneal lymph nodes above the renal hilum [1], diffuse small bowel adhesions/thickening [1], moderate or severe ascites [2], gallbladder fossa/liver intersegmental fissure lesion [2], lesser sac lesion >1 cm [2], and root of SMA lesion [4] [9]. When applied to our cohort, the ability of the model to discriminate between patients was excellent; the c-index for the validation cohort was 0.762 (95% CI, 0.703–0.815) compared to 0.72 as reported in the original paper. Calibration estimates are presented in Table 5 using the collapsed predictive score categories reported in the original paper. The model provided accurate predictions for patients with total predictive scores of 6–8; the model predicted 87% of the patients in this category to have gross residual disease and the observed rate was 86.1%. However, for patients with scores of 0–2, 3–5, and ≥9 the model over-predicted the rate having gross residual disease (45% predicted vs. 22.9% actual, 68% vs. 53.7%, and 96% vs. 79.3%, respectively).

Considering that the prevalence of patients with RD >1 cm was considerably lower in the validation cohort compared to the model development cohort (9.4% vs. 25.4%), this may partially explain why the predicted rates of RD>1 cm were higher than the observed rates. Therefore we explored the methods described by Janssen et al. for updating the prediction models by calculating a correction factor for the intercept [10]. However, the new intercepts derived from the correction factors did not alter the original scoring, so the calibration remains unchanged.

Discussion

Predicting residual disease after primary cytoreductive surgery would improve treatment planning. Patients unlikely to have successful resection could move directly to alternative approaches. Similarly patients deemed resectable would be triaged to primary surgical resection. Use of such a tool could also be useful in determining which patients may benefit from transfer to centers specializing in surgical cytoreduction. Toward that end, we have evaluated the external validity of two published models that utilize preoperative and radiologic variables in an independent population of advanced stage ovarian cancer. We observed that 1) our cohort and the published cohort had similar rates of disease as measured by strict radiologic criteria, 2) the model to predict RD >1 cm resection did not validate in our external cohort, and 3) the model to predict presence of gross RD did validate but over-predicted the presence of gross RD in 3 of the 4 subcategories of the predictive scores.

There have been many attempts to use radiologic criteria to determine resectability[6,7,11,12]. A study from our own institution of 87 patients only found diffuse

peritoneal thickening and ascites to be predictive of RD >1 cm resection[6]. We have previously published on this same cohort of patients as presented in this paper, but only analyzed a small number of radiologic variables, which again only found diffuse peritoneal thickening, ascites, and omental cake as relevant variables[7]. Neither of these models have been validated. Three other models were externally tested by Rutten et al. and found to have limited predictive ability and the authors question the use of such models[13]. This raises the point that most prediction models will succeed or fail based on the underlying rates of cytoreduction used for the test cohort. When tested in centers with dissimilar rates of cytoreduction, any model is likely to perform poorly. The models tested in this cohort have the advantage of being multi-institutionally derived and have utilized a large number of radiologic variables.

While CT scans are one of multiple tools available to gynecologic oncologists to predict resectability, there are limitations to this approach. Any model for predicting surgical outcome will have false negatives and false positives, and so clinical experience must be used in addition to prediction tools. In addition, radiologists must work closely with surgeons on evaluating CT scans for resectability. The c-index for the validated model is still less than ideal if we were to use CT models on their own. These tools are not intended to substitute for clinical judgment, but can be used as an adjunct to counsel patients or even consider referral of patients to other centers. We prefer to consider the question of resectability separate from surgical fitness, viewing resectability as a surgical outcome that is determined by anatomic location of disease and spread pattern. Surgical fitness is better considered separately to identify those patients who can tolerate a maximal surgical effort. We argue that one weakness of the proposed models is the incorporation of age and functional status, which are more related to patient factors and the ability to tolerate surgery than technical resectability [14,15]. Models that answer these two questions independently may be more useful for treatment planning. Future directions in the development of models addressing resectability should include technologic advances such as artificial intelligence in CT scan reading. These tools may improve the predictability of CT scans.

Strengths of this study include a large cohort of patients from a single institution that has an aggressive upfront approach to primary cytoreduction, similar to the context in which the models were derived. The two cohorts of patients were similar with respect to age, stage, and histology, as well as presence of radiologic criteria. This study and the original studies were performed at tertiary care centers with specialized radiologists dedicated to body imaging, so findings may not be reproducible in other settings. There were differences with regards to overall rate of RD0, RD1, and RD >1 cm resection: these differences between centers is a possible explanation of why the RD >1 cm model was not validated. Again, we stress that such models will be most relevant when underlying rates of cytoreduction are similar. In addition, in this cohort, rates of RD0 resection increased over time from 33% in 2003–2006 to 54% in 2007–2011 owing to a change in philosophy, so the RD0 rate may not be completely explained by resectability, and may in part be reflective of surgical effort and goal [16]. This also highlights the shifting reference points of when R0 is possible with increased training and attention on that goal [17]. The setting for this validation study is also a weakness, and the results may not be generalizable to community center or smaller centers. Further, the evaluation of CT scans in this detailed manner takes significant time

and effort and may not be feasible in every clinical setting. Most centers do not have specialized radiologists that will be trained in this methods, limiting clinical utility for some.

In conclusion, our results demonstrate that published models perform differently depending on the treating center and the patient cohort when clinical variables are included in the model. The published multi-institutional model to predict gross RD appears to validate, and should be useful as an adjunct to clinical judgment to help predict residual disease after primary cytoreductive surgery, especially when the goal is RD0. These models will help guide clinical decision making and counseling for patients with advanced ovarian cancer, however are not sufficient on their own at this point.. They can also be useful in triaging patients to centers specializing in surgical resection: patients with higher scores predicting acceptable rates of RD0 will require high complexity operations. Finally, future directions in the triage of patients will need to incorporate more than just anatomic distribution. Our work looking at the relationship between molecular subtype and surgical outcomes may help further improve these models [18].

Acknowledgments

All authors from Memorial Sloan Kettering Cancer Center were funded in part through the NIH/NCI cancer center support grant P30 CA008748.

References

[1]. Wright AA, Bohlke K, Armstrong DK, Bookman MA, Cliby WA, Coleman RL, et al., Neoadjuvant Chemotherapy for Newly Diagnosed, Advanced Ovarian Cancer: Society of Gynecologic Oncology and American Society of Clinical Oncology Clinical Practice Guideline, Journal of Clinical Oncology. 34 (2016) 3460–3473. doi:10.1200/JCO.2016.68.6907. [PubMed: 27502591]

[2]. Fagotti A, Vizzielli G, Fanfani F, Costantini B, Ferrandina G, Gallotta V, et al., Introduction of staging laparoscopy in the management of advanced epithelial ovarian, tubal and peritoneal cancer: Impact on prognosis in a single institution experience, Gynecologic Oncology. 131 (2013) 341–346. doi:10.1016/j.ygyno.2013.08.005. [PubMed: 23938372]

[3]. Fagotti A, Ferrandina G, Fanfani F, Garganese G, Vizzielli G, Carone V, et al., Prospective validation of a laparoscopic predictive model for optimal cytoreduction in advanced ovarian carcinoma, American Journal of Obstetrics and Gynecology. 199 (2008) 642.e1–642.e6. doi:10.1016/j.ajog.2008.06.052. [PubMed: 18801470]

[4]. Petrillo M, Vizzielli G, Fanfani F, Gallotta V, Cosentino F, Chiantera V, et al., Definition of a dynamic laparoscopic model for the prediction of incomplete cytoreduction in advanced epithelial ovarian cancer: Proof of a concept. Gynecologic Oncology. 139 (2015) 5–9. doi:10.1016/j.ygyno.2015.07.095. [PubMed: 26196319]

[5]. Rutten MJ, van Meurs HS, van de Vrie R, Gaarenstroom KN, Naaktgeboren CA, Van Gorp T, et al., Laparoscopy to Predict the Result of Primary Cytoreductive Surgery in Patients With Advanced Ovarian Cancer: A Randomized Controlled Trial, Journal of Clinical Oncology. 35 (2017) 613–621. doi:10.1200/JCO.2016.69.2962. [PubMed: 28029317]

[6]. Dowdy SC, Mullany SA, Brandt KR, Huppert BJ, Cliby WA, The utility of computed tomography scans in predicting suboptimal cytoreductive surgery in women with advanced ovarian carcinoma, Cancer. 101 (2004) 346–352. doi:10.2214/ajr.165.4.7676985. [PubMed: 15241833]

[7]. Janco JMT, Glaser G, Kim B, McGree ME, Weaver AL, Cliby WA, et al., Development of a prediction model for residual disease in newly diagnosed advanced ovarian cancer, Gynecologic Oncology. 138 (2015) 70–77. doi:10.1016/j.ygyno.2015.04.013. [PubMed: 25913130]

[8]. Suidan RS, Ramirez PT, Sarasohn DM, Teitcher JB, Mironov S, Iyer RB, et al., A multicenter prospective trial evaluating the ability of preoperative computed tomography scan and serum

Ca-125 to predict suboptimal cytoreduction at primary debulking surgery for advanced ovarian, fallopian tube, and peritoneal cancer, Gynecologic Oncology. 134 (2014) 455–461. doi:10.1016/j.ygyno.2014.07.002. [PubMed: 25019568]

[9]. Suidan RS, Ramirez PT, Sarasohn DM, Teitcher JB, Iyer RB, Zhou Q, et al., A multicenter assessment of the ability of preoperative computed tomography scan and Ca-125 to predict gross residual disease at primary debulking for advanced epithelial ovarian cancer, Gynecologic Oncology. 145 (2017) 27–31. doi:10.1016/j.ygyno.2017.02.020. [PubMed: 28209497]

[10]. Janssen KJM, Vergouwe Y, Kalkman CJ, Grobbee DE, Moons KGM, A simple method to adjust clinical prediction models to local circumstances. Can J Anesth/J Can Anesth. 56 (2009) 194–201. doi:10.1161/01.CIR.99.16.2098.

[11]. Horowitz NS, Maxwell GL, Miller A, Hamilton CA, Rungruang B, Rodriguez N, et al., Predictive Modeling for determinination of microscopic residual disease at primary cytoreduction: An NRG Oncology/Gynecologic Oncology Group 182 Study, Gynecologic Oncology. 148 (2018) 49–55. doi:10.1016/j.ygyno.2017.10.011. [PubMed: 29174555]

[12]. Rutten MJ, van de Vrie R, Bruining A, Spijkerboer AM, Mol BW, Kenter GG, et al., Predicting Surgical Outcome in Patients With International Federation of Gynecology and Obstetrics Stage III or IV Ovarian Cancer Using Computed Tomography, International Journal of Gynecological Cancer. 25 (2015) 407–415. doi:10.1097/IGC.0000000000000368. [PubMed: 25695545]

[13]. Rutten IJG, van de Laar R, Kruitwagen RFPM, Bakers FCH, Ploegmakers MJM, Pappot TWF, et al., Prediction of incomplete primary debulking surgery in patients with advanced ovarian cancer: An external validation of three models using computed tomography, Gynecologic Oncology. 140 (2016) 22–28. doi:10.1016/j.ygyno.2015.11.022. [PubMed: 26607779]

[14]. Kumar A, Langstraat CL, DeJong SR, McGree ME, Bakkum-Gamez JN, Weaver AL, et al., Functional not chronologic age: Frailty predicts outcomes in advanced ovarian cancer, Gynecologic Oncology. 147 (2017) 104–109. doi:10.1016/j.ygyno.2017.07.126. [PubMed: 28734497]

[15]. Kumar A, Janco JM, Mariani A, Bakkum-Gamez JN, Langstraat CL, Weaver AL, et al., Risk prediction model of severe postoperative complications after primary debulking surgery for advanced ovarian cancer, Gynecologic Oncology. 140 (2016) 15–21. doi:10.1016/j.ygyno.2015.10.025. [PubMed: 26541980]

[16]. Wallace S, Kumar A, Gree MM, Weaver A, Mariani A, Langstraat C, et al., Efforts at maximal cytoreduction improve survival in ovarian cancer patients, even when complete gross resectin is not feasible, Gynecologic Oncology. 145 (2017) 21–26. doi:10.1016/j.ygyno.2017.01.029. [PubMed: 28159407]

[17]. Sioulas VD, Schiavone MB, Kadouri D, Zivanovic O, Roche KL, O'Cearbhaill R, et al., Optimal primary management of bulky stage IIIC ovarian, fallopian tube and peritoneal carcinoma: Are the only options complete gross resection at primary debulking surgery or neoadjuvant chemotherapy? Gynecologic Oncology. 145 (2017) 15–20. doi:10.1016/j.ygyno.2017.02.023. [PubMed: 28238354]

[18]. Torres D, Kumar A, Bakkum-Gamez JN, Weaver AL et al. "Mesenchymal molecular subtype is an independent predictor of severe postoperative complications after primary debulking surgery for advanced ovarian cancer." Gynecologic Oncology. 152 (2019) 223–227. doi:10.1016/j.ygyno.2018.11.019 [PubMed: 30503050]

Highlights

- CT and clinical attributes can predict the presence of gross residual disease at the completion of primary debulking surgery.
- CT models are limited in their ability to predict a suboptimal cytoreduction (RD > 1 cm).
- Radiologic models are clinically useful in the management of advanced ovarian cancer.

Table 1.

Patient characteristics in model development and validation cohort

Characteristic	Mayo validation cohort (N=276)	Published model development cohort* (N=350)
Age (years), median (range)	64 (21–91)	61 (34–86)
Primary site of disease, N (%)		
Ovary	186 (67.4)	264 (75.4)
Peritoneal	72 (26.1)	44 (12.6)
Fallopian tube	18 (6.5)	42 (12.0)
FIGO stage, N (%)		
IIIA/B	0	8 (2.3)
IIIC	210 (76.1)	248 (70.9)
IV	66 (23.9)	94 (26.9)
FIGO grade, N (%)		
1 or 2	13 (4.7)	19 (5.4)
3	259 (93.8)	328 (93.7)
Unknown	4 (1.4)	3 (0.9)
Histology, N (%)		
Endometrioid/clear cell	17 (6.1)	2 (0.6)
Serous	236 (85.5)	314 (89.7)
Mixed/Other	23 ((8.3)	34 (9.7)
Preoperative CA-125 (U/mL), median (range)	729 (10–45400)	860 (9–38100)
ASA score, N (%)		
1	3 (1.1)	10 (2.9)
2	139 (50.4)	158 (45.1)
3	133 (48.2)	179 (51.1)
4	1 (0.4)	3 (0.9)
Residual disease (RD), N (%)		
RD0, no visible disease	136 (49.3)	117 (33.4)
RD1, 0.1–1.0 cm	114 (41.3)	144 (41.1)
RD >1 cm	26 (9.4)	89 (25.4)

Abbreviations: ASA, American Society of Anesthesiologists; FIGO, International Federation of Gynecology and Obstetrics; IQR, interquartile range; SD, standard deviation.

*Data from Suidan 2014 publication

Table 2.

Presence of each radiologic criteria using the definitions created for the model to predict suboptimal cytoreduction (RD >1 cm)

Characteristic	Criteria present in Mayo validation cohort (N=276)	Criteria present in published model development cohort * (N=350)
Subcapsular liver lesion or perihepatic lesion >1 cm	113 (40.9%)	121 (34.6%)
Liver intraparenchymal lesion >1 cm	12 (4.3%)	9 (2.6%)
Retroperitoneal lymph nodes above the renal hilum (including supradiaphragmatic) >1 cm	59 (21.4%)	72 (20.6%)
Gastrohepatic ligament/porta hepatis lesion >1 cm	54 (19.6%)	73 (20.9%)
Gallbladder fossa lesion >1 cm	41 (14.9%)	25 (7.1%)
Liver intersegmental fissure lesion >1 cm	21 (7.6%)	48 (13.7%)
Lesser sac lesion >1 cm	30 (10.9%)	35 (10.0%)
Spleen intraparenchymal lesion >1 cm	3 (1.1%)	7 (2.0%)
Perisplenic lesion >1 cm	55 (19.9%)	59 (16.9%)
Root of the SMA lesion >1 cm	6 (2.2%)	8 (2.3%)
Small bowel mesentery lesion >1 cm	58 (21.0%)	61 (17.4%)
Omental lesion >1 cm	221 (80.1%)	212 (60.6%)
Ascites (moderate-severe)	97 (35.1%)	154 (44.0%)
Diffuse small bowel adhesions/thickening	53 (19.2%)	24 (6.9%)
Presacral extraperitoneal disease >1 cm	7 (2.5%)	4 (1.1%)
Tumor invading anterior abdominal wall >1 cm	13 (4.7%)	11 (3.1%)
Pulmonary metastasis (lung bases)	3 (1.1%)	13 (3.7%)
Pleural metastasis (lung bases)	9 (3.3%)	17 (4.9%)

* Data from Suidan 2014 publication

Abbreviations: RD, residual disease; SMA, superior mesenteric artery

Table 3.

Presence of each radiologic criteria using the definitions created for the model to predict gross residual disease model

Characteristic	Criteria present in Mayo validation cohort (N=276)	Criteria present in published model development cohort* (N=350)
Subcapsular liver lesion or perihepatic lesion	126 (45.7%)	147 (42.0%)
Liver intraparenchymal lesion	15 (5.4%)	11 (3.1%)
Retroperitoneal lymph nodes above the renal hilum (including supradiaphragmatic)	61 (22.1%)	94 (26.9%)
Gastrohepatic ligament/porta hepatis lesion	55 (19.9%)	88 (25.1%)
Gallbladder fossa lesion/liver intersegmental fissure lesion	64 (23.2%)	85 (24.3%)
Lesser sac lesion >1 cm	30 (10.9%)	35 (10.0%)
Spleen intraparenchymal lesion	5 (1.8%)	9 (2.6%)
Lesion in splenic hilum/ligaments	74 (26.8%)	70 (20.0%)
Root of the SMA lesion	7 (2.5%)	13 (3.7%)
Small bowel mesentery lesion	68 (24.6%)	67 (19.1%)
Omental lesion	232 (84.1%)	220 (62.9%)
Ascites (moderate-severe)	97 (35.1%)	154 (44.0%)
Diffuse small bowel adhesions/thickening	53 (19.2%)	24 (6.9%)
Presacral extraperitoneal disease	8 (2.9%)	4 (1.1%)
Tumor invading anterior abdominal wall	15 (5.4%)	16 (4.6%)
Pulmonary metastasis (lung bases)	3 (1.1%)	13 (3.7%)
Pleural metastasis (lung bases)	9 (3.3%)	17 (4.9%)

*Data from Suidan 2017 publication

Abbreviations: SMA, superior mesenteric artery

Table 4.

Calibration estimates for the predictive model for suboptimal cytoreduction (RD >1 cm)

Total predictive value score	Total patients N (%)	Optimal N	Suboptimal N	Observed suboptimal rate in the Mayo validation cohort (95% CI)*	Predicted suboptimal rate in the published model development cohort
0	17/276 (6.2%)	16	1	5.9% (0.2–28.7%)	5%
1–2	74/276 (26.8%)	69	5	6.8% (2.2–15.1%)	10%
3–4	63/276 (22.8%)	62	1	1.6% (0.04–8.5%)	17%
5–6	55/276 (19.9%)	49	6	10.9% (4.1–22.3%)	34%
7–8	38/276 (13.8%)	30	8	21.1% (9.6–37.3%)	52%
≥9	29/276 (10.5%)	24	5	17.2% (5.9–35.8%)	74%

*Exact 95% confidence intervals for a binominal proportion.

Table 5:

Calibration estimates for the predictive model for gross residual disease (RD)

Total predictive value score	Total patients N (%)	No gross RD N	Gross RD N	Observed gross RD rate in the Mayo validation cohort (95% CI)*	Predicted gross RD rate in the published model development cohort
0–2	96/276 (34.8%)	74	22	22.9% (15.0–32.6%)	45%
3–5	108/276 (39.1%)	50	58	53.7% (43.9–63.4%)	68%
6–8	43/276 (15.6%)	6	37	86.1% (72.1–94.7%)	87%
≥9	29/276 (10.5%)	6	23	79.3% (60.3–92.0%)	96%

*Exact 95% confidence intervals for a binominal proportion.