HHS Public Access

Author manuscript

Oncology. Author manuscript; available in PMC 2017 January 19.

Published in final edited form as:

Oncology. 2016 ; 90(2): 59–68. doi:10.1159/000442710.

Clinical and emergent biomarkers and their relationship to prognosis of ovarian cancer

Aminah Jatoi, M.D.[1,6], **Robert A. Vierkant, M.S.**[2], **Kieran M. Hawthorne, M.S.**[2], **Matthew S. Block, M.D., Ph.D.**[1], **Susan J. Ramus, Ph.D.**[3], **Nicholas B. Larson, Ph.D.**[2], **Brooke L. Fridley, Ph.D.**[4], and **Ellen L. Goode, Ph.D.**[5]

[1]Department of Oncology, Mayo Clinic, Rochester, Minnesota

[2]Department of Biomedical Statistics and Informatics, Mayo Clinic, Rochester, Minnesota

[3]Department of Preventive Medicine, Keck School of Medicine at the University of Southern California

[4]Department of Biostatistics, University of Kansas, Kansas City, Kansas

[5]Department of Health Sciences Research, Mayo Clinic, Rochester, Minnesota

Abstract

Objective—Ovarian cancer is the most lethal gynecological malignancy, but information relevant to prognosis and outcomes remain unknown. Here we used statistical methods to focus specifically on interactions between candidate, prognostic variables.

Methods and Results—Univariate, multivariate, and elastic net modeling of 42 variables were applied to a cohort of 542 ovarian cancer patients with 393 episodes of cancer recurrence/death. In univariate analyses, overexpression of TFF3, MDM2, and p53 were associated with improved recurrence-free survival. In multivariate analyses adjusted for age, histology, stage, grade, ascites, and residual disease, over-expression of PR appeared to provide a protective effect (HR for >50% of cells positive 0.64 [95% CI 0.44–0.94] compared to <1%), and TFF3 showed a nonlinear association. Importantly, we observed no interactions among variables. However, patients with tumors with moderate TFF3 expression were at marginally increased risk of recurrence, and patients with tumors with high expression were at similar to slightly lower risk, compared to those with tumors with no TFF3 expression.

Conclusions—Although no interactions among variables were observed, this study provides important precedent for seeking out interactions between clinical and tumor variables in future studies.

Keywords

ovarian cancer; modeling; prognosis; interactions

[6]Address correspondence to: Aminah Jatoi, M.D.; 200 First Street SW, Rochester, Minnesota 55905.

INTRODUCTION

Ovarian cancer is the most lethal gynecological malignancy [1]. Although previous studies have identified tumor characteristics associated with the development of ovarian cancer, fewer studies have explored the prognostic significance of such factors within the context of multivariate models that include clinical characteristics and potential interactions among all such variables.

To our knowledge, an appraisal of interactions among previously-reported clinical and tumor-related prognostic factors has not yet been undertaken in ovarian cancer. The inability of standard analysis tools to model large sets of predictors has precluded such efforts. Yet such interactions may exist and could conceivably enhance the accuracy of extant prognostic information. As an example, paclitaxel, an agent commonly used to treat ovarian cancer, appears to induce tumor expression of the myeloid differentiation primary response gene (MyD88), which, in turn, appears to lead to chemotherapy resistance [2,3]. Estrogen and progesterone appear to inhibit MyD88, thereby illustrating how clinical and tumor-related factors – chemotherapy, MyD88, and estrogen/progesterone – might together plausibly give rise to interactions that modify clinical outcomes. Determining the existence and extent of such putative interactions could enhance our ability to predict cancer recurrence with greater accuracy.

This study was undertaken with a twofold purpose. First, we examined prognostic associations among tumor markers. These markers included TFF3, WT1, p16, MDM2, and p53 and were chosen because of no well-established interactions. The rationale is that this lack of established interactions would enhance the novelty of findings that may emerge from our second purpose. Second, we examined associations of ovarian cancer survival with pairwise and higher order interactions of clinical factors and tumor expression variables in a series of ovarian cancer cases using contemporary elastic net and classification tree analysis tools in an effort to capture interactions among variables.

METHODS

Overview of Patients and Data Sources

This study was conducted at the Mayo Clinic after Institutional Review Board (IRB) approval. It focused on women 18 years of age or older diagnosed with pathologically-confirmed primary invasive epithelial ovarian cancer, fallopian tube cancer, or primary peritoneal cancer. Because these three malignancies are treated in a uniform manner, herein, we refer to all three as “epithelial ovarian cancer.”

Data were acquired from multiple sources: 1) patient-completed risk factor questionnaires; 2) formalin-fixed paraffin-embedded tumor tissue; and 3) patients’ medical records and additional pathological review with abstracted details such as tumor histology, type of surgery, and type of chemotherapy. Information on cancer recurrence was updated from the Mayo Clinic electronic medical record and defined based on date of starting cancer treatment for recurrent cancer. A follow-up, mailed questionnaire also elicited information

from patients about whether their cancer had recurred with subsequent medical record confirmation when available.

Clinical Factors

Clinical factors of interest consisted of patient age at cancer diagnosis, tumor histology, tumor stage, tumor grade, whether or not ascites was present at surgery, and extent of postoperative residual disease. In addition, patient age at menarche, use of oral contraceptives, parity, education level, and smoking status were also explored. All these clinical variables were chosen because previous studies had suggested prognostic relevance [4–6]. Of note, the majority of patients received cancer therapy in keeping with established guidelines with debulking surgery followed by 18 weeks of platinum-based chemotherapy.

Potential Tumor Prognostic Factors

Tumor-related factors were chosen either because a specific variable had previously been demonstrated to have a prognostic association or because of a purported mechanism-based role in ovarian cancer development or tumor growth. These factors include the following: ARID1A, beta 2 defensin, CD8, CD68, DKK1, ER, fibrinogen, FOLR1, gp96, heparan sulfate, heparanase, high-mobility group box 1, HNF1B, hsp60, hsp70, IκBα, IκBβ, MDM2, MMP9, MyD88, p16, p50, p53, p65, phospho-IκBα, phospho-p65, PR, TFF3, TLR4, vimentin, and WT1 [7–10]. It should be noted that the main emphasis of this paper was on the statistical methodology that centered on identifying and understanding variable interactions; for this reason, variables were not derived from well-established pathways with already well-demonstrated interactions between variables.

Tissue microarrays (TMAs) were created from formalin-fixed, paraffin-embedded tumors. Slides were immunostained with primary antibodies that recognized the proteins described above after optimization of staining conditions on positive control tissues. Slides were then scored by two reviewers. Details are available upon request. The strongest protein expression over multiple cores was used for scoring, and discrepancies were resolved by a gynecologic pathologist.

Statistical Methods: Imputation of Missing Values

Prior to analysis, imputation of missing clinical, lifestyle and tumor expression values was performed using the MICE package in R 2.15.0. All variables with at least one missing value were imputed using a regression model that included all other independent variables as predictors. Dichotomous variables (high versus low tumor grade, presence of ascites, ever used oral contraceptives, ever smoked) were imputed using a binary logistic regression model; multilevel nominal characteristics (parity/age at first birth, education, and age at menarche) using unordered, multinomial logistic regression; and ordinal characteristics (extent of residual disease and each of the tumor expression variables) using proportional odds logistic regression. Imputation was performed five times, resulting in five complete datasets.

Statistical Methods: Univariate and Multivariate Analyses

Cox proportional hazards regression was used to calculate survival estimates, with left truncation to account for delayed study enrollment and right censoring at 10 years to minimize competing causes of death. The main effect associations for each of the 42 potential prognostic variables were first examined in univariate analyses. Separate models were fit for each of the five complete imputed data sets, and unified parameter estimates and 95% confidence intervals were calculated with standard multiple imputation methods [11].

Next, a multivariate Cox regression model was fit to examine independent associations of clinical, lifestyle and tumor expression variables with recurrence free survival. The large number of predictors (42 variables with 51 corresponding degrees-of-freedom) relative to the sample size of 542 subjects prohibited the simultaneous inclusion of all variables in one model, and the introduction of multiple imputed data sets dramatically increased the complexity of common variable selection approaches such as stepwise regression. To overcome this, the Group LASSO (GL) regularization models were used to select variables to enter the final multivariate model [12]. Briefly, LASSO is a statistical regularization method that imposes penalty terms to regression-based parameter estimates to prevent over-fitting of a model [13]. These penalty terms serve the function of shrinking the parameter estimates toward zero (the null hypothesis). LASSO differs from ridge regression, another statistical regularization method, in that parameter estimates based on the former can be shrunk completely back to zero whereas estimates from the latter will always be non-zero. This adds a variable selection component to LASSO: estimates that are shrunk all the way to zero are effectively removed from the model. Cross-validation is used to tune and optimize the LASSO penalty terms. GL is an extension of LASSO that allows a single penalty term to be simultaneously imposed the n-1 parameter estimates that compose a given n-level nominal categorical variable, thus preserving the internal structure of that variable. Cox regression-based GL was carried out on each of the five imputed complete data sets using the analysis tool SGL. Variables that were retained in at least four of the five imputed data sets were considered predictive of recurrence free or overall survival and were included in one final traditional (that is, non-LASSO-based) Cox regression model to aid with interpretation of the parameter estimates. The following six variables were also included in this final model due to their commonly recognized effect on prognosis: age, stage, grade, histology, presence of ascites and extent of residual disease. Unified parameter estimates and corresponding 95% confidence intervals were again calculated using standard multiple imputation analysis methods.

Statistical Methods: Interaction Testing

Pairwise interactions were examined for potential prognostic variables with respect to recurrence free survival to assess the existence of effect modification. Interactions for each of the 861 possible variable pairs (42 choose 2) were examined separately by fitting main effects and interactions and testing the statistical significance of the interactions. Due to sparsity of data for some histologic categories, mucinous, clear cell, and other histologies were combined into one group.

Elastic net survival regression methods were used to simultaneously examine the 861 paired interaction effects [14]. Briefly, elastic net is a family of regularization models that applies penalty functions that are a linear combination of LASSO and ridge regression. These functions can range from 0 to 1, with those closer to 1 behaving more like LASSO in that fewer variables are retained in the final model, and those closer to 0 behaving more like ridge regression in that more variables are retained. A series of elastic net models were run on each of the five imputed data sets, varying the penalty function from 0.1 to 1.0. All 42 main effect terms were initially forced into the model and shrunk only the interaction terms, in an effort to retain the structural hierarchy of interaction models. A secondary two-step approach was applied because of concerns that simultaneous inclusion of all prognostic variables would result in an over-parameterized model. A model that allowed all main effects and interactions to be shrunk or eliminated was fit. All retained as either a main effect or interaction term were included as main effects in a subsequent model. This initial approach was then applied to a smaller model with a forced reduction in the reduced number of main effect terms and with allowance for their corresponding pairwise interactions to be shrunk. Analyses were carried out using the glmnet package [15,16].

Finally, survival-based classification trees were used to examine the possibility of higher order interaction terms [17]. This approach first examines all possible dichotomous partitions of the data based on the 42 prognostic variables and chooses the one which best discriminates survival. These two corresponding partitions are then examined and split into sub-partitions. The procedure continues recursively until a full tree is built. Cross-validation was used to prune the tree by determining the number of partitions that minimizes reclassification error.

RESULTS

Distribution of Clinical, Lifestyle, and Tumor Expression Factors

A total of 542 patients were included in the analysis, with 393 events, which included 260 episodes of cancer recurrence and an additional 133 deaths. Median recurrence-free survival was 2.1 years. Median follow up time of those patients still alive at the time of this report was 7.4 years. Median follow up time of those patients who had recurrent disease or who were dead was 1.4 years. Patient baseline characteristics and the 42 candidate prognostic variables are listed in Table 1.

Univariate and Multivariate Analyses

In univariate analyses, after Bonferroni correction, patient age at cancer diagnosis, histology, cancer stage, cancer grade, presence of ascites, and presence of postoperative residual disease showed a statistically significant association with worse recurrence-free survival (Table 1). No lifestyle factors, including oral contraceptive use, were associated with prognosis at $p<0.05$. Overexpression of WT1 and p16 in tumors were also associated with worse recurrence-free survival. In contrast, overexpression of TFF3, MDM2, and p53 were associated with improved recurrence-free survival. Results using imputed data were similar (Supplemental Table 2).

In Cox regression multivariate analyses, patient age at cancer diagnosis, tumor grade, and presence of ascites were marginally associated with prognosis, while tumor histology was not significantly associated with cancer recurrence (Table 2). Over-expression of PR appeared to be a protective prognostic variable in multivariate Cox regression models (HR for >50% of cells positive 0.64 [95% CI 0.44–0.94] compared to <1% positive). Associations with TFF3 were more complex: women with moderate expression were at marginally increased risk of recurrence, and women with high expression were at similar to slightly lower risk, compared to women with no expression.

Potential Interactions of Variables

In traditional analyses of all possible pairwise interactions, pairs ($p<0.001$) that were associated with recurrence-free survival included histology with extent of postoperative residual disease, parity/age at first birth with extent of postoperative residual disease, and histology with MDM2 (Table 3). These findings suggest that non-serous histology is most protective among women with no residual disease or with positive MDM2 expression, and that increased parity may be detrimental among women with no residual disease. However, none of these prognostic interactions remained statistically significant after accounting for multiple testing.

Using elastic net methodology, we observed null results with no single interaction remaining in the final model for more than one of the five imputed datasets for any of the penalty levels tested. Both approaches (shrinking only interaction terms or the two-step approach) resulted in the same conclusions. Cross-validation analyses with a survival-based classification tree suggested an optimal tree size of one node; no single data split, let alone combination of splits, dramatically improved discrimination of recurrence-free survival.

DISCUSSION

This study examined the prognostic significance of numerous clinical and tumor protein expression variables and their potential interactions in ovarian cancer patients. As expected, we observed that cancer stage, ascites, postoperative residual disease, and overexpression of PR were associated with worse recurrence-free survival. In addition, this study revealed two unexpected findings. First, using both traditional and contemporary analysis methods, we found limited pair-wise or higher order interactions with respect to ovarian cancer recurrence-free survival among the 42 clinical, lifestyle, and tumor expression factors examined. After adjusting for multiple testing, using the elastic net method, and confirming findings by means of a survival-based classification tree, we found no evidence of interactions between any of these variables. Some interactions may merit consideration in larger studies. Nonetheless, this study provides important precedent for assessing interactions among multiple, diverse groups of variables in an effort to better understand ovarian cancer prognosis.

Interestingly, the elastic net analysis methodology used in this study is timely in view of numerous ongoing studies that seek to examine interactions among an expansive number of variables – particularly genomic factors – and their associations with respect to cancer risk. Although the current study did not include genomic data, other investigations may use

similar methodology when analyzing large data sets to demonstrate the complexity of genetic factors or epistasis. Future research might use similar methodology to examine putative interactions among a variety of factors associated with recurrence-free survival, including genomic ones. Thus, we view our findings as an invitation to further explore potential interactions.

The second important finding of our paper centers on TFF3 expression – a marker that appears to have protective effects on epithelial cells and that has been relatively understudied in ovarian cancer – was associated with a variable but statistically significant risk of cancer recurrence based on extent of expression. This protein has begun to receive increasing attention in gastrointestinal cancers, as a potential marker for gastric cancer screening and therefore may merit further study in ovarian cancer [18,19]. Further validation of this incidental finding in conjunction with confirmation of effect sizes might lead to further understanding of the clinical relevance of this preliminary observation.

This study has both strengths and limitations. One limitation is that although our sample size appears robust by some standards, the number of variables introduced into our models ultimately led to diminished power as a result of the multiple testing burden. Despite this limitation, our efforts to champion the elastic net methodology with the inclusion of diverse groups of variables may prompt others to undertake a similar approach within larger multi-institutional data sets. Another limitation is that for certain relevant variables, such as BRCA1 and 2 mutation status, we had too little data to incorporate into our models. In terms of strengths, we believe again that the application of the elastic net methodology to such a diverse group variables with the goal of better understanding the prognosis of ovarian cancer patients is unique and merits more widespread use.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

This work was funded by the Minnesota Ovarian Cancer Alliance, R01-CA122443, P30-CA15083, and P50-CA136393.

References

1. Siegel R, Ma J, Zou Z, Jemal A. Cancer statistcs 2014. CA Cancer J Clin. 2014; 64:9–29. [PubMed: 24399786]
2. Kelly MG, Alvero AB, Chen R, et al. TLR-4 signaling promotes tumor growth and paclitaxel chemoresistance in ovarian cancer. Cancer Res. 2006; 66:3859–68. [PubMed: 16585214]
3. Wang AC, Ma YB, Wu FX, et al. TLR4 induces tumor growth and inhibits paclitaxel activity in MyD88-positive human ovarian carcinoma. Oncol Lett. 2014; 7:871–877. [PubMed: 24527095]
4. Gram IT, Lukanova A, Brill I, et al. Cigarette smoking and risk of histologic subtypes of epithelial ovarian cancer in the EPIC cohort. Int J Cancer. 2012; 130:2204–10. [PubMed: 21678398]
5. Kjaerbye-Thygesen A, Frederiksen K, Hogdall EV, et al. Do risk factors for epithelial ovarian cancer have an impact on prognosis? Focus on previous pelvic surgery and reproductive variables. Eur J Gynaecol Oncol. 2006; 27:467–72. [PubMed: 17139980]

6. Wahner Hendrickson AE, Hawthorne KM, Goode EL, et al. Assessment of published models and prognostic variables in epithelial ovarian cancer at Mayo Clinic. Gynecologic Oncology. 2015; 137:77–85. [PubMed: 25620544]
7. Kobel M, Kalloger SE, Lee S, et al. Biomarker-based ovarian carcinoma typing: a histologic investigation in the ovarian tumor tissue analysis consortium. Cancer Epidemiol Biomarkers Prev. 2013; 22:1677–86. [PubMed: 23880734]
8. Köbel M, Madore J, Ramus SJ, Clarke BA, Pharoah PDP, Deen S, et al. Evidence for a time-dependent association between FOLR1 expression and survival from ovarian carcinoma: Implications for clinical testing. An Ovarian Tumor Tissue Analysis consortium study. Br J Cancer. in press.
9. Sieh W, Kobel M, Longacre TA, et al. Hormone receptor expression and ovarian cancer survival: an Ovarian Tumor Tissue Analysis consortium study. Lancet Oncol. 2013; 14:853–62. [PubMed: 23845225]
10. Shen H, Fridley BL, Song H, et al. Epigenetic analysis leads to identification or HNF1B as a subtype-specific susceptibility gene for ovarian cancer. Nat Commun. 2013; 4:1628. [PubMed: 23535649]
11. Rubin DB, Schenker N. Multiple imputation for interval estimation from simple random samples with ignorable nonresponse. J Am Statist Assoc. 1986; 81:366–374.
12. Simon N, Friedman J, Hastie T, Tibshirani R. A sparse-group lasso. Journal of Computational and Graphical Statistics. 2013; 22:231–245.
13. Tibshirani R. Regression shrinkage and selection via the lasso: a retrospective. J R Statist Soc B. 2011; 73:273–282.
14. Zou H, Hastie T. Regularization and Variable Selection via the Elastic Net. Journal of the Royal Statistical Society, Series B. 2005; 67:301–320.
15. Friedman J, Hastie T, Tibshirani R. Regularization paths for generalized linear models via coordinate descent. J Statistical Software. 2010; 33:1–22.
16. Simon N, Friedman J, Hastie T, Tibshirani R. Regulation paths for Cox's proportional hazards model via coordinate descent. J Statistical Software. 2011; 39:1–13.
17. Therneau, TM.; Atkinson, EJ. Introduction to Recursive Partitioning Using the RPART Routines. Department of Health Science Research, Mayo Clinic; Rochester, Minnesota: 1997. Technical Report Series 61
18. Ding A, Zhao W, shi X, et al. Impact of NPM, TFF3 and TACC1 on the prognosis of patients with primary gastric cancer. PLoS One. 2013; 8:e82136. [PubMed: 24358147]
19. Huang Z, Zhang X, Lu H, et al. Serum trefoil factor is a promising non-invasive biomarker for gastric cancer screening: a monocentric cohort study in China. BMC Gastroenterol. 2014; 14:74. [PubMed: 24720760]

Table 1

Distributions of Patient Characteristics and Univariate Associations with Ovarian Cancer

Variable	Level	n	%	Person-years	HR (95% CI)	P-value
Clinical Factors						
Age at cancer diagnosis (years)	<50	96	65.63	373	ref	<.01
	50–59	139	66.19	518	1.02 (0.74–1.41)	
	60–69	149	71.14	509	1.15 (0.84–1.57)	
	70+	158	83.54	422	1.62 (1.20–2.19)	
Tumor histology	Serous	374	83.16	996	ref	<.01
	Endometrioid	91	41.76	499	0.30 (0.21–0.42)	
	Clear Cell	35	54.28	136	0.50 (0.32–0.80)	
	Mucinous	18	33.33	110	0.22 (0.10–0.49)	
	Other	24	79.17	82	0.80 (0.51–1.28)	
Cancer stage	1	87	27.59	568	ref	<.01
	2	38	47.37	211	1.88 (1.02–3.47)	
	3	328	81.40	886	5.85 (3.84–8.91)	
	4	89	94.38	158	10.10 (6.38–16.0)	
Grade	Low	91	37.36	534	ref	<.01
	High	449	79.73	1276	3.57 (2.51–5.09)	
Ascites present at surgery	No	164	48.78	807	ref	<.01
	Yes	275	83.64	695	2.85 (2.21–3.69)	
Postoperative residual disease	None	238	51.26	1137	ref	<.01
	≤1cm	152	92.11	308	3.52 (2.74–4.52)	
	>1cm	70	91.43	107	4.80 (3.51–6.55)	
Lifestyle/Reproductive Factors						
Age at menarche (years)	≤11	67	70.15	234	ref	0.62
	12	114	70.18	402	1.01 (0.71–1.45)	

Variable	Level	n	%	Person-years	HR (95% CI)	P-value
	13	122	66.39	446	0.89 (0.62–1.28)	
	14+	113	74.34	370	1.10 (0.77–1.57)	
Oral contraceptive use	Never	224	73.21	726	ref	0.20
	Ever	278	70.50	1004	0.87 (0.71–1.08)	
Parity/age at first birth	Nulliparous	90	68.89	344	ref	0.32
	1–2, ≤20 years	41	65.85	156	0.96 (0.61–1.50)	
	1–2, >20 years	156	67.95	539	1.06 (0.77–1.45)	
	3+, ≤20 years	94	73.40	300	1.25 (0.89–1.77)	
	3+, >20 years	124	79.03	396	1.30 (0.94–1.78)	
Education	Less than high school	32	68.75	119	ref	0.70
	Finished high school	160	73.75	549	1.18 (0.75–1.87)	
	More than high school	280	70.71	958	1.10 (0.71–1.71)	
Smoking status	Never	292	71.23	1041	ref	0.31
	Ever	213	72.30	693	1.11 (0.90–1.37)	
Tumor Protein Expression						
ARID1A	<1%–50% positive	55	58.18	252	ref	0.02
	>50% positive	449	74.16	1457	1.56 (1.08–2.24)	
Beta 2 defensin	None	20	85.00	48	ref	0.08
	Weak	185	77.30	651	0.68 (0.41–1.13)	
	Moderate/Strong	264	69.32	877	0.59 (0.36–0.97)	
CD68	None/Weak	138	68.12	523	ref	0.14
	Moderate	232	75.43	760	1.23 (0.96–1.58)	
	Strong	87	77.01	259	1.34 (0.98–1.84)	
CD8	None/Weak	317	73.50	1043	ref	0.10
	Moderate	117	74.36	388	0.98 (0.77–1.25)	
	Strong	42	59.52	185	0.63 (0.42–0.97)	

Oncology. Author manuscript; available in PMC 2017 January 19.

Variable	Level	n	%	Person-years	HR (95% CI)	P-value
DKK1	Negative (<1%)	368	75.00	1169	ref	<.01
	1–50% positive	98	75.51	331	0.99 (0.76–1.28)	
	>50% positive	24	41.67	135	0.34 (0.17–0.66)	
ER	Negative (<1%)	123	73.17	391	ref	0.04
	1–50% positive	91	84.62	267	1.21 (0.89–1.65)	
	>50% positive	286	68.88	1034	0.87 (0.67–1.12)	
Fibrinogen	None	425	73.65	1423	ref	0.89
	Weak/Moderate/Strong	38	73.68	119	1.03 (0.70–1.51)	
FOLR1	Negative (<1%)	88	59.09	373	ref	0.04
	Weak cyto, membr	189	73.02	670	1.29 (0.94–1.78)	
	Moderate/strong cyto, heterogeneous memb	159	79.25	461	1.59 (1.15–2.21)	
	Full diffuse cyto, consistent strong memb	69	78.26	192	1.45 (0.98–2.14)	
gp96	None	16	93.75	38	ref	0.06
	Weak	153	76.47	501	0.69 (0.41–1.19)	
	Moderate/Strong	320	68.75	1135	0.58 (0.35–0.98)	
Heparan sulfate	None/Weak	198	67.17	707	ref	0.13
	Moderate/Strong	268	77.24	864	1.18 (0.95–1.47)	
Heparanase	None	170	77.06	553	ref	0.23
	Weak	239	70.71	810	0.82 (0.65–1.04)	
	Moderate/Strong	50	72.00	181	0.83 (0.57–1.20)	
High-mobility group box 1	None	154	77.92	519	ref	0.10
	Weak	265	68.30	920	0.80 (0.64–1.01)	
	Moderate/Strong	42	83.33	127	1.07 (0.73–1.55)	
HNF1B	Negative (<1%)	384	76.82	1178	ref	<.01
	1–50% positive	61	62.30	252	0.70 (0.50–0.98)	
	>50% positive	55	58.18	245	0.61 (0.42–0.87)	
hsp60	Weak	122	67.21	483	ref	0.09

Oncology. Author manuscript; available in PMC 2017 January 19.

Variable	Level	n	%	Person-years	HR (95% CI)	P-value
	Moderate/Strong	363	75.21	1143	1.24 (0.96–1.59)	
hsp70	Low	164	64.63	638	ref	<.01
	Medium	181	81.22	516	1.51 (1.18–1.94)	
	High	158	71.52	542	1.12 (0.86–1.46)	
IkappaBalpha	None	128	78.91	423	ref	0.24
	Weak	248	72.18	852	0.82 (0.65–1.05)	
	Moderate/Strong	106	68.87	346	0.81 (0.60–1.09)	
IKKB	None	97	77.32	334	ref	0.49
	Weak	268	70.15	935	0.85 (0.65–1.11)	
	Moderate/Strong	109	72.48	346	0.92 (0.67–1.26)	
MDM2	Negative (<1%)	340	79.71	982	ref	<.01
	≥1% positive	152	58.55	658	0.58 (0.45–0.74)	
MMP9	None	177	79.66	579	ref	0.07
	Weak	259	68.73	888	0.77 (0.62–0.97)	
	Moderate/Strong	28	75.00	103	0.76 (0.48–1.20)	
MyD88	None/Weak	96	59.38	405	ref	0.01
	Moderate	209	73.68	691	1.46 (1.07–1.98)	
	Strong	159	78.62	480	1.62 (1.18–2.23)	
p16	Negative (<1%)	25	56.00	103	ref	<.01
	1–50% positive	242	67.36	951	1.27 (0.73–2.19)	
	>50% positive	232	80.17	630	1.89 (1.09–3.25)	
p50	None	77	84.42	238	ref	0.02
	Weak	240	75.00	779	0.84 (0.64–1.12)	
	Moderate/Strong	142	64.79	525	0.64 (0.47–0.89)	
p53	Negative (<1%)	99	78.79	289	ref	<.01
	1–50% positive	166	54.82	761	0.52 (0.38–0.70)	
	>50% positive	230	84.78	606	1.12 (0.86–1.46)	

Oncology. Author manuscript; available in PMC 2017 January 19.

Variable	Level	n	%	Person-years	HR (95% CI)	P-value
p65	None/Weak	40	65.00	163	ref	0.15
	Moderate/Strong	452	73.45	1491	1.35 (0.90–2.03)	
Phospho-IkappaBalpha	None	216	76.39	729	ref	0.19
	Weak/Moderate/Strong	261	68.97	899	0.87 (0.70–1.07)	
Phospho-p65	None	50	80.00	172	ref	0.66
	Weak	297	71.72	1033	0.85 (0.61–1.20)	
	Moderate/Strong	134	72.39	427	0.88 (0.61–1.27)	
PR	Negative (<1%)	297	79.80	827	ref	<.01
	1–50% positive	116	75.00	417	0.79 (0.61–1.00)	
	>50% positive	77	48.05	383	0.40 (0.28–0.56)	
TFF3	Negative (<1%)	327	81.04	951	ref	<.01
	1–50% positive	105	70.48	364	0.82 (0.63–1.06)	
	>50% positive	62	40.32	347	0.32 (0.21–0.48)	
TLR4	None/Weak	164	65.85	633	ref	<.01
	Moderate	215	74.42	731	1.23 (0.96–1.57)	
	Strong	90	80.00	236	1.60 (1.19–2.16)	
Vimentin	Negative (<1%)	354	78.25	1081	ref	<.01
	1–50% positive	93	61.29	384	0.63 (0.47–0.83)	
	>50% positive	44	65.91	166	0.65 (0.44–0.97)	
WT1	Negative (<1%)	141	51.77	681	ref	<.01
	1–50% positive	79	78.48	269	1.98 (1.40–2.78)	
	>50% positive	283	82.69	736	2.48 (1.89–3.24)	

Abbreviations: HR, hazard ratio; CI, confidence interval; memb., membrane staining; cyto, cytoplasmic staining.

Table 2

Multivariate Associations of Clinical and Tumor Expression Variables with Ovarian Cancer Recurrence-Free Survival

Variable	Level	n	%	Person-years	HR (95% CI)
Clinical Factors					
Age at cancer diagnosis (years)	<50	96	65.63	373	ref
	50–59	139	66.19	518	0.99 (0.71–1.38)
	60–69	149	71.14	509	0.81 (0.58–1.12)
	70+	158	83.54	422	1.31 (0.95–1.79)
Tumor histology	Serous	374	83.16	996	ref
	Endometrioid	91	41.76	499	1.14 (0.44–2.93)
	Clear Cell	35	54.28	136	0.81 (0.54–1.2)
	Mucinous	18	33.33	110	1.14 (0.69–1.89)
	Other	24	79.17	82	1.27 (0.79–2.05)
Cancer stage	1	87	27.59	568	ref
	2	38	47.37	211	1.72 (0.9–3.3)
	3	328	81.40	886	3.18 (1.86–5.44)
	4	89	94.38	158	4.77 (2.69–8.46)
Grade	Low	91	37.36	534	ref
	High	449	79.73	1276	1.28 (0.84–1.96)
Ascites present at surgery	No	164	48.78	807	ref
	Yes	275	83.64	695	1.47 (1.13–1.91)
Postoperative residual disease	None	238	51.26	1137	ref
	≤1cm	152	92.11	308	1.54 (1.16–2.04)
	>1cm	70	91.43	107	2.52 (1.79–3.54)
Tumor Protein Expression					
PR	Negative (<1%)	297	79.80	827	ref
	1–50% positive	116	75.00	417	0.92 (0.71–1.21)

Variable	Level	n	%	Person-years	HR (95% CI)
	>50% positive	77	48.05	383	0.64 (0.44–0.94)
TFF3	Negative (<1%)	327	81.04	951	ref
	1–50% positive	105	70.48	364	1.36 (1.03–1.81)
	>50% positive	62	40.32	347	0.84 (0.53–1.33)

Abbreviations: HR, hazard ratio; CI, confidence interval. Tumor expression variables were chosen using Group LASSO Cox proportional hazards regression analysis. Clinical variables (patient age at cancer diagnosis, histology, stage, grade, ascites, and extent of postoperative residual disease) were included in final multivariate model regardless of LASSO results. Hazard ratios and confidence intervals estimated using Cox regression of multiple imputed data, adjusted for all variables shown.

Table 3

Pairwise Interactions (P<0.001) and Association with Ovarian Cancer Recurrence-Free Survival

		n	%	Person-years	HR (95% CI)	P-value
Residual disease * Histology						0.0002
Residual disease	Histology					
None	Serous	127	67.72	476	ref	
	Endometrioid	61	24.59	393	0.26 (0.15–0.44)	
	Other	50	42.00	267	0.49 (0.31–0.80)	
≤1cm	Serous	133	91.73	274	ref	
	Endometrioid	10	90.00	23	1.01 (0.51–2.00)	
	Other	9	100.00	11	1.70 (0.86–3.35)	
>1cm	Serous	61	91.80	97	ref	
	Endometrioid	5	100.00	3	1.84 (0.66–5.08)	
	Other	4	75.00	7	0.75 (0.23–2.40)	
MDM2 * Histology						0.0006
MDM2	Histology					
Negative (<1%)	Serous	265	84.15	696	ref	
	Endometrioid	41	65.85	154	0.60 (0.40–0.90)	
	Other	34	61.76	132	0.56 (0.36–0.88)	
≥1% positive	Serous	72	83.33	193	ref	
	Endometrioid	44	20.45	302	0.12 (0.06–0.25)	
	Other	36	55.56	163	0.45 (0.27–0.75)	
Residual disease * Parity/Age at first birth						0.0003
Residual disease	Parity/Age at first birth					
None	Nulliparous	41	46.34	222	ref	
	1–2, ≤20 years	21	52.38	103	1.19 (0.57–2.51)	
	1–2, >20 years	81	44.44	397	1.03 (0.59–1.79)	
	3+, ≤20 years	32	53.13	138	1.42 (0.74–2.73)	
	3+, >20 years	49	61.22	220	1.53 (0.86–2.73)	

		n	%	Person-years	HR (95% CI)	P-value
≤1cm	Nulliparous	25	88.00	48	ref	
	1–2, ≤20 years	11	72.73	35	0.51 (0.23–1.14)	
	1–2, >20 years	29	93.10	57	1.00 (0.57–1.75)	
	3+, ≤20 years	32	96.88	61	1.04 (0.60–1.80)	
	3+, >20 years	42	92.86	88	0.93 (0.55–1.57)	
>1cm	Nulliparous	10	100.00	12	ref	
	1–2, ≤20 years	4	100.00	2	4.15 (1.28–13.43)	
	1–2, >20 years	18	100.00	17	1.39 (0.63–3.04)	
	3+, ≤20 years	18	66.67	56	0.34 (0.14–0.78)	
	3+, >20 years	13	100.00	13	1.03 (0.44–2.39)	

Abbreviations: HR, hazard ratio; CI, confidence interval.