

Development and validation of a novel 11-gene prognostic model for serous ovarian carcinomas based on lipid metabolism expression profile

Mingjun Zheng, Heather Mullikin, Anna Hester, Bastian Czogalla, Helene Heidegger, Theresa Vilsmaier, Aurelia Vattai, Anca Chelariu-Raicu, Udo Jeschke, Fabian Trillsch, Sven Mahner, Till Kaltofen

Angaben zur Veröffentlichung / Publication details:

Zheng, Mingjun, Heather Mullikin, Anna Hester, Bastian Czogalla, Helene Heidegger, Theresa Vilsmaier, Aurelia Vattai, et al. 2020. "Development and validation of a novel 11-gene prognostic model for serous ovarian carcinomas based on lipid metabolism expression profile." *International Journal of Molecular Sciences* 21 (23): 9169.
<https://doi.org/10.3390/ijms21239169>.



Article

Development and Validation of a Novel 11-Gene Prognostic Model for Serous Ovarian Carcinomas Based on Lipid Metabolism Expression Profile

Mingjun Zheng ¹, Heather Mullikin ¹, Anna Hester ¹, Bastian Czogalla ¹ , Helene Heidegger ¹, Theresa Vilsmaier ¹, Aurelia Vattai ¹, Anca Chelariu-Raicu ¹, Udo Jeschke ^{1,2} , Fabian Trillsch ¹ , Sven Mahner ¹ and Till Kaltofen ^{1,*}

¹ Department of Obstetrics and Gynecology, University Hospital, LMU Munich, Maistrasse 11, 80337 Munich, Germany; Mingjun.Zheng@med.uni-muenchen.de (M.Z.); heather.mullikin@gmail.com (H.M.); Anna.Hester@med.uni-muenchen.de (A.H.); Bastian.Czogalla@med.uni-muenchen.de (B.C.); Helene.Heidegger@med.uni-muenchen.de (H.H.); Theresa.Vilsmaier@med.uni-muenchen.de (T.V.); Aurelia.Vattai@med.uni-muenchen.de (A.V.); Anca.Chelariuraicu@med.uni-muenchen.de (A.C.-R.); Udo.Jeschke@med.uni-muenchen.de (U.J.); Fabian.Trillsch@med.uni-muenchen.de (F.T.); Sven.Mahner@med.uni-muenchen.de (S.M.)

² Department of Obstetrics and Gynecology, University Hospital Augsburg, Stenglinstrasse 2, 86156 Augsburg, Germany

* Correspondence: Till.Kaltofen@med.uni-muenchen.de

Received: 19 September 2020; Accepted: 27 November 2020; Published: 1 December 2020



Abstract: (1) Background: Biomarkers might play a significant role in predicting the clinical outcomes of patients with ovarian cancer. By analyzing lipid metabolism genes, future perspectives may be uncovered; (2) Methods: RNA-seq data for serous ovarian cancer were downloaded from The Cancer Genome Atlas and Gene Expression Omnibus databases. The non-negative matrix factorization package in programming language R was used to classify molecular subtypes of lipid metabolism genes and the limma package in R was performed for functional enrichment analysis. Through lasso regression, we constructed a multi-gene prognosis model; (3) Results: Two molecular subtypes were obtained and an 11-gene signature was constructed (PI3, RGS, ADORA3, CH25H, CCDC80, PTGER3, MATK, KLRB1, CCL19, CXCL9 and CXCL10). Our prognostic model shows a good independent prognostic ability in ovarian cancer. In a nomogram, the predictive efficiency was notably superior to that of traditional clinical features. Related to known models in ovarian cancer with a comparable amount of genes, ours has the highest concordance index; (4) Conclusions: We propose an 11-gene signature prognosis prediction model based on lipid metabolism genes in serous ovarian cancer.

Keywords: ovarian neoplasms; lipid metabolism; genes; The Cancer Genome Atlas (TCGA); Gene Expression Omnibus (GEO)

1. Introduction

Epithelial ovarian cancer (EOC) is one of the most lethal gynecological malignancies worldwide [1]. It has a high mortality, constituting 3.3% of all malignant diseases and claiming 5.6% of gynecological cancer-related deaths of women in Germany. [2]. Although the prognosis has been improved to a certain degree by surgical treatment, platinum-based chemotherapy, bevacizumab and poly ADP ribose polymerase inhibitors, the 5-year survival of patients with advanced stage EOC is poor at only 20–30% [3,4]. Therefore, when investigating new therapeutic options it is of clinical importance to identify reliable prognostic markers or models to more accurately study their role in the occurrence and development of EOC.

To date, many driver genes have already been identified. For instance, BRCA1, BRCA2, p53, KRAS, PIK3CA, CTNNB1 and PTEN can be implicated in the development of EOC [5]. EPOR is known for its biological effect on tumor growth [6], TTC30A and LRRC8D regulate the expression of host proteins [7], Tfcp2l1 may be involved in the differentiation of ovarian tumor stem cells [8] and the expression of FOXM1 is correlated with chemotherapy resistance and poor prognosis in patients with non-serous epithelial EOC [9].

Recent studies suggest that lipid metabolism disorders may represent important metabolic markers for cancer cells in general. Metabolic reprogramming, including changes in lipid metabolism, can occur in tumor cells and the tumor microenvironment and has impact on the growth, proliferation, invasion and metastasis of cancer cells [10–13]. Braicu et al. [14] conducted a comprehensive lipidomics analysis on the serum samples of 147 EOC patients and 98 control subjects with benign ovarian tumors or non-tumorous diseases. They revealed that a variety of lipid molecules could act as prognostic markers for EOC, due to their superior efficacy compared to CA125. Furthermore, EOC cells are known to be more aggressive after reprogramming lipid metabolism in an ascites microenvironment. Targeting the signal transduction axis of lipid metabolism can effectively prevent peritoneal metastasis of EOC under experimental circumstances [15]. Niemi et al. [10] showed that changes in lipid metabolism can occur in various stages of EOC and can become intensified in a consistent pattern with advanced cancer stages.

However, the role and the prognostic value of genes related to lipid metabolism in EOC remain to be clarified, since a lack of large-scale EOC sample populations, especially for validation sets, constrains the reliability and validity of previous research results. However, in an era of big data, the emergence of genome-sequencing technologies and data [16–19] may help in tumor diagnosis and prognosis prediction [20].

Accordingly, we collected genes related to lipid metabolism and constructed molecular subtypes of serous EOC based on lipid metabolism genes according to The Cancer Genome Atlas (TCGA) and the Gene Expression Omnibus (GEO) databases. After this, we established an 11-gene signature prognosis prediction model to validate its performance in a large panel of tumors.

2. Results

2.1. Molecular Subtypes of Lipid Metabolism-Related Genes

2.1.1. Identification of Two Molecular Subtypes

After preprocessing, a total of 751 lipid metabolism genes from our serous EOC samples qualified for subsequent analysis (Supplementary Table S1). We then conducted a univariable Cox analysis using the *coxph* function to obtain 64 prognostic genes ($p < 0.05$). The expression matrix of these genes was obtained and the TCGA samples were divided into two clusters through the non-negative matrix factorization (NMF) algorithm (Supplementary Figure S1). The levels of expression of the 64 genes in the two subtypes are shown in Figure 1a and they differ significantly between cluster 1 (C1) and cluster 2 (C2). Furthermore, the log-rank test showed a significant difference in the overall survival (OS) between these two groups ($p = 0.026$) (Figure 1b) with a better prognosis among C2: 36.82 months in C1 vs. 43.61 months in C2. When we analyzed the disease-specific survival, we were able to confirm a significant difference between both clusters ($p = 0.033$) (data not shown).

2.1.2. Relationship between Two Subtypes and Immunity

It is known that the expression of immune genes is associated with different genomic aberrations in gynecological cancers [21] and Yang et al. [22] showed the dependence of tumor-infiltrating lymphocyte type on different types of EOC. To include an immunological viewpoint on our two clusters we used the Tumor Immune Estimation Resource (TIMER) tool. We then compared the immune scores for six different lymphocytes between C1 and C2 (Figure 2, Supplementary Table S2). Overall, the median

immune scores of the six types of lymphocytes are significantly higher in the C1 than in the C2 subtype ($p < 0.01$): B cells 0.076 vs. 0.072, CD4+ T cells 0.152 vs. 0.135, CD8+ T cells 0.181 vs. 0.147, neutrophil cells 0.143 vs. 0.109, macrophages 0.072 vs. 0.053, dendritic cells 0.517 vs. 0.453.

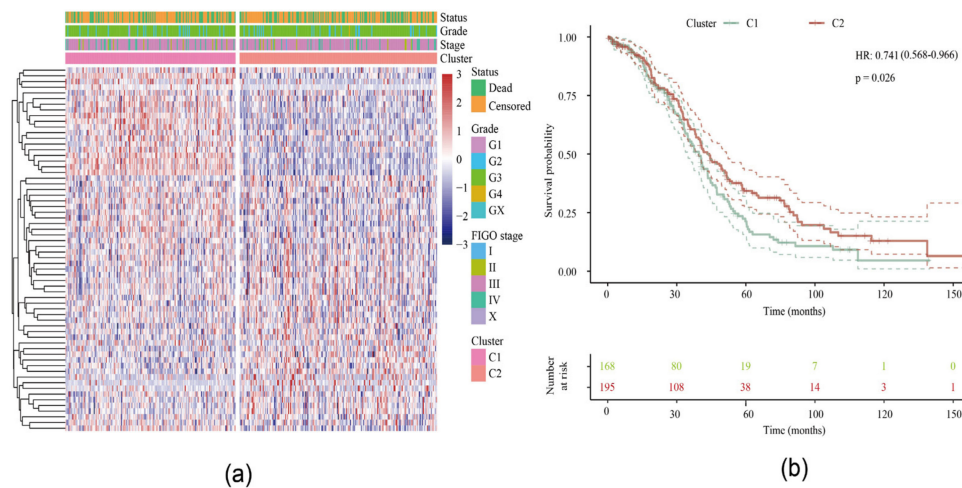


Figure 1. Identification of two molecular subtypes: (a) Heat map of clustering of 64 prognosis-related genes. (b) Survival curve of molecular subtypes including hazard ratio (HR) with 95% confidence interval (CI) and p -value.

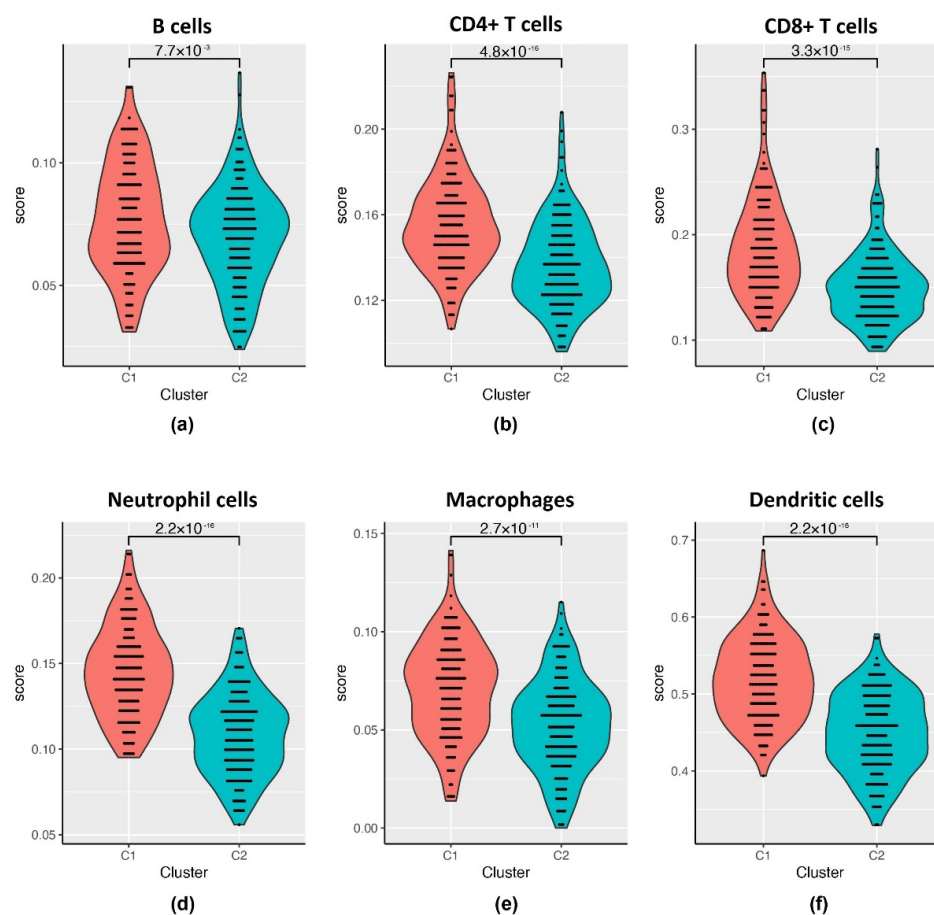


Figure 2. Tumor Immune Estimation Resource (TIMER) immune scores for cluster 1 (C1) and cluster 2 (C2) subtypes with the p -value in between for: (a) B cells; (b) CD4+ T cells; (c) CD8+ T cells; (d) neutrophil cells; (e) macrophages; (f) dendritic cells.

2.2. Analysis of DEGs between Subtypes

2.2.1. The DEGs in C2 Subtype Were Mainly Downregulated

A total of 925 differentially expressed genes (DEGs) between the two subtypes were identified (Supplementary Table S3). As shown in Supplementary Figure S2, in C2 subtype, 193 genes were upregulated in comparison to C1, whereas 732 genes were downregulated when compared to C1.

2.2.2. DEGs Are Enriched in Tumor-Related Pathways

We performed Kyoto Encyclopedia of Genes and Genomes (KEGG) and Gene Ontology (GO) functional enrichment analysis on these 925 DEGs using the clusterProfiler package. The DEGs were collectively enriched to 1871 biological process annotations and there were 749 significant annotations with a false discovery rate (FDR) < 0 (Supplementary Table S4a). For visualization, we selected the top ten functional annotations according to FDR. As shown in Figure 3a, genes are significantly enriched e.g., in regulation of leukocyte activation, T cell activation, regulation of lymphocyte activation and many more. Similarly, 66 significant functional annotations were enriched in the molecular function region (FDR < 0.01) and 58 significant functional annotations were enriched in the cellular component region (FDR < 0.01) (Figure 3b,c, Supplementary Table S4b,c). Under the scope of the KEGG database, DEGs were significantly enriched in 39 pathogenetic pathways, e.g., for rheumatoid arthritis or inflammatory bowel disease (Figure 3d, Supplementary Table S4d).

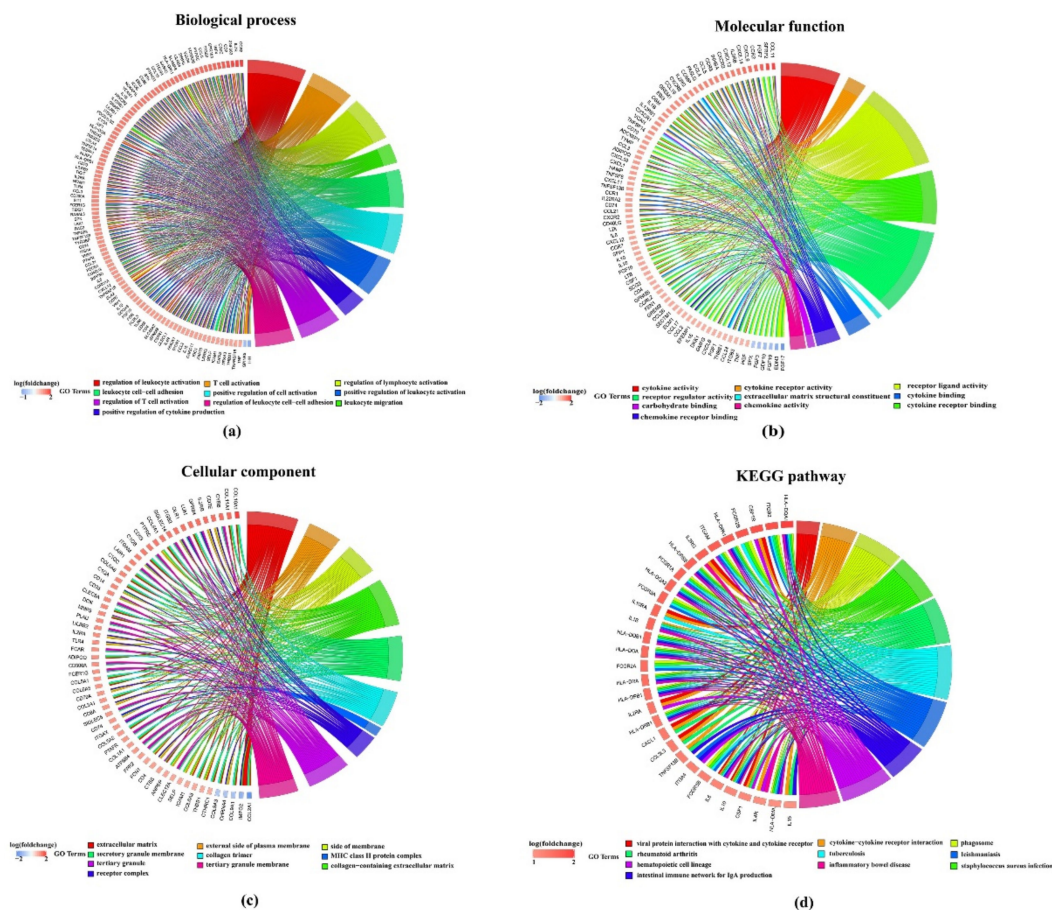


Figure 3. Kyoto Encyclopedia of Genes and Genomes (KEGG) and Gene Ontology (GO) functional enrichment analysis on 925 differentially expressed genes (DEGs) with each top ten annotations in: (a) Biological process subsection of GO; (b) molecular function subsection of GO; (c) cellular component subsection of GO; (d) KEGG pathways.

2.3. Construction of a Prognostic Risk Model

2.3.1. Randomly Grouping of Training and Testing Cohort

The final training cohort had a total of 253 samples and the testing cohort had 110 samples in total. The difference between them was analyzed using a chi-square test. The results showed that the grouping was reasonable and no significant differences were found between the groups comparing event rate, Fédération Internationale de Gynécologie et d'Obstétrique (FIGO) stage, age and grade (Table 1).

Table 1. Sample information of The Cancer Genome Atlas (TCGA) training and testing cohorts.

Clinical feature		Training Cohort	Testing Cohort	X-Squared	p-Value
Event	Censored	99	42	0.0028	0.9575
	Dead	154	68		
FIGO stage	I	0	1	3.9129	0.4179
	II	15	5		
	III	198	87		
	IV	37	17		
	None	3	0		
Age	≤60	132	59	0.0202	0.8870
	>60	121	51		
Grade	G1	1	0	2.7958	0.5926
	G2	29	13		
	G3	217	93		
	G4	0	1		
	None	6	3		

2.3.2. Univariable Analysis of Training Cohort

We used the training cohort to conduct a univariable analysis on each gene by using survival coxph function package. A *p*-value less than 0.05 was selected as the threshold. We found 30 prognosis-related significant DEGs (Supplementary Figure S3, Supplementary Table S5).

2.3.3. Construction of the 11-Gene Signature Using Lasso Regression

Lasso Cox regression analysis was performed to compress the 30 genes from Section 2.3.2. As seen in Supplementary Figure S4a, with decreased lambda, the number of independent variable coefficients approaching zero increased gradually. The model is optimal, which means stable, when lambda = 0.0686 (Supplementary Figure S4b). Therefore, we selected 11 genes under the condition of lambda = 0.0686 as the target genes (Supplementary Table S6): PI3, RGS1, ADORA3, CH25H, CCDC80, PTGER3, MATK, KLRB1, CCL19, CXCL9 and CXCL10. These 11 genes were analyzed by multivariate Cox analysis to obtain each coefficient.

2.3.4. Construction and Evaluation of a Risk Model

The 11-gene risk model was established according to the following formula:

$$\begin{aligned}
 \text{risk score (RS)} = & \\
 & (0.0004 * \text{expression level of PI3}) \\
 & + (0.0044 * \text{expression level of RGS1}) \\
 & + (0.0227 * \text{expression level of ADORA3}) \\
 & + (0.0103 * \text{expression level of CH25H}) \\
 & + (0.0078 * \text{expression level of CCDC80}) \\
 & + (0.0510 * \text{expression level of PTGER3}) \\
 & + (0.0357 * \text{expression level of MATK}) \\
 & + (-0.0640 * \text{expression level of KLRB1}) \\
 & + (-0.0294 * \text{expression level of CCL19}) \\
 & + (-0.0020 * \text{expression level of CXCL9}) \\
 & + (-0.0006 * \text{expression level of CXCL10}).
 \end{aligned}$$

The risk score (RS) of each sample was calculated and consequently the median RS was applied as the threshold to subdivide the training cohort into a high-risk group (HRG) and a low-risk group (LRG). Considering the overall distribution of the sample's OS, we evaluated the 2-year, 3-year and 5-year predictive effect of the model. In the receiver operating characteristic (ROC) curve the 5-year area under the curve (AUC) was 0.724 in the training cohort. We observed a significant difference in the Kaplan-Meier (KM) curve between the HRG and the LRG (Figure 4a).

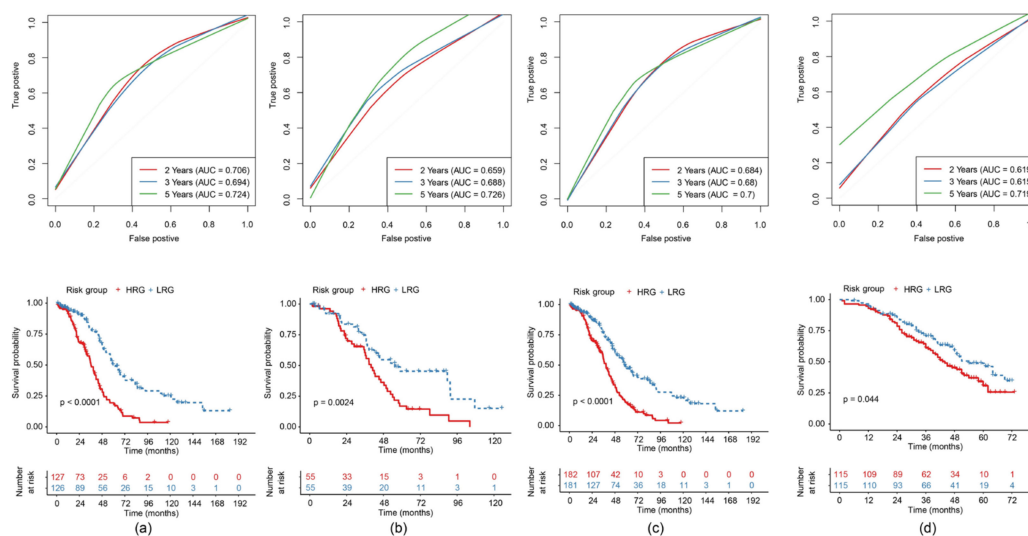


Figure 4. Construction and evaluation of the 11-gene risk model with each receiver operating characteristic (ROC) curve (area under the curve (AUC) of the 2-year, 3-year and 5-year predictive effect) above and the KM (Kaplan-Meier) curve comparing high-risk group (HRG) and low-risk group (LRG) beneath for: (a) TCGA training cohort; (b) TCGA testing cohort; (c) TCGA-(epithelial ovarian cancer (EOC)) cohort; (d) GEO Series 32026 (GSE32026) cohort.

To verify the stability and reliability of the model, we also calculated the prediction performance of this model in the testing cohort for the 2-, 3- and 5-year predictive effect. The result of the testing cohort showed a 5-year AUC of 0.726 and confirmed a significant difference between the HRG and the LRG in the KM curve (Figure 4b). The results of the ROC curve analysis and the KM curve of the whole TCGA-EOC cohort (Figure 4c), as well as the independent external validation through GEO Series 32026 (GSE32026), saw identical results (Figure 4d).

Among these 11 genes, expression levels of PI3, RGS1, ADORA3, CH25H, CCDC80, PTGER3 and MATK were upregulated in the HRG compared with the LRG. In contrast, the expression levels of KLRB1, CCL19, CXCL9 and CXCL10 were upregulated in the LRG, showing a consistent pattern within the training and the validation cohorts (Supplementary Figure S5).

In order to evaluate the stability of the model, we conducted 1000 random samplings at different proportions from all the TCGA-EOC samples. We found significance in 997 out of 1000 times when the sampling ratio was 0.5 (Supplementary Figure S6). This confirmed a lower sampling bias.

2.4. Univariable and Multivariable Analysis of Gene Signature

To identify the independence of the 11-gene signature model in clinical application, we conducted univariable and multivariable Cox regression analysis to investigate the relevant hazard ratio (HR), 95% confidence interval (CI) of HR and the *p*-value. We systematically analyzed the clinical information of TCGA patients including age, FIGO stage, grade and our RS of the 11-gene signature (Table 2). Univariable Cox regression analysis found that the RS was significantly related to survival (HR = 1.593, 95% CI: 1.377-1.843, *p* = 3.77E-10). Moreover, the corresponding multivariable Cox regression analysis found that the RS also correlated significantly with survival (HR = 1.534, 95% CI: 1.322-1.780, *p* = 1.65E-08) (Supplementary Figure S7).

Table 2. Univariable and multivariable analysis of the TCGA cohort.

Clinical Feature	Univariable Analysis			Multivariable Analysis		
	HR	95% CI	p-Value	HR	95% CI	p-Value
FIGO stage	1.929	0.856–4.349	0.1130	1.705	0.742–3.917	0.2080
Age	1.186	0.791–1.778	0.4090	1.087	0.718–1.646	0.6940
Grade	1.291	0.992–1.680	0.0576	1.193	0.910–1.566	0.2020
RS	1.593	1.377–1.843	<0.001	1.534	1.322–1.780	<0.001

2.5. Survival Curves of Risk Models in Different Clinical Subgroups

In order to verify the effect of our model on clinical subgroup characteristics, we classified the TCGA-EOC cohort according to the different clinical characteristics from Table 1. Significant differences were found between the HRG and the LRG in FIGO stage III and IV ($p < 0.05$) (Figure 5a,b). Due to an insufficient amount of stage I and II samples, we did not analyze them. G1 samples were also not examined because of the lack of data. G2 sample differences between both groups were not significant (Figure 5c), whereas G3 samples showed a significant difference ($p < 0.01$) (Figure 5d). Patient samples ≤ 60 years as well as >60 years showed a significant difference between the HRG and the LRG ($p < 0.0001$) (Figure 5e,f).

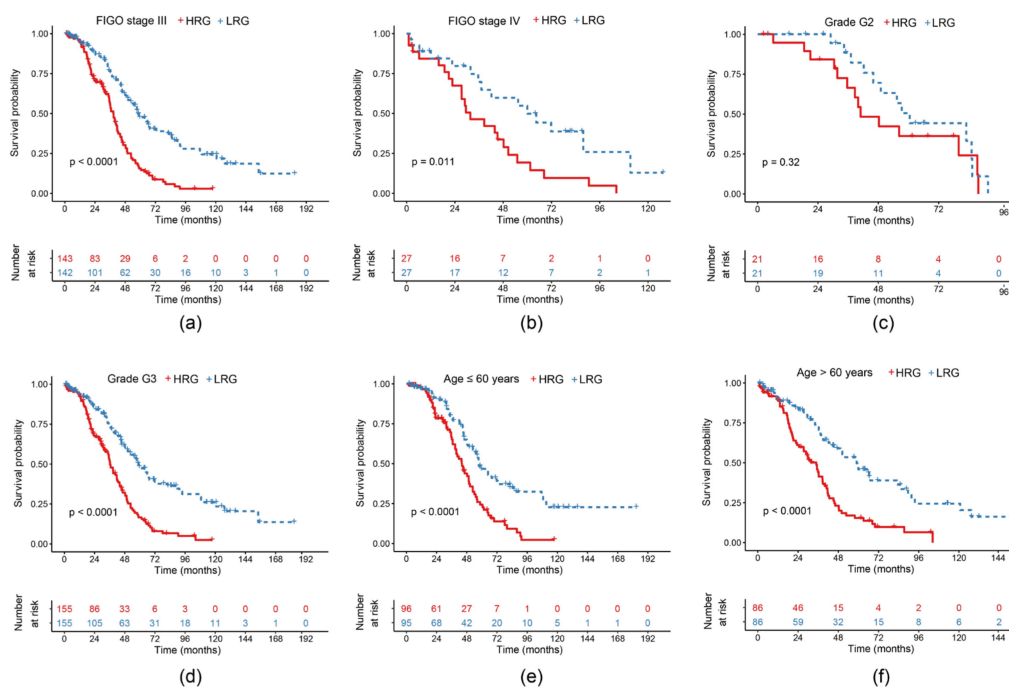


Figure 5. KM curves of overall survival (OS) of the 11-gene risk model in different clinical subgroups: (a) Fédération Internationale de Gynécologie et d’Obstétrique (FIGO) stage III; (b) FIGO stage IV; (c) grade G2; (d) grade G3; (e) age ≤ 60 years; (f) age > 60 years.

2.6. Survival Prognosis on Different Mutation Subtypes in the Risk Model

To verify the effectiveness of our model on different common mutation subtypes of EOC, the TCGA-EOC cohort was classified according to different single nucleotide variant types. In TCGA there are a total of 436 exon sequencing samples. Here, we saw 411 altered ones ($=94.27\%$). Mutations of TP53 and TTN dominated this classification (Figure 6a). Consequently, we conducted the KM curves for the 363 RNA-Seq samples and found that, regardless of with/without a TP53/TNN mutation, prognosis in the HRG was worse compared to the LRG (Figure 6b–e).

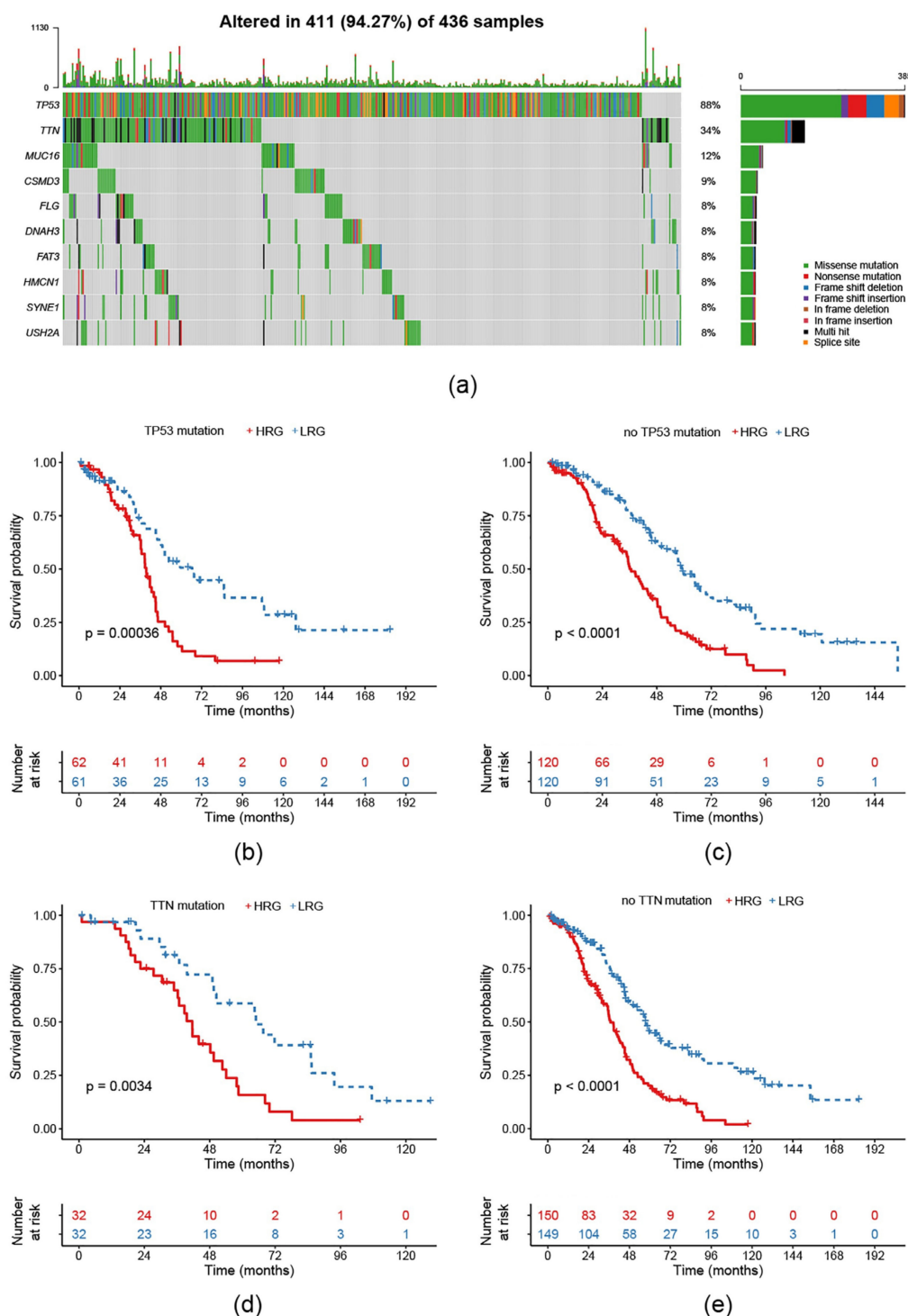


Figure 6. Analysis of the 11-gene signature model in different single nucleotide mutations: (a) Distribution of common single nucleotide mutations in EOC with 94.27% samples altered; (b) KM curve of TP53-mutated samples; (c) KM curve of not TP53-mutated samples; (d) KM curve of TTN-mutated samples; (e) KM curve of not TP53-mutated samples.

2.7. Construction of Nomogram Model Based on RS and Clinical Features

We combined the traditional clinical features FIGO stage, age and grade with our RS to construct a nomogram model to predict the OS of EOC patients (Figure 7a). In the modeling results the RS has the greatest impact on survival prediction. Calibration plots were used to visualize the performances

of the nomograms. The 2-year, 3-year and 5-year calibration plots demonstrated the performance of our model (Figure 7b).

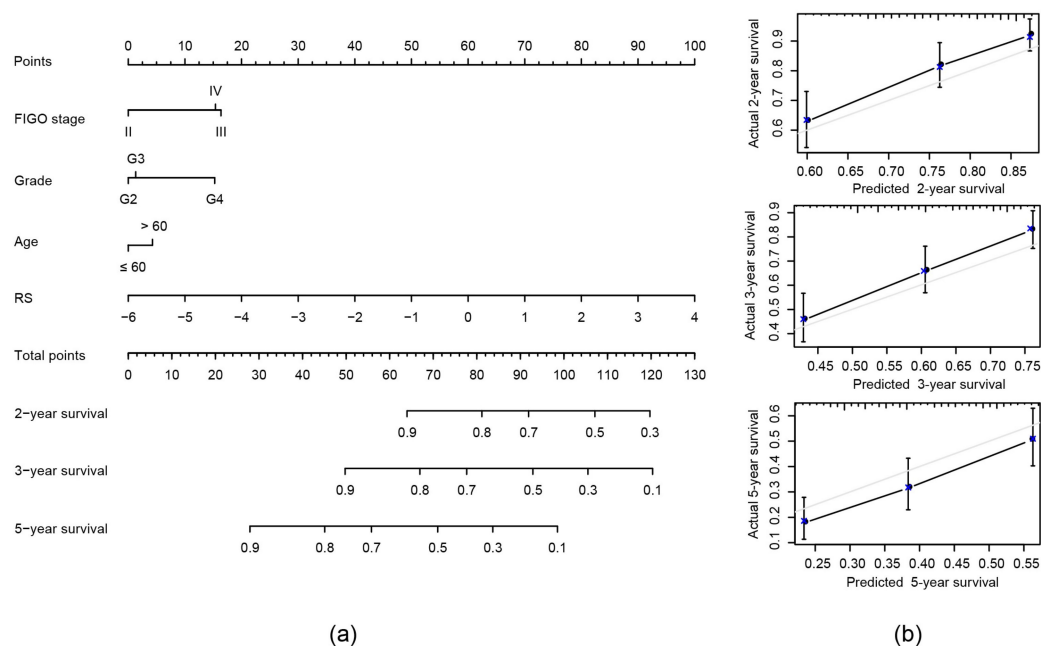


Figure 7. Construction of nomogram model: (a) Nomogram predicting 2-, 3- and 5-year OS for patients with EOC. The nomogram is applied by adding up the points identified on the points scale for each variable to a total points amount. Finally, beneath the total points, the probability of 2-, 3- or 5-year survival is projected on the bottom scales. (b) Calibration curves for nomogram predicted 2-, 3- and 5-year OS for patients with EOC in relation to actual survival.

Under consideration of the nomogram, we saw its notably superior predictive ability compared to clinical features themselves or the RS alone. The concordance index (C-index) of the nomogram was the highest (0.663) compared to the other variables, as seen in Table 3.

Table 3. Comparison of concordance indexes (C-indexes) between clinical features, risk score (RS) and nomogram.

Variables	C-Index	95% CI of C-Index	p-Value
FIGO stage	0.609	0.523–0.696	0.013
Age	0.619	0.547–0.690	0.001
Grade	0.593	0.508–0.678	0.032
RS	0.658	0.602–0.684	<0.001
Nomogram	0.663	0.625–0.701	<0.001

2.8. ROC Curve and DCA of Nomogram Model

To demonstrate putative advantages of the nomogram model, we compared the 2-, 3- and 5-year ROC curves of the single variables against the nomogram curve. The highest AUC each was seen for the nomogram model (Figure 8a–c). Furthermore, a decision curve analysis (DCA) confirmed our expectations. The net benefit in 2-, 3- and 5-year predictions was the highest in the combined nomogram model compared to the single variable models (Figure 8d–f). These methods showed the improved clinical utility of our nomogram model.

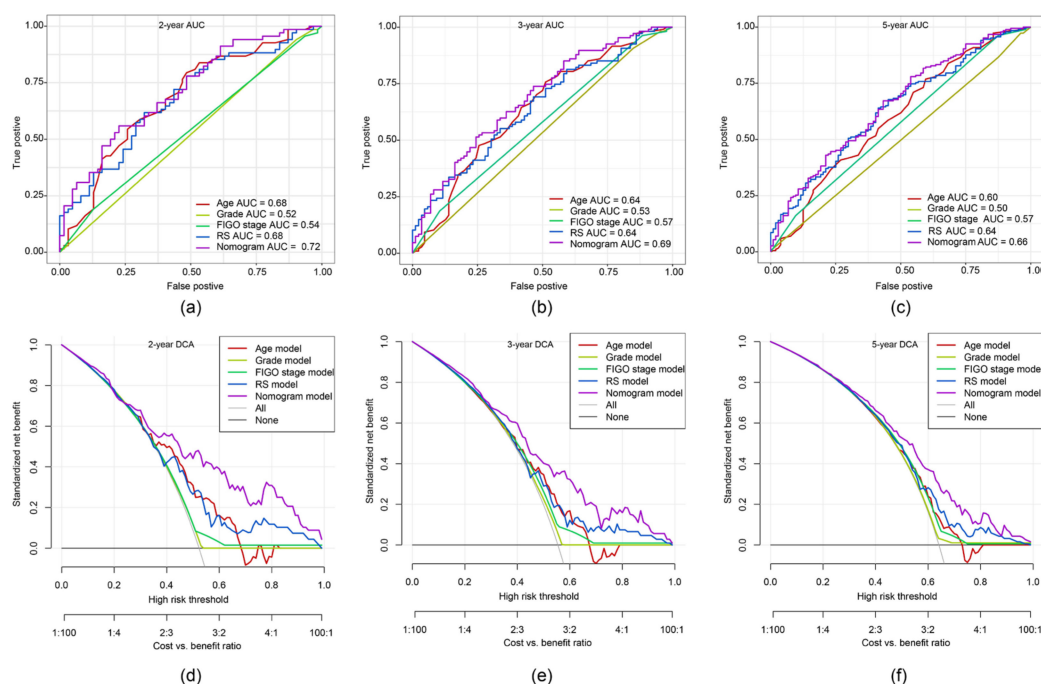


Figure 8. The time-dependent ROC curve and decision curve analysis (DCA) on the nomogram compared to single variable models. Through the ROC curves the accuracy of the models was tested for: (a) 2-year survival; (b) 3-year survival; (c) 5-year survival. The DCA curves can evaluate the clinical benefit of the nomograms and the scope of application. Black indicates that all samples are negative and none are treated, therefore the net benefit is 0. Grey indicates that all samples are positive and all are treated. The x-axis represents threshold probabilities of patients having: (d) 2-year survival; (e) 3-year survival; (f) 5-year survival.

2.9. Comparison of the 11-Gene Risk Model with Other Models

Five prognosis-related risk models were selected, including a 19-gene signature from Yang et al. [23], a 32-gene signature from Willis et al. [24], a 10-gene signature from Wang et al. [25], a 7-gene signature from Sabatier et al. [26] and a 101-gene signature from Millstein et al. [27] to compare with our 11-gene model. To ensure comparability, we calculated a RS of the TCGA-EOC cohort for all five models using the same methods as in our gene signature but based on the corresponding genes of each model. As described, samples were divided into a HRG and a LRG with the median as the threshold. The ROC and KM curves of the five models are shown in Figure 9a–e. Only the AUC of Millstein et al. [27] averaged above our model. No significant difference in prognosis was found among the 7-gene signature, whereas all others confirmed significant differences between the HRG and the LRG.

To compare the predictive performance of these models on EOC, we used the restricted mean survival (RMS) package in R [28] to calculate the C-index of all six models including the 11-gene model. The highest C-index was seen in the 101-gene model, while our C-index ranks second (Figure 9f). We used the RMS time to evaluate the predictive effect of the six models at different time points. The RMS curves showed that the six models had an overlap of 58 months. Under the condition of <58 months, our 11-gene risk model performed better than the models from Yang et al. [23], Willis et al. [24], Wang et al. [25] and Millstein et al. [27] (Figure 9g). Thus, our risk model is more suitable to evaluate the data of <5-year OS.

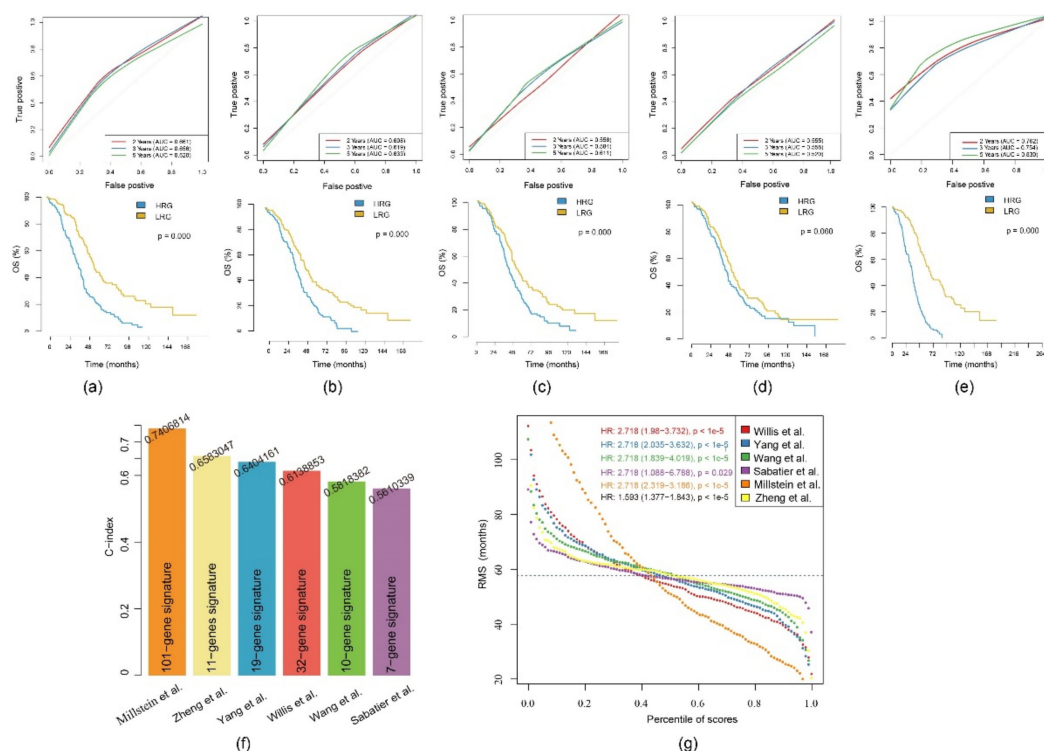


Figure 9. Comparison of the 11-gene risk model with other models: (a) The ROC and KM curves of the 19-gene signature (Yang et al.); (b) the ROC and KM curves of the 32-gene signature (Willis et al.); (c) the ROC and KM curves of the 10-gene signature (Wang et al.); (d) the ROC and KM curves of the 7-gene signature (Sabatier et al.); (e) the ROC and KM curves of the 101-gene signature (Millstein et al.). (f) Concordance indexes (C-indexes) of the six prognostic risk models; (g) Restricted mean survival (RMS) time curve of the six prognostic risk models (the dashed line indicates the overlap of 58 months).

2.10. Expression of a Gene Product from the 11-Gene Signature in an EOC Cohort

In a representative EOC cohort from our hospital we measured the expression of the prostaglandin E2 receptor 3 (EP3) encoded by PTGER3, which is upregulated in the HRG compared to the LRG. An immunohistochemistry (IHC) score >1 represents elevated expression of EP3, while an IHC score ≤1 shows low expression. In parallel to the 11-gene signature, higher expression of PTGER3's gene product EP3 is correlated with poor OS in both the whole cohort and the serous subgroup (Figure 10). Even without having found any significance, this finding supports the functionality of the 11-gene signature and could act as a basis for further confirmation of the 11-gene signature in a clinical context.

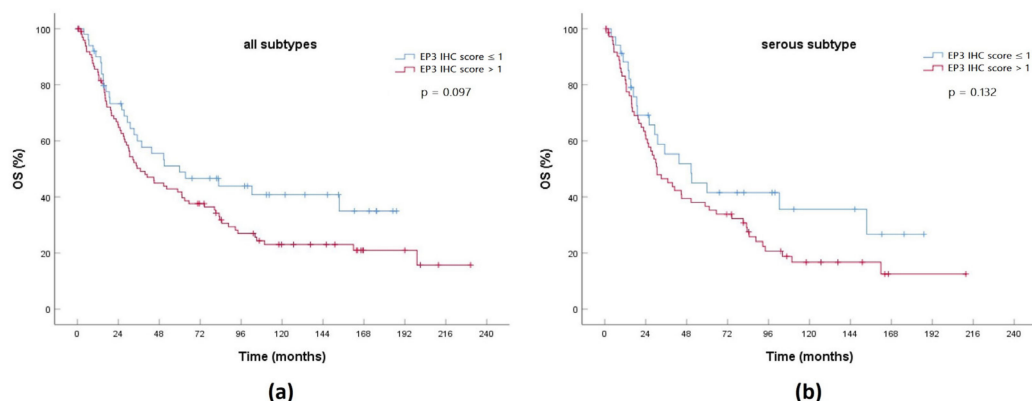


Figure 10. KM curve comparing OS between high prostaglandin E2 receptor 3 (EP3) expression (immunohistochemistry (IHC) score >1) and low EP3 expression (IHC score ≤1) in an EOC cohort for: (a) all histological subtypes ($n = 151$); (b) serous subtype ($n = 108$).

2.11. Translational Level Validation Related to Signature Genes

In order to analyze the translational levels of more signature genes besides PTGER3, the Human Protein Atlas (HPA) database was used. ADORA3, CH25H, CCL19, CXCL9 and CXCL10 were not recorded in the database. The results of PI3, RGS1, CCDC80, PTGER3, MATK and KLRB1 are shown in Figure 11a–f. We found that the expression intensity and quantity of PI3, RGS1, PTGER3 and CCDC80 in ovarian cancer tissue was higher than that in normal ovarian tissue. In contrast, KLRB1's expression intensity and quantity in normal tissue was higher than that in the tumor tissue. Both findings concur with the expression profile in our 11-gene signature. The expression intensity of MATK in ovarian cancer tissue was lower than that in normal tissue, but had a higher expression quantity, which cannot be clearly correlated with the polarity of our signature.

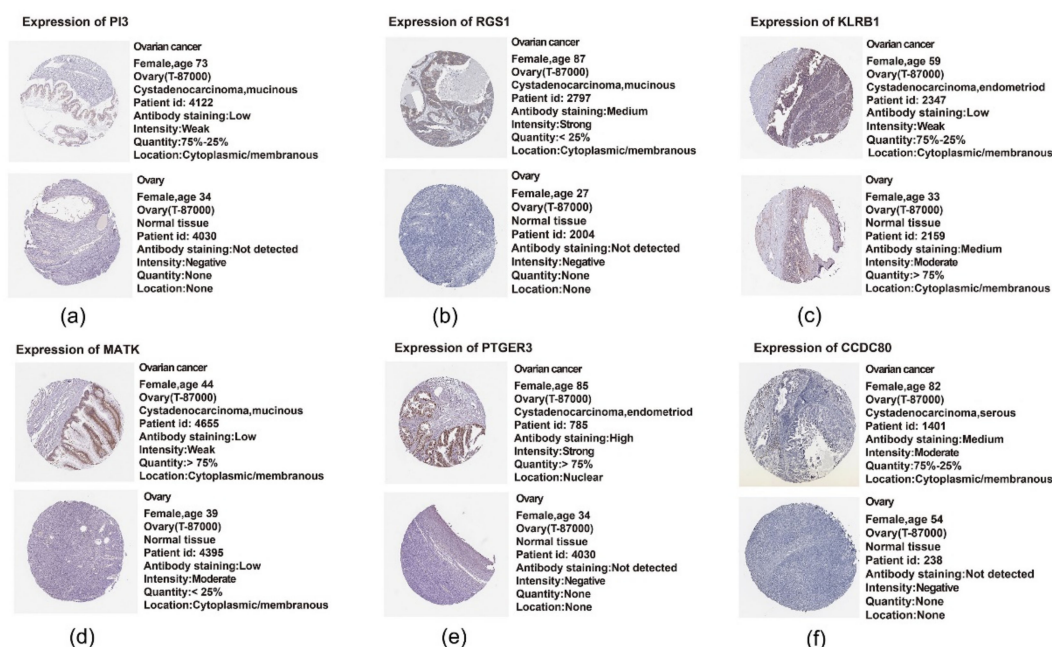


Figure 11. Translational level validation of six genes from the 11-gene signature using the Human Protein Atlas (HPA): (a) expression of PI3; (b) expression of RGS1; (c) expression of KLRB1; (d) expression of MATK; (e) expression of PTGER3; (f) expression of CCDC80.

3. Discussion

Due to the lack of early detection and prevention, 70% of EOC patients present in an advanced stage with distant metastases upon diagnosis, making ovarian carcinoma the leading cause of death among malignant gynecological tumors [1,2,29]. Traditional prognostic criteria are not sufficient in accurately predicting the survival of an individual patient. Multiple large cancer databases, such as TCGA and GEO, offer researchers the opportunity to analyze gene expression data and the corresponding clinical information on a large scale [30,31]. Until now, previous studies on gene expression have seen modifications in the molecular signature between benign and malignant tumors or low and advanced tumor stages [32–34]. Meanwhile, lipidomic analyses of serum samples have confirmed differences in the lipid profile depending on the tumor's dignity [14] and have even been shown to prevent peritoneal metastases when targeting the lipid metabolism signaling axis [15].

Consequently, in this study, 363 EOC samples from the TCGA were subdivided, based on 751 lipid metabolism-related genes, into two subtypes. We report that the prognosis of the C1 subtype is significantly poorer than that of the C2 subtype. This finding suggests that lipid-based molecular subtypes can be used, to a certain extent, as an indication for evaluating the prognosis of patients.

To study the individual role of all the genes of both subtypes, we obtained 925 DEGs, of which 193 were upregulated and 732 were downregulated in the C2 compared to the C1 subtype. These genes were mainly active in the regulation of leukocyte activation, T cell activation, regulation of lymphocyte activation and other immunological functions. The immune cell infiltration scores in C1 were found to be significantly higher than in the C2 subtype. Numerous other studies have pointed out the prognostic significance of tumor-infiltrating lymphocytes in other various cancers [35–38]. It has been reported that tumor infiltration by a subpopulation of CD4⁺ T cells with immunosuppressive properties predicted reduced survival in EOC [39,40]. Therefore, we can infer that the C1 subtype has a worse prognosis partly because the proportion of CD4⁺ and CD8⁺ T cells in the C1 subtype is larger than that in the C2 subtype, as an excessive immune enhancement process might also be a sign of poor prognosis for patients.

Currently, studies on the effect of lipid metabolism on tumor immune functions are being carried out to examine a potential link between both lipids and immune regulation. Interestingly, Wefers et al. [41] discovered that the dysregulation of lipid metabolism in the ascites of EOC patients can affect the immune system by regulating T cell proliferation.

Out of the 925 DEGs, we constructed an 11-gene prognostic risk model based on the genes PI3, RGS1, ADORA3, CH25H, CCDC80, PTGER3, MATK, KLRB1, CCL19, CXCL9 and CXCL10. This model shows a strong robustness and can be used in the prognosis predictions of EOC patients. Calibration plots demonstrated that our nomogram is superior in terms of predictive performance when compared to the grading and FIGO stage. Traditional scores like TNM or FIGO depend on an anatomical spread and, therefore, cannot reflect the biological heterogeneity of EOC [3], which may affect their accuracy.

This is the first prognostic model based on lipid metabolism expression profile. Compared with five other prognostic risk models [23–27] for EOC, the predictive effect of our model at different time points shows that within a survival period of less than 58 months, our 11-gene risk model is the most powerful. Although the model from Millstein et al. [27] has a very high 5-year AUC, it should be noted that their model involves a very large amount of genes, indicating higher costs and consequently reduced clinical utility. Among these 101 genes, the top five genes were TAP1, ZFHX4, CXCL9, FBN1 and PTGER3 ($p < 0.001$), which is interesting because PTGER3 and CXCL9 are in our 11-gene model as well.

While other models did not use the lipid metabolism approach, they also used mainly TCGA data as a base [23–25,27]. Only Sabatier et al. [26] used their own patient cohort. In the 19-gene signature, they initially performed a combination with clinical data, which we did as well in the nomogram [23]. In the model from Wang et al. [25] they already evaluated ten biomarkers from the candidate genes and achieved a 100% accuracy. To optimize an early diagnosis of EOC this access via biomarkers

remains promising and encouraged us to take a more detailed look into some of the 11 genes selected in our model.

PI3 is located on chromosome 20q 12-13.1 [42,43] and encodes elafin, also known as peptidase inhibitor 3. It is reported to be highly expressed in high-grade serous EOC and is associated with a poor prognosis [44,45]. Wei et al. [46] suggested that elafin selectively regulates the sensitivity of EOC cells to genotoxic drug-induced apoptosis. Our results show that the higher the risk value, the higher the expression of PI3 and the poorer the prognosis of EOC patients, which is consistent with the experimental results.

The regulator of G protein signaling 1 is encoded through RGS1, located on chromosome 1q 31.2 [47]. There is increasing evidence for aberrantly differentiated expression of certain members of this protein family in various cancers and their capability of mediating the proliferation or migration of cancer cells [48]. A study had shown that RGS1 is highly expressed in advanced cervical cancer and is associated with cancer progression [49]. So far, besides our identification of RGS1 as a negative prognosticator in this 11-gene model, no other study reported a role of RGS1 in EOC.

Adenosine receptors are a class of purinergic G protein-coupled receptors with adenosine as an endogenous ligand [50], and ADORA3 codes for one of them. In humans, they are involved in the induction of p53-mediated apoptosis. Consequently, in lung cancer cell lines it is used as a target for antibody-based therapy in p53 mutant tumors [51]. In parallel to RGS1, the biological role of ADORA3 in EOC has not been clarified yet but it should be a target for future research due to its high expression in ovarian tissue, as seen in databases.

The gene product of CH25H is cholesterol 25-hydroxylase, which catalyzes the formation of 25-hydroxycholesterol and thereby results in an inhibitive effect on cholesterol biosynthetic enzymes. It is derived and secreted by U87MG and GM133 glioblastoma cell lines and may be involved in the recruitment of immune-competent cells [52,53]. Mitempergher et al. [54] found that CH25H expression is correlated with the prognosis of breast cancer patients and is an independent predictor of distant metastasis, which is consistent with our data.

The prostaglandin receptor EP3, encoded by PTGER3, is one of the four identified receptors that mediate the effects of prostaglandin E2 [55]. In previous work, our IHC assay showed that EP3 is highly expressed in tissues of clear cell ovarian carcinomas and is a prognostic factor in tumor-associated mucin-1 negative EOC [56]. In this model, high expression of EP3 was grouped into the HRG, indicating poor prognosis, which concurs with the previous experimental results. An analysis of the whole patient cohort from Czogalla et al. [56] for both all histologic subtypes and serous subtype also confirmed the correlation between EP3 as a “high-risk” gene and the clinical data. Additionally to EOC, in our lab, EP3 has been identified as an independent risk factor for the survival prognosis of patients with other gynecological malignancies, such as cervical [57], endometrial [58] and breast cancer [59].

In a meta-analysis of the pan-carcinoma resources and expression characteristics of 18,000 human tumors, Gentles et al. [60] found that KLRB1 is the most favorable pan-cancer prognosis gene and is a marker for enhanced innate immune characteristics in different T cell subsets. Consistent with the results of previous studies, we found that the high expression of KLRB1 was located in the LRG with a good prognosis.

However, several limitations in the current research should be considered. Firstly, the TCGA database is mainly constrained to Caucasian and African populations. Therefore, a robust nomogram should be further validated within multicenter clinical trials and prospective studies. Secondly, we do not have experimental data for the majority of these genes to prove the correlation between the 11-gene signature and the prognosis of EOC. Some of these genes have yet to be reported in the context of EOC. Moreover, we need more external independent datasets to further verify our signature even if we proved the robustness of our signature in the GSE32026 verification cohort.

A follow-up study to analyze the translational levels of the proteins belonging to these 11 genes is under work. In clinical environments, this gene signature can primarily be used as an additional tool, as it still must be validated in large patient cohorts. In actuality, the signature might support the shared

decision-making for or against an incriminating therapy in special patient groups, e.g., very old patients or others with relevant comorbidities. Another aspect in regards to personalized treatment could be the patient's RS as an indicator for the adaptation of gynecological aftercare intervals. Finally, biomarkers or therapeutic drugs resulting from our or other gene signatures mentioned would naturally be the biggest therapeutic gain but need a lot of further research. Nevertheless, in an upcoming era of next generation sequencing and gene expression profiling, which we already use as standard critical diagnostic tests in breast cancer treatment, multiple prognostic gene signatures will grow in relevance in our clinical decision-making.

4. Materials and Methods

4.1. Ovarian Cancer Cohort Source and Preprocessing

The gene expression profiles and clinical follow-up information of EOC were downloaded using the TCGA Genomic Data Commons Application Programming Interface. This cohort contains the expression profile information of 379 RNA-Seq samples of serous EOC. GSE32026 data, downloaded from GEO, covering 260 samples. Human lipid metabolism-related pathways were downloaded from Molecular Signature Database (version 7.0). Herein, a total of 776 lipid metabolism genes (Supplementary Table S7) were sorted out from the six lipid metabolism pathways from the databases KEGG and Reactome (Table 4).

Table 4. Lipid metabolism-related pathways in the Molecular Signature Database.

Pathway	Database	Gene Count
Peroxisome proliferator activated receptor alpha	Reactome	119
Metabolism of lipids	Reactome	738
Sphingolipid metabolism	Reactome	89
Transcriptional regulation of white adipocyte differentiation	Reactome	84
Glycerophospholipid metabolism	KEGG	77
Fatty acid metabolism	Reactome	77
		Total: 1184
		Unique: 776

The RNA-Seq data from the TCGA were preprocessed by removing the samples without clinical follow-up information, removing the genes with fragments per kilobase of exon model per million reads mapped less than one and retaining the lipid metabolism gene expression profile. The same procedures were performed on the GSE32026 cohort and the benign tissue samples were removed. A total of 363 samples from the TCGA cohort along with 230 samples from the GSE32026 cohort remained (Table 5).

Table 5. Clinical features of the remaining TCGA and GEO Series 32026 (GSE32026) cohort.

Clinical Feature		TCGA	GSE32026
Event	Censored	141	116
	Dead	222	114
FIGO stage	I	1	
	II	20	
	III	285	
	IV	54	
	None	3	
Age	≤60	191	
	>60	172	
Grade	G1	1	
	G2	42	
	G3	310	
	G4	1	
	None	9	

4.2. Identification of Molecular Subtypes Using NMF Algorithm

We extracted the expression of these 776 lipid metabolism genes from the TCGA expression profile data and retained the samples with a gene expression level above zero; 751 genes remained. Univariable Cox analysis was performed using the `coxph` function in R package to determine the genes that are correlated with the prognosis expressed as OS of EOC ($p < 0.05$). The NMF was used to cluster the EOC samples under the condition of standard NMF “brunet” for 50 iterations by extracting biological correlation coefficients and internal feature structures of the gene expression matrix. The number of clusters k was identified from 2–10. The average contour width of the common member matrix was determined using the NMF package and the minimum member of each subclass was set to 10. We calculated the optimal number of clusters. The selection was based on the following parameters: cophenetic, residual sum of squares and silhouette.

4.3. Comparison of Molecular Subtype Immune Scores with TIMER

TIMER is a web server for comprehensive analysis of tumor-infiltrating immune cells [61]. The levels of six tumor-infiltrating immune subsets are precalculated for 10,897 tumors from 32 cancer types. It provides six major analytic modules that allow users to interactively explore the associations between immune infiltrates and a wide spectrum of factors, including gene expression, clinical outcome, somatic mutations and somatic copy number alterations.

4.4. Functional Analysis of DEGs

We used the `limma` package (version 3.42) in R to calculate the DEGs between the different molecular subtypes and filtered these genes in accordance with the threshold of a $FDR < 0.05$ and $|\log_2(\text{foldchange})| > 1$ [62]. Further analysis of the DEGs was performed using `clusterProfiler` package (version 3.13) [63]. KEGG and GO functional enrichment analysis was conducted. KEGG, a database for the systematic analysis of gene functions, associates genomic information with gene functions and aims to reveal the genetic and chemical blueprint of life. GO enrichment analysis covers three areas including cell components, molecular function, and biological processes.

4.5. Sample Preparation

Firstly, the 363 samples in the TCGA cohort were divided into a training cohort and a validation cohort. In order to prevent the bias of random allocation from undermining the stability of the subsequent modeling, all samples were put back into random groups for 100-times in advance. Herein, the group sampling of the training and validation cohorts was performed in the ratio of 7:3. The most suitable training and validation cohorts were selected according to the following conditions:

- The two cohorts are similar in age distribution, FIGO stage, follow-up time and death rate of patients.
- The gene expression profiles of the two data sets that were randomly grouped were close in the number of classified samples.

Finally, the training set had a total of 253 samples and the validation set had 110 samples in total. The difference between the training set and the validation set was analyzed using a chi-square test.

4.6. Lasso Regression Analysis

We further compressed the genes using lasso regression to reduce the number of genes in the risk model. The lasso method is a compression estimation used to build a more refined model by constructing a penalty function, thereby compressing some coefficients and setting some coefficients to zero [64]. Therefore, it retains the advantage of contraction in subsets. It is a biased estimation that can be used to process complex collinearity data and can realize the simultaneous selection of variables during parameter estimation to optimize the multicollinearity problem in regression analysis. When applying the `glmnet`

package for lasso Cox regression analysis, we used 3-fold cross validation for model construction and analyzed the CI in each lambda [65].

4.7. Stability Assessment of Risk Model

In order to evaluate the impact of random sampling on the stability of the model, we conducted 1000 instances of random sampling at various proportions in all TCGA samples to evaluate the times of significant difference in the prognosis of the HRG and the LRG samples.

4.8. Construction of Nomogram Combined with RS and Clinical Features

Nomogram is a method to display the results of the risk model intuitively and effectively. It is conveniently applied in the prediction of the outcome. It uses the length of the line to represent the different variables, thereby exhibiting the effect of different variable values on the outcome. We used the TCGA-EOC cohort to build a nomogram that combines FIGO stage, age, grade and RS.

4.9. Analysis of DCA

DCA is a simple method to evaluate clinical predictive models, diagnostic tests and molecular markers. It was initially used as a novel analytical technique that incorporated the clinical consequences of a decision to quantify the clinical utility of a prediction nomogram. Therefore, the DCA can decide whether the predictive nomogram is clinically useful or not. The best model is one with a high net benefit as calculated within the favorable probability.

4.10. IHC of EP3 in an EOC Patient Cohort

The specimens derived from a cohort consisting of 151 patients with epithelial EOC (serous: $n = 108$, endometrioid: $n = 20$, clear cell: $n = 11$, mucinous $n = 12$) who underwent radical cytoreductive surgery in our department between 1990 and 2002. Diagnoses were done by a specialized gynecologic pathologist. Advanced disease (FIGO IIB-IV) presented in approximately three quarters of the specimens. Mean age at primary diagnosis was 58.9 years. All patients, except FIGO stage IA, received adjuvant platinum-based chemotherapy. Lifetime data were taken from our patient charts, the Munich Cancer Registry and aftercare calendars. The study was carried out in compliance with the guidelines of the Helsinki Declaration of 1964 (last revision October 2018). All participants gave their written informed consent.

The procedure of IHC has been previously described in our lab [57–59]. We stained tissue microarrays of the EOC samples of paraffin-embedded and formalin-fixed tissues after epitope retrieval with primary polyclonal anti-EP3 rabbit IgG (Abcam, Cambridge, UK). Afterwards, detection was performed via polymer method with Zytomed Plus HRP Polymer System mouse/ rabbit (Zytomed Systems, Berlin, Germany) and the chromogen diaminobenzidine (Dako, Hamburg, Germany).

The IHC staining was assessed semiquantitatively, according to Remmele and Stegner [66], using the IHC score. EP3 expression was regarded as negative with an IHC score 0–1 and as positive with an IHC score >1 . Expression-dependent differences in OS were tested by chi-square statistic of the Log-Rank test (Mantel-Cox) in KM curves with SPSS Statistics 25 (IBM, Chicago, IL, USA).

4.11. Translational Level Validation of Signature Genes

The HPA was initiated in 2003 and shows the expression and localization of human proteins across tissues and organs. It is based on deep RNA-seq from 37 major tissue types and IHC on tissue microarrays containing 44 different tissue types. Altogether, 76 different cell types, covering all major parts of the human body, have been analyzed manually and the data are presented as histology-based annotation of protein expression levels. The antibody-based protein profiles are qualitative and describe the spatial distribution, cell type specificity and the rough relative abundance of proteins in these tissues, whereas the mRNA data provide quantitative data on the average gene expression within an entire

tissue. For each gene, the immunohistochemical staining profile is matched with mRNA data and gene/protein characterization data to yield an “annotated protein expression” profile.

5. Conclusions

In conclusion, we propose the first 11-gene signature (PI3, RGS1, ADORA3, CH25H, CCDC80, PTGER3, MATK, KLRB1, CCL19, CXCL9 and CXCL10) prediction model based on lipid metabolism-related genes in EOC. Despite different drawbacks of the current analysis, this model may be an interesting approach for a molecular diagnostic test assessing the prognosis and possible risk factors of EOC patients. Furthermore, the development of biomarkers based on this gene signature could represent a perspective for the clinical care of cancer patients.

Supplementary Materials: The following are available online at <http://www.mdpi.com/1422-0067/21/23/9169/s1>.

Author Contributions: Conceptualization, M.Z., U.J., F.T., S.M. and T.K.; data curation, M.Z.; formal analysis, M.Z. and T.K.; investigation, M.Z., A.H., B.C., H.H., T.V., A.V., A.C.-R. and T.K.; methodology, M.Z. and T.K.; project administration, U.J., F.T., S.M. and T.K.; resources, M.Z., B.C., H.H., T.V., A.V., A.C.-R. and T.K.; software, M.Z. and A.H.; supervision, U.J., F.T., S.M. and T.K.; validation, M.Z., U.J. and T.K.; visualization, M.Z. and T.K.; writing—original draft, M.Z.; writing—review and editing, H.M. and T.K. All authors have read and agreed to the published version of the manuscript.

Funding: We acknowledge financial support by the China Scholarship Council for Mingjun Zheng and by the Friedrich-Baur-Stiftung (registration number 33/19) for Till Kaltofen.

Acknowledgments: The results shown here are in whole or part based upon data generated by TCGA (<https://www.cancer.gov/tcga>), KEGG (<https://www.genome.jp/kegg>), GO (<http://geneontology.org>), GEO (<https://www.ncbi.nlm.nih.gov/geo>), Reactome (<https://reactome.org>), Molecular Signature Database (<https://www.gsea-msigdb.org/gsea/msigdb/index.jsp>) and HPA (<http://www.proteinatlas.org>).

Conflicts of Interest: F.T.: Research support, advisory board, honoraria and travel expenses from AstraZeneca, Clovis, Medac, PharmaMar, Roche and Tesaro. S.M.: Research support, advisory board, honoraria and travel expenses from AbbVie, AstraZeneca, Clovis, Eisai, GlaxoSmithKline, Medac, MSD, Novartis, Olympus, PharmaMar, Pfizer, Roche, Sensor Kinesis, Teva and Tesaro. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, or in the decision to publish the results. The other authors declare no conflict of interest.

Abbreviations

AUC	area under the curve
C1	cluster 1
C2	cluster 2
C-index	concordance index
CI	confidence interval
DCA	decision curve analysis
DEG	differentially expressed gene
EOC	epithelial ovarian cancer
FDR	false discovery rate
FIGO	Fédération Internationale de Gynécologie et d'Obstétrique
GEO	Gene Expression Omnibus
GO	Gene Ontology
GSE32026	GEO Series 32026
HPA	Human Protein Atlas
HR	hazard ratio
HRG	high-risk group
IHC	immunohistochemistry
KEGG	Kyoto Encyclopedia of Genes and Genomes
KM	Kaplan-Meier
LRG	low-risk group
NMF	non-negative matrix factorization
OS	overall survival
EP3	prostaglandin E2 receptor 3
RMS	restricted mean survival
ROC	receiver operating characteristic
RS	risk score
TCGA	The Cancer Genome Atlas
TIMER	Tumor Immune Estimation Resource

References

1. Siegel, R.L.; Miller, K.D.; Jemal, A. Cancer statistics 2019. *CA Cancer J. Clin.* **2019**, *69*, 7–34. [\[PubMed\]](#)
2. Waldmann, A.; Eisemann, N.; Katalinic, A. Epidemiology of malignant cervical, corpus uteri and ovarian tumours-current data and epidemiological trends. *Geburtshilfe Frauenheilkd* **2013**, *73*, 123–129. [\[PubMed\]](#)
3. Jayson, G.C.; Kohn, E.C.; Kitchener, H.C.; Ledermann, J.A. Ovarian cancer. *Lancet* **2014**, *384*, 1376–1388. [\[PubMed\]](#)
4. Howlander, N.; Noone, A.M.; Krapcho, M.; Miller, D.; Brest, A.; Yu, M.; Ruhl, J.; Tatalovich, Z.; Mariotto, A.; Lewis, D.R.; et al. (Eds.) *SEER Cancer Statistics Review, 1975–2016*; based on November 2018 SEER data submission; National Cancer Institute: Bethesda, MD, USA, 2019.
5. Lech, A.; Daneva, T.; Pashova, S.; Gagov, H.; Crayton, R.; Kukwa, W.; Czarnecka, A.M.; Szczylik, C. Ovarian cancer as a genetic disease. *Front. Biosci.* **2013**, *18*, 543–563.
6. Eng, K.H.; Weir, I.; Tsuji, T.; Odunsi, K. Immuno-stimulatory/regulatory gene expression patterns in advanced ovarian cancer. *Genes Cancer* **2015**, *6*, 399–407.
7. Mitterpergher, L. Genomic characterization of high-grade serous ovarian cancer: Dissecting its molecular heterogeneity as a road towards effective therapeutic strategies. *Curr. Oncol. Rep.* **2016**, *18*, 44.
8. Cai, S.Y.; Yang, T.; Chen, Y.; Wang, J.W.; Li, L.; Xu, M.J. Gene expression profiling of ovarian carcinomas and prognostic analysis of outcome. *J. Ovarian Res.* **2015**, *8*, 50.
9. Tassi, R.A.; Todeschini, P.; Siegel, E.R.; Calza, S.; Cappella, P.; Ardighieri, L.; Cadei, M.; Bugatti, M.; Romani, C.; Bandiera, E.; et al. FOXM1 expression is significantly associated with chemotherapy resistance and adverse prognosis in non-serous epithelial ovarian cancer patients. *J. Exp. Clin. Cancer Res.* **2017**, *36*, 63.
10. Niemi, R.J.; Braicu, E.I.; Kulbe, H.; Koistinen, K.M.; Sehouli, J.; Puistola, U.; Mäenpää, U.J.; Hilvo, M. Ovarian tumours of different histologic type and clinical stage induce similar changes in lipid metabolism. *Br. J. Cancer* **2018**, *119*, 847–854.
11. Nickels, J.T. New links between lipid accumulation and cancer progression. *J. Biol. Chem.* **2018**, *293*, 6635–6636.
12. Xu, Y. Lysophospholipid signaling in the epithelial ovarian cancer tumor microenvironment. *Cancers* **2018**, *10*, 227.
13. Zhao, G.; Cardenas, H.; Matei, D. Ovarian cancer-why lipids matter. *Cancers* **2019**, *10*, 1870.
14. Braicu, E.I.; Darb-Esfahani, S.; Schmitt, W.D.; Koistinen, K.M.; Heiskanen, L.; Pöhö, P.; Budczies, J.; Kuhberg, M.; Dietel, M.; Frezza, C.; et al. High-grade ovarian serous carcinoma patients exhibit profound alterations in lipid metabolism. *Oncotarget* **2017**, *8*, 102912–102922. [\[PubMed\]](#)
15. Chen, R.R.; Yung, M.M.; Xuan, Y.; Zhan, S.; Leung, L.L.; Liang, R.R.; Leung, T.H.Y.; Yang, H.; Xu, D.; Sharma, R.; et al. Targeting of lipid metabolism with a metabolic inhibitor cocktail eradicates peritoneal metastases in ovarian cancer cells. *Commun. Biol.* **2019**, *2*, 281.
16. Cui, T.; Zhang, L.; Huang, Y.; Yi, Y.; Tan, P.; Zhao, Y.; Hu, Y.; Xu, L.; Li, E.; Wang, D. MNDR v2.0: An updated resource of ncRNA–disease associations in mammals. *Nucleic Acids Res.* **2017**, *46*, D371–D374.
17. Liang, Z.Y.; Lai, H.Y.; Yang, H.; Zhang, C.J.; Yang, H.; Wei, H.H.; Chen, X.X.; Zhao, Y.W.; Su, Z.D.; Li, W.C.; et al. Pro54DB: A database for experimentally verified sigma-54 promoters. *Bioinformatics* **2017**, *33*, 467–469.
18. Hu, B.; Zheng, L.; Long, C.; Song, M.; Li, T.; Yang, L.; Zuo, Y. EmExplorer: A database for exploring time activation of gene expression in mammalian embryos. *Open Biol.* **2019**, *9*, 190054.
19. Wang, Z.; Gerstein, M.; Snyder, M. RNA-Seq: A revolutionary tool for transcriptomics. *Nat. Rev. Genet.* **2009**, *10*, 57–63.
20. Yang, H.; Lv, H.; Ding, H.; Chen, W.; Lin, H. iRNA-2OM: A sequence-based predictor for identifying 2'-O-methylation sites in homo sapiens. *J. Comput. Biol.* **2018**, *25*, 1266–1277.
21. Safonov, A.; Jiang, T.; Bianchini, G.; Györfy, B.; Karn, T.; Hatzis, C.; Pusztai, L. Immune gene expression is associated with genomic aberrations in breast cancer. *Cancer Res.* **2017**, *77*, 3317–3324.
22. Yang, L.; Wang, S.; Zhang, Q.; Pan, Y.; Lv, Y.; Chen, X.; Zuo, Y.; Hao, D. Clinical significance of the immune microenvironment in ovarian cancer patients. *Mol. Omics* **2018**, *14*, 341–351. [\[CrossRef\]](#) [\[PubMed\]](#)
23. Yang, R.; Xiong, J.; Deng, D.; Wang, Y.; Liu, H.; Jiang, G.; Peng, Y.; Peng, X.; Zeng, X. An integrated model of clinical information and gene expression for prediction of survival in ovarian cancer patients. *Transl. Res.* **2016**, *172*, 84–95. [\[CrossRef\]](#) [\[PubMed\]](#)
24. Willis, S.; Villalobos, V.M.; Gevaert, O.; Abramovitz, M.; Williams, C.; Sikic, B.I.; Leyland-Jones, B. Single gene prognostic biomarkers in ovarian cancer: A meta-analysis. *PLoS ONE* **2016**, *11*, e0149183. [\[CrossRef\]](#) [\[PubMed\]](#)

25. Wang, L.; Wang, L.; Ma, L.; Liu, J.; Ma, S. Identifying gene signature for the detection of ovarian cancer based on the achieved related genes. *Gynecol. Obstet. Investig.* **2017**, *82*, 361–370. [[CrossRef](#)] [[PubMed](#)]
26. Sabatier, R.; Finetti, P.; Bonense, J.; Jacquemier, J.; Adelaide, J.; Lambaudie, E.; Viens, P.; Birnbaum, D.; Bertucci, F. A seven-gene prognostic model for platinum-treated ovarian carcinomas. *Br. J. Cancer* **2011**, *105*, 304–311. [[CrossRef](#)] [[PubMed](#)]
27. Millstein, J.; Budden, T.; Goode, E.L.; Anglesio, M.S.; Talhouk, A.; Intermaggio, M.P.; Leong, H.S.; Chen, S.; Elatre, W.; Gilks, B.; et al. Prognostic gene expression signature for high-grade serous ovarian cancer. *Ann. Oncol.* **2020**, *31*, 1240–1250. [[CrossRef](#)]
28. Royston, P.; Parmar, M.K.B. Restricted mean survival time: An alternative to the hazard ratio for the design and analysis of randomized trials with a time-to-event outcome. *BMC Med. Res. Methodol.* **2013**, *13*, 152. [[CrossRef](#)]
29. Herzog, T.J. Recurrent ovarian cancer: How important is it to treat to disease progression? *Clin. Cancer Res.* **2004**, *10*, 7439–7449. [[CrossRef](#)]
30. Nagasawa, S.; Ikeda, K.; Horie-Inoue, K.; Sato, S.; Itakura, A.; Takeda, S.; Hasegawa, K.; Inoue, S. Systematic identification of characteristic genes of ovarian clear cell carcinoma compared with high-grade serous carcinoma based on RNA-sequencing. *Int. J. Mol. Sci.* **2019**, *20*, 4330. [[CrossRef](#)]
31. Seborova, K.; Vaclavikova, R.; Soucek, P.; Elsnerova, K.; Bartakova, A.; Cernaj, P.; Bouda, J.; Rob, L.; Hruda, M.; Dvorak, P. Association of ABC gene profiles with time to progression and resistance in ovarian cancer revealed by bioinformatics analyses. *Cancer Med.* **2019**, *8*, 606–616. [[CrossRef](#)]
32. Zhang, Z.; Huang, K.; Gu, C.; Zhao, L.; Wang, N.; Wang, X.; Zhao, D.; Zhang, C.; Lu, Y.; Meng, Y. Molecular subtyping of serous ovarian cancer based on multi-omics data. *Sci. Rep.* **2016**, *6*, 26001. [[CrossRef](#)] [[PubMed](#)]
33. Tothill, R.W.; Tinker, A.V.; George, J.; Brown, R.; Fox, S.B.; Lade, S.; Johnson, D.S.; Trivett, M.K.; Etemadmoghadam, D.; Locandro, B.; et al. Novel molecular subtypes of serous and endometrioid ovarian cancer linked to clinical outcome. *Clin. Cancer Res.* **2008**, *14*, 5198–5208. [[CrossRef](#)] [[PubMed](#)]
34. Zhang, D.; Chen, P.; Zheng, C.H.; Xia, J. Identification of ovarian cancer subtype-specific network modules and candidate drivers through an integrative genomics approach. *Oncotarget* **2016**, *7*, 4298–4309. [[CrossRef](#)] [[PubMed](#)]
35. Clemente, G.C.; Mihm, C.M.; Zurrida, S.; Collini, P.; Cascineelli, N. Prognostic value of tumor infiltrating lymphocytes in the vertical growth phase of primary cutaneous melanoma. *Cancer* **1996**, *77*, 1303–1310. [[CrossRef](#)]
36. Rozek, L.S.; Schmit, S.L.; Greenson, J.K.; Tomsho, L.P.; Rennert, H.S.; Rennert, G.; Gruber, S.B. Tumor-infiltrating lymphocytes, crohn's-like lymphoid reaction, and survival from colorectal cancer. *J. Natl. Cancer Inst.* **2016**, *108*, djw027. [[CrossRef](#)] [[PubMed](#)]
37. Naito, Y.; Saito, K.; Shiiba, K.; Ohuchi, A.; Saigenji, K.; Nagura, H.; Ohtani, H. CD8+ T cells infiltrated within cancer cell nests as a prognostic factor in human colorectal cancer. *Cancer Res.* **1998**, *58*, 3491–3494.
38. Nakano, O.; Sato, M.; Naito, Y.; Suzuki, K.; Orikasa, S.; Aizawa, M.; Suzuki, Y.; Shintaku, I.; Nagura, H.; Ohtani, H. Proliferative activity of intratumoral CD8(+) T-lymphocytes as a prognostic factor in human renal cell carcinoma: Clinicopathologic demonstration of antitumor immunity. *Cancer Res* **2001**, *61*, 5132–5136.
39. Sato, E.; Olson, S.H.; Ahn, J.; Bundy, B.; Nishikawa, H.; Qian, F.; Junbluth, A.A.; Frosina, D.; Gnjjatic, S.; Ambrosone, C.; et al. Intraepithelial CD8+ tumor-infiltrating lymphocytes and a high CD8+/regulatory T cell ratio are associated with favorable prognosis in ovarian cancer. *Proc. Natl. Acad. Sci. USA* **2005**, *102*, 18538–18543. [[CrossRef](#)]
40. Diederichsen, A.C.; Hjelmborg, J.V.B.; Christensen, P.B.; Zeuthen, J.; Fenger, C. Prognostic value of the CD4+/CD8+ ratio of tumour infiltrating lymphocytes in colorectal cancer and HLA-DR expression on tumour cells. *Cancer Immunol. Immunother.* **2003**, *52*, 423–428. [[CrossRef](#)]
41. Wefers, C.; Duiveman-de Boer, T.; Zusterzeel, P.L.M.; Massuger, L.F.A.G.; Fuchs, D.; Torensma, R.; Wheelock, C.E.; de Vries, I.J.M. Different lipid regulation in ovarian cancer: Inhibition of the immune system. *Int. J. Mol. Sci.* **2018**, *19*, 273. [[CrossRef](#)]
42. Schalkwijk, J.; Wiedow, O.; Hirose, S. The trappin gene family: Proteins defined by an N-terminal transglutaminase substrate domain and a C-terminal four-disulphide core. *Biochem. J.* **1999**, *340*, 569–577. [[CrossRef](#)] [[PubMed](#)]
43. Clauss, A.; Lilja, H.; Lundwall, Å. A locus on human chromosome 20 contains several genes expressing protease inhibitor domains with homology to whey acidic protein. *Biochem. J.* **2002**, *368*, 233–242. [[CrossRef](#)] [[PubMed](#)]

44. Labidi-Galy, S.I.; Clauss, A.; Ng, V.; Duraisamy, S.; Elias, K.M.; Piao, H.Y.; Bilal, E.; Davidowitz, R.A.; Lu, Y.; Badalian-Very, G.; et al. Elafin drives poor outcome in high-grade serous ovarian cancers and basal-like breast tumors. *Oncogene* **2015**, *34*, 373–383. [[CrossRef](#)] [[PubMed](#)]
45. Clauss, A.; Ng, V.; Liu, J.; Piao, H.; Russo, M.; Vena, N.; Sheng, Q.; Hirsch, M.S.; Bonome, T.; Matulonis, U.; et al. Overexpression of elafin in ovarian carcinoma is driven by genomic gains and activation of the nuclear factor κ B pathway and is associated with poor overall survival. *Neoplasia* **2010**, *12*, 161–172. [[CrossRef](#)]
46. Wei, H.; Hellström, K.E.; Hellström, I. Elafin selectively regulates the sensitivity of ovarian cancer cells to genotoxic drug-induced apoptosis. *Gynecol. Oncol.* **2012**, *125*, 727–733. [[CrossRef](#)]
47. Bou, D.J.; Lee, J.K. Regulator of G-Protein signaling 1 (RGS1). In *Encyclopedia of Signaling Molecules*, 2nd ed.; Choi, S., Ed.; Springer International Publishing: Basel, Switzerland, 2018.
48. Sethakorn, N.; Dulin, N.O. RGS expression in cancer: Oncomining the cancer microarray data. *J. Recept. Signal. Transduct. Res.* **2013**, *33*, 166–171. [[CrossRef](#)]
49. Wong, Y.F.; Cheung, T.H.; Tsao, G.S.; Lo, K.W.; Yim, S.F.; Wang, V.W.; Heung, M.M.; Chan, S.C.; Chan, L.K.; Ho, T.W.; et al. Genome-wide gene expression profiling of cervical cancer in Hong Kong women by oligonucleotide microarray. *Int. J. Cancer* **2006**, *118*, 2461–2469. [[CrossRef](#)]
50. Fredholm, B.B.; IJzerman, A.P.I.; Jacobson, K.A.; Klotz, K.N.; Linden, J. International Union of Pharmacology–XXV. Nomenclature and classification of adenosine receptors. *Pharmacol. Rev.* **2001**, *53*, 527–552.
51. Riccardo, F.; Arigoni, M.; Buson, G.; Zago, E.; Iezzi, M.; Longo, D.; Carrara, M.; Fiore, A.; Nuzzo, S.; Bicciato, S.; et al. Characterization of a genetic mouse model of lung cancer: A promise to identify non-small cell lung cancer therapeutic targets and biomarkers. *BMC Genom.* **2014**, *15*, S1. [[CrossRef](#)]
52. Doms, A.; Sanabria, T.; Hansen, J.N.; Altan-Bonnet, N.; Holm, G.H. 25-Hydroxycholesterol production by the cholesterol-25-Hydroxylase interferon-stimulated gene restricts mammalian Reovirus infection. *J. Virol.* **2018**, *92*, e01047-18. [[CrossRef](#)]
53. Bauman, D.R.; Bitmansour, A.D.; McDonald, J.G.; Thompson, B.M.; Liang, G.; Russell, D.W. 25-Hydroxycholesterol secreted by macrophages in response to toll-like receptor activation suppresses immunoglobulin A production. *Proc. Natl. Acad. Sci. USA* **2009**, *106*, 16764–16769. [[CrossRef](#)] [[PubMed](#)]
54. Mittempergher, L.; Saghatchian, M.; Wolf, D.M.; Michiels, S.; Canisius, S.; Dessen, P.; Delaloge, S.; Lazar, V.; Benz, S.C.; Tursz, T.; et al. A gene signature for late distant metastasis in breast cancer identifies a potential mechanism of late recurrences. *Mol. Oncol.* **2013**, *7*, 987–999. [[CrossRef](#)] [[PubMed](#)]
55. Sugimoto, Y.; Narumiya, S. Prostaglandin E receptors. *J. Biol. Chem.* **2007**, *282*, 11613–11617. [[CrossRef](#)] [[PubMed](#)]
56. Czogalla, B.; Kuhn, C.; Heublein, S.; Schmöckel, E.; Mayr, D.; Kolben, T.; Trillsch, F.; Burges, A.; Mahner, S.; Jeschke, U.; et al. EP3 receptor is a prognostic factor in TA-MUC1-negative ovarian cancer. *J. Cancer Res. Clin. Oncol.* **2019**, *145*, 2519–2527. [[CrossRef](#)] [[PubMed](#)]
57. Heidegger, H.; Dietlmeier, S.; Ye, Y.; Kuhn, C.; Vattai, A.; Aberl, C.; Jeschke, U.; Mahner, S.; Kost, B. The prostaglandin EP3 receptor is an independent negative prognostic factor for cervical cancer patients. *Int. J. Mol. Sci.* **2017**, *18*, 1571. [[CrossRef](#)]
58. Zhu, J.; Trillsch, F.; Mayr, D.; Kuhn, C.; Rahmeh, M.; Hofmann, S.; Vogel, M.; Mahner, S.; Jeschke, U.; Schönfeldt, V. Prostaglandin receptor EP3 regulates cell proliferation and migration with impact on survival of endometrial cancer patients. *Oncotarget* **2018**, *9*, 982–994. [[CrossRef](#)]
59. Semmlinger, A.; von Schoenfeldt, V.; Wolf, V.; Meuter, A.; Kolben, T.M.; Kolben, T.; Zeder-Goes, C.; Weis, F.; Gallwas, J.; Wuerstlein, R.; et al. EP3 (prostaglandin E2 receptor 3) expression is a prognostic factor for progression-free and overall survival in sporadic breast cancer. *BMC Cancer* **2018**, *18*, 431. [[CrossRef](#)]
60. Gentles, A.J.; Newman, A.M.; Liu, C.L.; Bratman, S.V.; Feng, W.; Kim, D.; Nair, V.S.; Xu, Y.; Khuong, A.; Hoang, C.D.; et al. The prognostic landscape of genes and infiltrating immune cells across human cancers. *Nat. Med.* **2015**, *21*, 938–945. [[CrossRef](#)]
61. Li, T.; Fan, J.; Wang, B.; Traugh, N.; Chen, Q.; Liu, J.S.; Li, B.; Liu, X.S. TIMER: A web server for comprehensive analysis of tumor-infiltrating immune cells. *Cancer Res.* **2017**, *77*, e108–e110. [[CrossRef](#)]
62. Ritchie, M.E.; Phipson, B.; Wu, D.; Hu, Y.; Law, C.W.; Shi, W.; Smyth, G.K. Limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res.* **2015**, *43*, e47. [[CrossRef](#)]
63. Yu, G.; Wang, L.G.; Han, Y.; He, Q.Y. ClusterProfiler: An R package for comparing biological themes among gene clusters. *Omics* **2012**, *16*, 284–287. [[CrossRef](#)] [[PubMed](#)]
64. Tibshirani, R. Regression shrinkage and selection via the lasso. *J. R. Stat. Soc. B* **1996**, *58*, 267–288. [[CrossRef](#)]

65. Friedman, J.; Hastie, T.; Tibshirani, R. Regularization paths for generalized linear models via coordinate descent. *J. Stat. Softw.* **2010**, *33*, 1–22. [[CrossRef](#)] [[PubMed](#)]
66. Remmele, W.; Stegner, H.E. Recommendation for uniform definition of an immunoreactive score (IRS) for immunohistochemical estrogen receptor detection (ER-ICA) in breast cancer tissue. *Pathologe* **1987**, *8*, 138–140. [[PubMed](#)]

Publisher’s Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



© 2020 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).