Risk of Ovarian Cancer Relapse Score
A Prognostic Algorithm to Predict Relapse Following Treatment for
Advanced Ovarian Cancer

Ivana Rizzuto, MRCOG,* Chara Stavraka, MRes, MRCP,* Jayanta Chatterjee, MRCOG,*
Jane Borley, PhD, MRCOG,* Thomas Glass Hopkins, MBBS,† Hani Gabra, PhD, FRCP,**†
Sadaf Ghaem-Maghami, PhD, MRCOG,*† Les Huson, PhD, CStat,‡
and Sarah P. Blagden, PhD, FRCP**†

Objective: The aim of this study was to construct a prognostic index that predicts risk of
relapse in women who have completed first-line treatment for ovarian cancer (OC).

Methods: A database of OC cases from 2000 to 2010 was interrogated for International
Federation of Gynecology and Obstetrics stage, grade and histological subtype of cancer,
preoperative and posttreatment CA-125 level, presence or absence of residual disease after
cytoreductive surgery and on postchemotherapy computed tomography scan, and time to
progression and death. The strongest predictors of relapse were included into an algorithm,
the Risk of Ovarian Cancer Relapse (ROVAR) score.

Results: Three hundred fifty-four cases of OC were analyzed to generate the ROVAR
score. Factors selected were preoperative serum CA-125, International Federation of
Gynecology and Obstetrics stage and grade of cancer, and presence of residual disease at
posttreatment computed tomography scan. In the validation data set, the ROVAR score had a
sensitivity and specificity of 94% and 61%, respectively. The concordance index for the
validation data set was 0.91 (95% confidence interval, 0.85-0.96). The score allows patient
stratification into low (<0.33), intermediate (0.34-0.67), and high (>0.67) probability of
relapse.

Conclusions: The ROVAR score stratifies patients according to their risk of relapse
following first-line treatment for OC. This can broadly facilitate the appropriate tailoring of
posttreatment care and support.

Key Words: CA-125, Ovarian cancer, Prognostic factors, Relapse, ROVAR

Received August 26, 2014, and in revised form October 23, 2014.
Accepted for publication December 7, 2014.

[Int J Gynecol Cancer 2015;25: 416–422]
E
ey year, 6500 women in the United Kingdom are diagnosed with ovarian cancer (OC), which is the fifth most common cause of cancer death in the female population.\footnote{1,2} The lethality of OC is attributable to many factors, from its often advanced stage at presentation to its high risk of relapse. Currently, there are no reliable means of predicting an individual’s risk of relapse after first-line treatment for OC. This limits clinicians’ ability to provide risk-appropriate surveillance follow-up and survivorship support. As a result, OC patients of all International Federation of Gynecology and Obstetrics (FIGO) stages are required to adhere to a generic but non-evidence-based outpatient follow-up schedule. At most centers, this comprises 3 monthly follow-ups for 2 years, 6 monthly for 2 years, and yearly thereafter for up to 5 or 10 years.\footnote{3,4} Lack of ability to predict relapse also precludes the provision of tailored survivorship support, which is consequently based around current symptom concerns rather than evaluation of long-term effects of treatment on, for example, cardiovascular or bone health. This is appropriate for the majority of patients whose survival prognosis is less than 5 years but inadequate for those who have been cured by treatment or can expect a long relapse-free interval. However, the ability to tailor risk-appropriate survivorship support and follow-up care is hampered by difficulties in reliability predicting risk of relapse in this setting.

Although FIGO stage at diagnosis is a clinically relevant means of predicting 5-year survival and thus providing guidance in the allocation of upfront treatment,\footnote{3–5} it is an unreliable predictor of relapse risk. To improve on FIGO stage alone, other statistical tools (4 nomograms and 2 scores) have added age, performance status, molecular/genetic and blood markers, presence of comorbidities, tumor grade and histology, residual tumor volume after surgery, the time interval between surgery and chemotherapy, and the presence or absence of ascites.\footnote{6–10} Although proven superior to FIGO staging at predicting 5-year survival from OC, these algorithms are again unsuitable for predicting relapse. Having observed that the presence of residual disease and/or persistent elevation in serum CA-125 on completion of postoperative chemotherapy correlated with risk of early relapse in our patients, we aimed to develop our own risk-of-relapse prognostic algorithm using clinicopathological information obtained before and after first-line treatment. Our purpose was to discriminate patients with a low risk of recurrence who would most benefit from long-term treatment effects assessment (and less frequent follow-up surveillance) from those for whom a symptom-focused survivorship support package and more frequent or responsive follow-up schedule were more appropriate.

**MATERIALS AND METHODS**

Patients were selected for inclusion in our study from our preexisting OC database composed of 500 women with a diagnosis of OC (fallopian tube, ovarian, and primary peritoneal cancers were all defined here as OC) and treated between 2000 and 2010 at Imperial College Healthcare NHS Trust (United Kingdom). The database had been collected prospectively from 2007 until 2010 and retrospectively from 2000 to 2007. Only patients who had received a combination of surgical cytoreduction and platinum-based chemotherapy were included in our study. Additional information from patients listed on the database was collected from the hospital electronic database and medical notes. Patients with missing information, a diagnosis of borderline OC, or those who did not undergo surgical treatment and/or chemotherapy postsurgery were excluded from analysis.

The patient characteristics we evaluated were preoperative serum CA-125 level, FIGO tumor stage evaluated as a categorical variable (levels 1, 2, 3, and 4 corresponding to FIGO stages I to IV), grade (1, 2, and 3 corresponding to histopathologic grades 1 to 3), tumor histological subtype (serous, endometrioid, mucinous, clear cell, and “others”), residual disease on postchemotherapy computed tomography (CT) scan (present or absent), and residual disease after cytoreductive surgery (“nonoptimal cytoreduction” meaning gross residual disease of greater than 1 cm in diameter, “optimal cytoreduction” when the residual disease was less than 1 cm, and “total cytoreduction” when there was no residual disease). Patients were defined as “CA-125 expressers” when the CA-125 at diagnosis was 200 U/mL or greater or “nonexpressers” if the CA-125 level was less than 200 U/mL at diagnosis. For each patient, the time to first relapse after first-line chemotherapy treatment and the time of death were recorded.

Logistic regression analyses were used to determine which prognostic factors were the best predictors of relapse. The full data set was split at random into a fitting data set (n = 254) and a validation data set (n = 100). A logistic regression model was applied to the fitting data set, and prognostic factors were selected using forward, backward, and stepwise selection procedures and $P = 0.05$ for inclusion in the model. Random splitting of the data set into fitting and validation sets was repeated 20 times, and the variables selected in each random partition were recorded. The final selected prognostic model included only variables found to be statistically significant by each selection method (forward/backward/stepwise) in a majority of the random partitions. To confirm this variable selection, a “best subsets” regression algorithm was also applied to the full data set to confirm the selection of covariates in the final model, and this procedure resulted in the selection of the same 4 covariates as the “best model” (using the Bayes information criterion). Performance of the model was further assessed using the “leave-one-out” cross-validation technique on the full data set. For final estimates of sensitivity and specificity in a validation data set, a final random partition of the data was then made, and these variables were included in a logistic regression model applied to the fitting data set only. For comparison with the results derived from the validation data set, the model was also applied to the full data set and performance statistics calculated. Finally, the Hosmer-Lemeshow test\footnote{11} was used to assess goodness of fit of the model. The logistic regression model (Table 2) was used to generate a multifactorial Risk of Ovarian Cancer Relapse (ROVAR) score to predict probability of relapse. In a final exploratory analysis, we also applied a “machine learning” (boosting) algorithm to a series of random partitions of the data and estimated the sensitivity and specificity.
of the resulting classifications for comparison with the logistic regression models.12

**RESULTS**

Of the 500 patients listed on the database, 354 met study eligibility criteria. The remaining 146 (29%) were excluded because of the missing information. The characteristics of these 354 patients (termed here the “ROVAR data set”) are summarized in Table 1. All had a diagnosis of OC, fallopian tube, or primary peritoneal cancer. The histological subtypes were serous, mucinous, clear cell, endometrioid adenocarcinoma, carcinosarcoma, or “others.” Included into “others” were transitional cell tumors (Brenner and non-Brenner types), squamous cell carcinoma, mixed epithelial tumors, sex cord gonadal stromal tumors, undifferentiated, and unclassified.

All patients in the ROVAR data set had undergone cytoreductive surgery aimed at removing all visible residual

---

**TABLE 1. Patient characteristics of the study population**

<table>
<thead>
<tr>
<th>Characteristics</th>
<th>n</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Stage</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>IA</td>
<td>2</td>
<td>21</td>
</tr>
<tr>
<td>IB</td>
<td>6</td>
<td>8.8</td>
</tr>
<tr>
<td>IC</td>
<td>65</td>
<td>20</td>
</tr>
<tr>
<td>II</td>
<td>32</td>
<td>50.3</td>
</tr>
<tr>
<td>III</td>
<td>178</td>
<td>29</td>
</tr>
<tr>
<td>IV</td>
<td>71</td>
<td>20</td>
</tr>
<tr>
<td><strong>Grade</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>23</td>
<td>6.5</td>
</tr>
<tr>
<td>2</td>
<td>101</td>
<td>28.5</td>
</tr>
<tr>
<td>3</td>
<td>230</td>
<td>65</td>
</tr>
<tr>
<td><strong>Histological type</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Serous tumors</td>
<td>234</td>
<td>63.3</td>
</tr>
<tr>
<td>Mucinous tumors</td>
<td>15</td>
<td>4.2</td>
</tr>
<tr>
<td>Endometrioid tumors</td>
<td>39</td>
<td>11</td>
</tr>
<tr>
<td>Mixed Mullerian malignant tumor</td>
<td>13</td>
<td>3.7</td>
</tr>
<tr>
<td>Clear cell tumors</td>
<td>33</td>
<td>9.3</td>
</tr>
<tr>
<td>Other*</td>
<td>20</td>
<td>8.5</td>
</tr>
<tr>
<td><strong>Surgical outcome</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nonoptimal (≥1 cm)</td>
<td>123</td>
<td>34.8</td>
</tr>
<tr>
<td>Optimal (&lt;1 cm)</td>
<td>147</td>
<td>41.5</td>
</tr>
<tr>
<td>Total macroscopic (no disease left)</td>
<td>84</td>
<td>23.7</td>
</tr>
<tr>
<td>CT postchemotherapy</td>
<td></td>
<td></td>
</tr>
<tr>
<td>No residual disease</td>
<td>170</td>
<td>48</td>
</tr>
<tr>
<td>Residual disease</td>
<td>184</td>
<td>52</td>
</tr>
<tr>
<td><strong>First relapse after treatment (surgery and chemotherapy)</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>253</td>
<td>71.5</td>
</tr>
<tr>
<td>No</td>
<td>101</td>
<td>28.5</td>
</tr>
<tr>
<td><strong>CA-125 &gt;200 U/mL at diagnosis</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>251</td>
<td>70.9</td>
</tr>
<tr>
<td>No</td>
<td>103</td>
<td>29.1</td>
</tr>
<tr>
<td><strong>CA-125 posttreatment</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>&gt;35 U/mL</td>
<td>98</td>
<td>27</td>
</tr>
<tr>
<td>&lt;35 U/mL</td>
<td>256</td>
<td>73</td>
</tr>
</tbody>
</table>

*Transitional cell tumors (Brenner and non-Brenner types), squamous cell carcinoma, mixed epithelial tumors, sex cord gonadal stromal tumors, undifferentiated, and unclassified.
disease and for correct surgical staging. Adequate non-fertility-sparing cytoreductive surgery consisted of peritoneal washings, bilateral salpingo-oophorectomy, hysterectomy, multiple peritoneal biopsies of all abdominal fields, at least infracolic omentectomy, and pelvic and para-aortic lymph nodes dissection up to the renal veins. Fertility-sparing surgery in patients with stage IA or stage IC was used only in case of unilateral ovarian involvement and favorable histology (ie, mucinous, serous, endometrioid, or mixed histology and grade 1 or 2); however, this was combined with complete surgical staging, which included a lymphadenectomy to exclude more advanced disease.

A total of 52 (15%) from the data set had undergone neoadjuvant chemotherapy and received chemotherapy prior to definitive surgical cytoreduction, whereas 302 (85%) had adjuvant chemotherapy after upfront surgery. Follow-up lasted more than 5 years (range, 1–120 months) for 275 patients (78%) in the data set. The relapse rate in the full data set was 71% at the time analysis was completed (October 31, 2013).

The final selected model included the following 4 variables: FIGO stage and grade at diagnosis, preoperative CA-125, and residual disease at the posttreatment CT scan. Elevated CA-125 at diagnosis and presence of measurable residual disease after posttreatment CT were no/yes variables. The sensitivity and specificity of the model to predict relapse as applied to the validation data set were 94% and 61%, respectively, and estimated on the full data set were 93% and 65%, respectively. Area under the receiver operating characteristic (ROC) curve for the validation data set was 0.91 (95% confidence interval, 0.85–0.96), and that for the full data set was 0.92 (95% confidence interval, 0.88–0.95). The area under the ROC curve for a model including FIGO stage alone was 0.80 (95% confidence interval, 0.75–0.85). The leave-one-out cross-validation technique gave an adjusted estimate of prediction error for the model of 10.7%. The positive predictive value, which is the proportion of predicted relapses that were true relapses, was 84% in the validation data set. The negative predictive value, which is the proportion of predicted no-relapse patients who were true nonrelapses, was 83% in the validation data set. The Hosmer-Lemeshow test resulted in $P = 0.4074$, indicating no evidence of lack of fit of the model.\(^\text{11}\)

Comparison between the observed and expected responses suggests that the model was well calibrated. The boosting algorithm that we applied had overall mean sensitivity of 86% and mean specificity of 64%, and so performed comparably, but slightly less well than the final logistic regression model that was used to compute the ROVAR score. This is in line with previously reported experience of machine-learning algorithms in this setting.

The ROVAR score derived from the final model is the predicted probability of relapse at 5 years from completion of combined (surgery and chemotherapy) treatment and ranges from 0.052 to 0.992 in the full data set. Using this score, patients can be stratified by risk of relapse using the following cutoff values: 0 to 0.33 = low risk, between 0.34 and 0.67 = intermediate

<table>
<thead>
<tr>
<th>Calculation Steps</th>
<th>ROVAR Score Calculation</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>Initial score</td>
</tr>
<tr>
<td>B</td>
<td>Tumor evident on posttreatment CT scan?</td>
</tr>
<tr>
<td>C</td>
<td>Elevated CA-125 at diagnosis?</td>
</tr>
<tr>
<td>D</td>
<td>FIGO stage</td>
</tr>
<tr>
<td>E</td>
<td>Tumor grade</td>
</tr>
<tr>
<td>F</td>
<td>Subtotal derived</td>
</tr>
</tbody>
</table>

ROVAR risk allocation
- Low risk of relapse 0–0.33
- Intermediate risk of relapse 0.34–0.67
- High risk of relapse $\geq 0.68$
1) Patient with stage IIIC (3), grade III ovarian cancer (3), with no CT-measurable disease after completion of treatment (no) and CA125 of 1500 kU/L (yes) at diagnosis, the following would be applied: Initial step -2.8899 + 0 (no residual disease) + 1.0541 (elevated CA125) + 1.8536 (stage III) + 1.1800 (grade 3) = score of 1.1978. Final step to give ROVAR score = 1/(1+exp(-1.1978)) = 0.76. This gives a high risk of relapse (as it is >0.68)

2) Patient with stage IC (1), grade 3 (3) ovarian cancer with CA125 of 50 kU/L (no) at diagnosis and no measurable disease on post-treatment CT scan (no), the following would be applied: Initial step -2.8899 +0 (no residual disease) + 0 (no elevated CA125) + 0 (stage I) + 1.1800 (grade 3) = subtotal score of -1.7099. Final step to give ROVAR score 1/(1+exp(-1.7099)) = 0.153. This gives a low risk of relapse (as it is less than 0.33)

FIGURE 1. Examples of the ROVAR score.

risk, and 0.68 or greater = high risk. The calculations to generate this score are shown in Table 2, with examples in Figure 1.

DISCUSSION

Knowledge of a patient’s likelihood of relapse from OC is important for the planning and implementation of follow-up care. Until now, the most reliable predictor of a patient’s relapse risk has been their FIGO stage at diagnosis. Although other scores (4 nomograms and 2 scores) have improved on the FIGO stage as predictors of 5-year survival, they mostly considered only advanced (stage III or IV) OC and were based around residual disease after surgery, rather than after completion of combination frontline (ie, surgery plus chemotherapy) treatment summarized in Table 3.13–16 For this reason, these scores cannot be applied to predict risk of relapse.

In light of the above, we developed the ROVAR score with the aim of designing a score that could be applicable across all FIGO stages of OC. We showed that 4 variables contributed most strongly to risk of relapse, and these were (1) tumor stage at diagnosis, (2) tumor grade at diagnosis, (3) CA-125 at diagnosis, and (4) presence of residual disease on CT scan after chemotherapy completion. When we applied the score in our data set to measure the risk of relapse, we obtained an area under the ROC curve of 0.91. When we applied the FIGO stage with the same aim and to the same data set, the ROC curve was 0.80. The ROVAR score was therefore a better predictor of relapse than FIGO stage alone. The sensitivity and specificity of the ROVAR score, estimated on the full data set, were 93% and 65%, respectively. It includes presence of residual disease on completion of first-line (surgery plus chemotherapy) treatment, which we found to be one of the best predictors of relapse. Using the ROVAR score, each patient’s risk of relapse can be stratified into low, intermediate, or high risk of relapse. This can be used to provide an estimate of risk of relapse from the end of treatment (surgery and first-line chemotherapy) onward.

We compared the ROVAR score to other scores and nomograms published (summarized in Table 3). Although specific to OC, all were designed using retrospective data to predict 5-year survival, not relapse-free survival. Most considered only advanced (stage II or IV) OC and were based around residual disease after surgery, without including the presence or absence of disease after chemotherapy treatment.

Our database was representative, we compared patient characteristics with those described in other published data and found them to be similar.17,18 Furthermore, more than 70% of our data set had advanced stage (III or IV) OC consistent with recent statistics.19 In our population, FIGO stage and histological grade were 2 variables significantly associated with risk of relapse. We did not find histological subtype to be an independent prognostic factor. This contradicts published studies showing that clear cell tumor has a poorer prognosis than same-stage serous tumor.20 Perhaps the histological subtype of OC was not a statistically significant independent prognostic factor in our population as the majority had serous histology. Only 17% of our cohort had rarer histological subtypes, such as clear cell tumors, carcinosarcoma, and mucinous tumors. Moreover, we showed tumor grade had prognostic significance, whereas other authors have reported that grade has prognostic significance only for serous tumors (between high- and low-grade disease), but is not of prognostic significance for mucinous, endometrioid, and clear cell tumors.21 This may also be explained by the low percentage of cases in our data set with a diagnosis of nonserous OC.

We found a strong association between CA-125 at diagnosis and risk of relapse, and this was confirmed by logistic regression. Our findings are in accordance with several epidemiologic studies on the association between CA-125 at diagnosis and survival from OC.22,23 In a recent multinational collaborative study involving 940 patients, the authors concluded that, especially in the context of serous OC, the higher the CA-125 at diagnosis, the greater the likelihood of the biomarker being independently associated with impaired disease-free and overall survival.24 We used a similar strategy and defined only those with CA-125 marker levels of greater than 200 U/mL as being true CA-125 “expressors” in view of the high number of serous OC patients included in our analysis. Interestingly, in our cohort, the level of CA-125 after chemotherapy treatment was less significantly associated with risk of relapse. This contradicts 3 small retrospective studies, which reported strong association between the CA-125 level postchemotherapy and prediction of progression-free survival and overall survival.25,26

In our model, we analyzed the residual disease defined on completion of cytoreductive surgery as well as that observed radiologically at treatment completion. We found that the presence or absence of measurable disease seen on CT scan after treatment completion had greater prognostic significance than the amount of measurable disease present on completion of surgery. This result appears reasonable, particularly in advanced stages because, in patients with inoperable extraperitoneal disease, radiological remission can be obtained only after chemotherapy, and the majority of
patients within our cohort had advanced disease. We did not include age or comorbidities in our score as these can be limiting factors to optimal surgery and chemotherapy, themselves significant independent prognostic factors when compared with age and comorbidities.27,28 We also did not include pretreatment or posttreatment performance status as it was not longitudinally documented in the patients’ records during their treatment and has been shown to be unreliable when objectively assessed.29

A limitation of our study is that, although accurate for the majority of patients, the ROVAR score has relatively low specificity, with a prediction error of 10.7%. Thus, a healthcare professional using this score has a 10% chance of incorrectly reassuring a patient that her risk of relapse is low or falsely concerning her by informing her that her risk of relapse is high. This is an important consideration when designing an intervention around risk of relapse. Another potential limitation is that patient treatment during 2000–2010 may have differed to our current standard of care. Generally, however, now as then, standard treatment for primary OC was a combination of surgical cytoreduction and platinum-based chemotherapy. Another limitation is that approximately 30% of our population was excluded from analysis because of missing data.

In summary, the ROVAR score is a four-variable algorithm designed to predict risk of relapse after first-line treatment for OC. Of the 4 variables, FIGO stage, tumor grade, and pretreatment CA-125 level have been previously shown to be prognostic for 5-year survival, and the fourth, the presence of measurable disease on posttreatment CT scan, has

![Table 3](Image)

*Serous, endometrioid, mucinous, clear cell, mixed mesodermal, adenocarcinoma, undifferentiated.
†Serous, endometrioid clear cell, mixed.
‡Serous, mucinous, undifferentiated, miscellaneous.
§Serous, endometrioid, clear cell, mixed.
never before been included within a prognostic score. The ROVAR score will require careful prospective validation in a large sample of OC patients before it is fully implemented. However, by allowing the stratification of patients into low, intermediate, or high risk of relapse once they have completed first-line treatment, the ROVAR score is a useful tool for allocating tailored, risk-appropriate survivorship and follow-up support for patients with OC.

REFERENCES