

Immune and Stromal Score-Related Gene Signatures For Risk Stratification of Gynaecologic Cancers

Li Gao

The First Affiliated Hospital of Guangxi Medical University

Zi-Qian Liang

The First Affiliated Hospital of Guangxi Medical University

Zhu-Xin Wei

The First Affiliated Hospital of Guangxi Medical University

Jun-Hong Chen

Guangxi Maternal and Child Health Hospital

Jing-Jing Zeng

The First Affiliated Hospital of Guangxi Medical University

Gang Chen (✉ chengang@gxmu.edu.cn)

Guangxi Medical University First Affiliated Hospital <https://orcid.org/0000-0002-4864-1451>

Primary research

Keywords: gynaecologic cancers, immune score, stromal score, tumour microenvironment, ESTIMATE

Posted Date: April 12th, 2021

DOI: <https://doi.org/10.21203/rs.3.rs-370788/v1>

License: © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Abstract

Background: The immune and stromal landscape of gynaecologic cancers is far from elucidated, and there is a lack of systematic prognostic analysis of the immune or stromal parts of the tumour microenvironment (TME) in all types of gynaecologic cancers. We aimed to establish immune and stromal score-based risk signatures for each type of gynaecologic cancer.

Methods: In the present study, using the algorithm of “Estimation of stromal and immune cells in malignant tumours using expression data”, we calculated immune and stromal scores for cervical squamous carcinoma (CESC), uteri corpus endometrial cancer (UCEC) and ovarian cancer (OVCA) from the cancer genome atlas (TCGA) to establish immune and stromal score-based risk signatures. Multi-omics analysis combining single nucleotide polymorphism (SNP), tumour mutation burden (TMB) and gene set variation analysis was further performed to investigate the mechanism of immune and stromal score-based risk signatures in gynaecologic cancers. Eventually, the prognostic value of some signature genes was examined in independent microarrays and aberrant expression of IHH in UCECs through immunohistochemistry.

Results: According to the results, the prognosis prediction ability of the immune and stromal score-related risk signatures constructed for pan-gynaecological cancers withstood the tests of tROC curves and Kaplan-Meier survival curves, proving to be effective in risk stratification. Three signature genes, including BNC1, CCDC80, and IHH, were verified in independent datasets to have prognostic value in gynaecologic cancers.

Conclusions: In summary, the immune and stromal score-related risk signature constructed in this study might be promising prognostic markers for pan-gynaecological cancers.

Background

Gynaecologic cancers are a set of cancers that arise from female reproductive organs, including the uterus, ovary, cervix, vagina, fallopian tube and vulva [1]. This group of tumours have similar embryonic origins from Mullerian ducts and was affected by hormones [2]. As one of the most frequent neoplasms worldwide, gynaecologic cancers are the main contributors to cancer-related mortality in women, taking a severe toll particularly on women in developing countries [3]. From the perspectives of incidence and mortality, cervix uteri neoplasms and OVCA are ranked as the deadliest gynaecological cancers [4]. Accurate prognosis evaluation of patients suffering from gynaecologic cancers was an important segment of oncological care, which propelled us to identify effective prognostic biomarkers for gynaecological cancers.

It has been gradually recognised in the past two decades that the tumour microenvironment (TME) plays an essential role in governing the biological behaviours of human cancers. Non-tumour components in the TME, such as immune and stromal cells, have emerged as promising predictors of prognosis for a wide variety of human cancers, including gynaecological cancers. Hinchcliff et al. discovered that

lymphocyte-specific kinase presented significant prognostic value in OVCA [5]. Ittiamornlert reported that the neutrophil-lymphocyte ratio could predict the clinical outcomes for stage IVB, persistent, or recurrent cervical cancer patients after receiving chemotherapy [6]. In endometrial cancer, intraepithelial CD8 + tumour-infiltrating T-lymphocytes indicate a favourable prognosis [7]. Despite great research progress on the crosstalk between non-tumour components of the microenvironment and cancer cells of gynaecological cancers, the immune and stromal landscape of gynaecologic cancers was far from elucidated, and there lacks systematic analysis of the prognostic value of immune or stromal parts of TME in all types of gynaecologic cancers.

In the present study, using the “Estimation of stromal and immune cells in malignant tumours using expression data (ESTIMATE)” algorithm proposed by Kosuke Yoshihara et al. [8], we calculated immune and stromal scores for cervical squamous carcinoma (CESC), uteri corpus endometrial cancer (UCEC) and ovarian cancer (OVCA) from the cancer genome atlas (TCGA) to compare the commonness and differences between the TME landscape of gynaecologic cancers, based on which immune and stromal score-based risk signature was established for each type of gynaecologic cancer. Multi-omics analysis combining gene set variation analysis (GSVA), single nucleotide polymorphism (SNP) and tumour mutation burden (TMB) was further performed to investigate the mechanism of immune and stromal score-based risk signature in gynaecologic cancers. Immunohistochemistry (IHC) was performed to evaluate the differential expression of one of the signature genes, IHH, in UCEC and non-cancerous endometrium tissues.

Material And Methods

1. Data download

RNA-seq data in the form of fragments per kilobase million (FPKM), count data, and clinical information on CESC, OVCA and UCEC were available from the Genomic Data Commons (GDC) Data Portal of the TCGA database. Normalised RNA-seq data in the form of $\log_2(\text{TPM}+0.001)$ was reckoned from FPKM RNA-seq data and was utilised for calculating immune and stromal scores. Each sample of gynaecologic cancers was assigned an immune score and a stromal score through the ESTIMATE package in R software v.3.6.1. Because the sample size of uterine carcinosarcoma in the TCGA database was not large enough for subsequent prognostic analysis, we did not include it in the current work. Independent sample's *t*-test and a Kruskal-Wallis test were performed in GraphpadPrism v.8.0.1. for comparing immune scores between different groups of clinical variables of the three gynaecologic cancers. Thresholds of statistical significance were $P < 0.05$.

Differentially expressed genes (DEG) related to immune and stromal scores

All samples of the three gynaecologic cancers were divided into high and low immune or stromal score groups according to the median value of immune and stromal scores of all samples. DEGs between the two groups of immune and stromal scores were identified using the count data of the three gynaecologic

cancers. The limma voom package in R software v.3.6.1 was employed for identification of DEGs [9]. Adjusted P value < 0.05 and $|\log_2FC| > 1$ were considered as the threshold for DEGs.

2. DEGs with prognostic value

Patients of the three gynaecologic cancers with more than 90 days of overall survival time were reserved for further survival analysis. Intersected genes of significantly upregulated DEGs from the high immune score group and significantly upregulated DEGs in the high stromal score group were submitted to univariate Cox regression analysis for selection of prognostic DEGs in three gynaecologic cancers ($P < 0.05$), which was conducted via survival package in R software v.3.6.1. We further conducted gene ontology (GO) analysis via the ClusterProfiler package of R software v.3.6.1 to investigate the functional enrichment of prognostic DEGs in biological process (BP), cellular component (CC) and molecular function (MF). Thresholds of significant GO terms were $P < 0.05$.

3. Immune and stromal score-based risk signature for pan-gynaecologic cancers

For each of the three gynaecologic cancers, multivariate Cox regression analysis was conducted in SPSS v.22.0 to choose the signature genes of the prognostic model. Significant prognostic DEGs from multivariate Cox regression analysis were defined as the signature genes ($P < 0.01$ for CESC, $P < 0.05$ for OVCA and UCEC). Immune and stromal score-based risk signatures composed of significant prognostic DEGs were built for three gynaecologic cancers with the formula: $\text{risk score} = \sum_i^n \log_2(TPM + 0.001) * \beta_i$ (β_i was the regression coefficient from multivariate Cox regression analysis). The prognostic prediction power of the risk signature was appraised by Kaplan-Meier survival curves and time-dependent receiver's operating characteristics (tROC) curves [10]. Independent sample *t*-tests and Kruskal-Wallis tests were performed in GraphpadPrism v.8.0.1. to compare risk scores between different groups of clinical variables of the three gynaecologic cancers. $P < 0.05$ was considered significant.

4. Immune infiltration of common prognostic DEGs in gynaecologic cancers

We took intersection of significant prognostic DEGs from multivariate Cox regression analysis of the three gynaecologic cancers. For genes from intersected parts, we consulted Tumour Immune Estimation Resource (TIMER, <https://cistrome.shinyapps.io/timer/>) to analyse the correlation between the expression of common prognostic DEGs and the tumour purity or the infiltration level of various immune cells in the three gynaecologic cancers.

5. Construction of nomogram combining immune-related risk signature and clinical features

The independence of the immune-related risk signature was evaluated by univariate and multivariate Cox regression analysis, considering the clinical features of the three gynaecologic cancers. Clinical features significantly correlated with survival from univariate Cox regression analysis ($P < 0.05$) and immune-related risk signature were combined into nomogram. Calibration curves were plotted to assess the prediction accuracy of the nomogram against real survival conditions at three and five years. Univariate

and multivariate Cox regression analyses for this part were carried out in SPSS v.22.0. Nomogram and calibration curves were drawn by Rms, foreign and survival packages in R software v.3.6.1.

6. GSVA analysis

Enrichment analysis of the KEGG pathway for three gynaecologic cancer samples included for survival analysis was analysed via GSVA package in R software v.3.6.1. Enrichment scores of KEGG pathways against in each sample were calculated according to the $\log_2(\text{TPM}+0.001)$ normalised expression data. The file "c2.cp.kegg.v6.2.symbols.gmt" from the Molecular Signature Database (MSigDB, <http://software.broadinstitute.org/gsea/msigdb/index.jsp>) was the reference gene set. The limma package in R software v.3.6.1 was utilised to screen out significant KEGG pathways between high- and low-risk patients (adjusted $P < 0.05$).

7. Mutation analysis for gynaecologic cancers

TMB refers to the rate of somatic nonsynonymous mutations occurring in the coding area of sequenced DNA per megabase [11]. The TMB of low- and high-risk gynaecologic cancer samples was estimated by the MutSigCV algorithm based on the whole exome sequencing data of the three gynaecologic cancers from TCGA [12]. The mean TMB of samples with available mutation data was set as the cutoff value for allocating gynaecologic patients to low and high TMB groups. Differential distribution of risk scores in high or low TMB groups of three gynaecologic cancers was compared by independent sample's t -test in GraphpadPrism v.8.0.1. SNP data of the three gynaecologic cancers were inferred from the mutation annotation files in TCGA using the method of VarScan in the TCGA database. A waterfall plot was created using "GenomeInfoDbData" and "GenVisR" packages in R software v.3.6.1 to illustrate the mutation type and its translational effect of signature genes in three gynaecologic cancers.

8. Validation of immune-related risk signatures for pan-gynaecologic cancers

GSE44001 and GSE9891 were two microarrays that contained disease-free survival information of 300 CESC patients and overall survival information of 285 OVCA patients. Part of the signature genes for CESC and OVCA were available in the two microarrays. The two microarrays served as the validation set for CESC and OVCA. Kaplan-Meier survival curves were plotted to assess the prognostic value of part of the signature genes in the two validation datasets. The median value of the normalised expression value was set as the threshold. No microarray with survival time and survival status for UCEC patients was found. To remedy the problem, we examined the expression value of part of the signature genes in 21 survivors and 24 non-survivors diagnosed with UCEC from GSE21882. Expression difference was measured by independent sample's t -test in SPSS v.22.0. $P < 0.05$ was considered statistically significant.

9. IHC of IHH in UCEC and paracarcinoma tissues

A total of 215 UCEC tissue samples and 48 non-cancer endometrium tissue samples were included for tissue microarray by Pantomics, Inc. Part of the samples (20 UCEC tissues and 15 non-cancer endometrium tissues) were collected from the UCEC patients of the First Affiliated Hospital of Guangxi

Medical University. Permission for the study was obtained from the ethics committee of the First Affiliated Hospital of Guangxi Medical University, and all patients signed informed consent forms.

The procedure of IHC experiments and the assessment criteria of the tissue slices were described in a previous study [13]. Expression data were analysed by an independent sample's *t*-test in GraphpadPrism. $P < 0.05$ was considered statistically significant.

Results

1. Immune scores and stromal scores for gynaecologic cancers

Immune and stromal scores were deduced from 306 CESC samples, 379 OVCA samples and 552 UCEC samples from TCGA. Intersected upregulated and downregulated immune or stromal score-related genes in the three gynaecologic cancers were displayed in a set of Venn plots (Additional file 1). Differential expression of these DEGs in high and low immune or stromal score groups is exhibited in Additional file 2. Because the intersection part of upregulated immune and stromal score-related DEGs occupied the majority of DEGs, we selected these DEGs to perform further survival analysis.

2. DEGs with prognostic value

According to the selection criteria, a total of 211 CESC patients, 351 OVCA patients and 456 UCEC patients were eligible for survival analysis. Altogether, 162 DEGs, 78 DEGs and 64 DEGs showed prognostic significance for CESC, OVCA and UCEC, respectively ($P < 0.05$). BP, CC and MF terms significantly assembled by these DEGs are presented in Additional file 3 and Additional file 4 and 5, which were mainly correlated with the immune system.

3. Immune and stromal score-based risk signature for pan-gynaecologic cancers

Three immune score-based risk signatures were constructed for each of the three gynaecologic cancers. Univariate and multivariate regression coefficients of signature genes for the three gynaecologic cancers are listed in Table 1. As revealed by 5-year tROC curves, the three immune and stromal score-based risk signatures all presented strong prognosis prediction ability, with the immune-score-based risk signature for CESC performing best (AUC = 0.854; Fig. 1). Kaplan-Meier survival curves indicated that high- or low-risk gynaecologic cancer patients stratified by immune and stromal score-based risk signature showed significantly different survival outcomes ($P < 0.001$, $P < 0.001$, $P = 0.003$; Fig. 2). Moreover, the immune and stromal score-based risk signatures were significantly associated with the worse clinical progression of gynaecologic cancers. CESC patients with distant metastasis, higher clinical stage and higher T stage had significantly higher risk scores than patients in control groups (Additional file 6A-C). OVCA patients with younger age, venous invasion and higher clinical stage had significantly higher risk scores than patients in control groups (Additional file 6D-F). UCEC patients with elder age, serous type, higher histological grade and higher clinical stage had significantly higher risk scores than patients in control groups (Additional file 6G-J).

Table 1. Univariate and multivariate Cox regression results for risk signature genes of three gynaecologic cancers

Cancer type	Gene ID	univariate Cox regression analysis				multivariate Cox regression analysis			
		HR	LL	UL	p value	HR	LL	UL	p value
CESC	RASAL3	0.820	0.690	0.974	0.024	3.48	1.614	7.505	0.001
CESC	TRGV3	0.774	0.647	0.926	0.005	0.861	0.781	0.95	0.003
CESC	SELP	0.869	0.767	0.985	0.028	0.784	0.66	0.931	0.006
CESC	CHST11	0.916	0.860	0.975	0.006	2.28	1.72	3.021	<0.001
CESC	HAVCR2	0.721	0.627	0.829	0.000	0.369	0.229	0.594	<0.001
CESC	TRBV28	1.269	1.029	1.564	0.026	0.351	0.225	0.546	<0.001
CESC	ARHGAP15	0.827	0.692	0.987	0.036	3.703	1.96	6.997	<0.001
CESC	SLITRK4	0.786	0.686	0.900	0.001	1.235	1.119	1.363	<0.001
CESC	TRBV124	0.842	0.711	0.998	0.047	0.888	0.813	0.97	0.008
OVCA	HTRA4	0.925	0.871	0.984	0.013	0.838	0.986	0.909	0.021
OVCA	BNC1	1.056	1.007	1.108	0.026	1.006	1.118	1.06	0.030
OVCA	CCDC80	1.123	1.020	1.236	0.018	1.056	1.333	1.187	0.004
OVCA	LILRA2	1.112	1.001	1.236	0.048	1.141	1.49	1.304	<0.001
OVCA	TRBV103	0.936	0.903	0.970	0.000	0.905	0.988	0.945	0.012
OVCA	IGHV371	0.954	0.922	0.986	0.005	0.926	0.998	0.961	0.041
OVCA	DIO3	0.946	0.899	0.995	0.031	0.881	0.988	0.933	0.017
UCEC	IFNG	0.902	0.844	0.964	0.002	0.882	0.824	0.944	<0.001
UCEC	IHH	0.909	0.853	0.969	0.004	0.893	0.839	0.95	<0.001

Note: CESC: cervical squamous cell carcinoma; OVCA: ovarian cancer; UCEC: uteri corpus endometrial cancer; HR: hazard ratio; LL: lower limit; UL: upper limit

4. Immune infiltration of common prognostic DEGs in gynaecologic cancers

Six genes, TRAV9-2, CD3G, TRBC2, LINC00861, MS4A1 and GIMAP7, were common genes of prognostic DEGs from three gynaecologic cancers. Three of the genes were available query terms in TIMER (CD3G, GIMAP7 and MS4A1). All three genes were negatively correlated with tumour purity and positively

correlated with infiltration of B cells, CD8+ T cells, CD4+ T cells, macrophages, neutrophils, and dendritic cells ($P < 0.05$, Additional file 7 and 8).

Construction of nomogram combining immune and stromal-related risk signature and clinical features

Univariate analysis and multivariate analysis for immune and stromal score-related risk signature and other clinical variables of CESC suggested that the prognostic value of immune and stromal score-related risk signature for CESC was not affected by clinical parameters ($P < 0.05$, Table 2). Although immune and stromal score-based risk signatures for OVCA and UCEC showed insignificant prognostic value in the coexistence of other clinical features, they were significant indicators of worse prognosis from univariate Cox regression analysis ($P < 0.05$, Table 3 and 4). Since there were too many missing values of lymph node metastasis and distant metastasis for CESC, we excluded the two variables from the construction of the nomogram. Only immune and stromal score-related risk signature, clinical stage and T stage were component variables for the nomogram of CESC. With regard to OVCA and UCEC, immune and stromal score-related risk signatures and all significant clinical variables from univariate Cox regression analysis were combined to construct a nomogram. Nomograms for the three gynaecologic cancers all highlighted the importance of immune and stromal score-related risk signatures in prognosis prediction of CESC, OVCA and UCEC (Additional file 9). Calibration curves at three and five years proved the accuracy and stability of the combined nomograms for CESC and UCEC (Fig. 3).

Table 2. Univariate and multivariate Cox regression analysis of the clinical parameters and immune and stromal-score-related risk signature for CESC patients

Clinical variables	Group	Univariate		Multivariate	
		HR (95% CI)	p-value	HR (95% CI)	p-value
Immune and stromal score-related risk signature	low	9.635	<0.001	42.652	0.014
	high	(4.784-19.406)		(2.114-860.720)	
Keratinizing squamous cell carcinoma present indicator	Keratinizing squamous cell carcinoma	1.323	0.450	1.702	0.679
	Non-keratinizing squamous cell carcinoma	(0.640-2.734)		(0.137-21.099)	
Neoplasm histologic grade	1-2	0.782	0.265	2.622	0.301
	3-4	(0.508-1.205)		(0.423-16.266)	
Age	≥50	1.231	0.434	1.483	0.796
	<50	(0.731-2.073)		(0.075-29.271)	
Clinical stage	I-II	2.445	0.001	0.548	0.809
	III-IV	(1.432-4.175)		(0.004-72.260)	
T stage	1-2	3.652	<0.001	45.288	0.777
	3-4	(1.831-7.285)		(0-1.29E+13)	
Lymph node metastasis	no	2.750	0.013	2.222	0.393
	yes	(1.233-6.134)		(0.356-13.862)	
Distant metastasis	no	3.645	0.024	2.053	0.999
	yes	(1.184-11.226)		(0-NA)	

Note: CI: confidence interval; HR: hazard ratio; NA: not available

Table 3. Univariate and multivariate Cox regression analysis of the clinical parameters and immune and stromal-score-related risk signature for OVCA patients

Clinical variables	Group	Univariate		Multivariate	
		HR (95% CI)	p-value	HR (95% CI)	p-value
Immune and stromal score-related risk signature	low	2.846	<0.001	4.274	0.153
	high	(2.095-3.865)		(0.582-31.370)	
Race	White	0.842	0.020	0.679	0.381
	Black or African American	(0.728-0.973)		(0.286-1.614)	
	Asian				
	American Indian or Alaska native				
	Native Hawaiian or other Pacific islander				
Neoplasm histologic grade	1-2	1.316	0.205	0.895	0.923
	3-4	(0.861-2.102)		(0.094-8.563)	
Age	≥50	1.159	0.442	1.681	0.717
	<50	(0.795-1.689)		(0.101-27.836)	
Tumor residual disease	1-10 mm	0.894	0.085	0.965	0.918
	11-20 mm	(0.786-1.016)		(0.487-1.909)	
	>20 mm				
	no macroscopic disease				
Anatomic neoplasm subdivision	bilateral	0.919	0.489	0.005	0.989
	left	(0.725-1.167)		(0-NA)	
	right				

Note: CI: confidence interval; HR: hazard ratio; NA: not available

Table 4. Univariate and multivariate Cox regression analysis of the clinical parameters and immune and stromal-score-related risk signature for UCEC patients

Clinical variables	Group	Univariate		Multivariate	
		HR (95% CI)	p-value	HR (95% CI)	p-value
Immune and stromal score-related risk signature	low	2.860	0.001	1.674	0.200
	high	(1.496-5.467)		(0.762-3.681)	
Histological type	Serous endometrial carcinoma	1.393	0.037	0.864	0.522
	endometrioid endometrial carcinoma	(1.02-1.901)		(0.551-1.353)	
	Mixed serous and endometrioid				
Ethnicity	Hispanic or Latino	0.568	0.580	0.532	0.545
	Not Hispanic or Latino	(0.077-4.200)		(0.069-4.110)	
Neoplasm histologic grade	1-2	2.654	0.009	2.300	0.096
	3-4	(1.275-5.521)		(0.877-6.135)	
Age	≥50	0.772	0.623	1.761	0.370
	<50	(0.276-2.164)		(0.511-6.072)	
Clinical stage	I-II	4.315	<0.001	2.621	0.010
	III-IV	(2.373-7.846)		(1.264-5.436)	

Note: CI: confidence interval; HR: hazard ratio; NA: not available

5. GSVA analysis

A total of 23, 110 and 113 KEGG pathways were differentially enriched between high- and low-risk groups of CESC, OVCA and UCEC patients, respectively (Fig. 4). For CESC and UCEC, most of the differentially enriched KEGG pathways were notably enriched in the low-risk group, while most of the 110 KEGG pathways were notably enriched in the high-risk group of OVCA. In particular, six KEGG pathways, including ECM receptor interaction, glycosaminoglycan biosynthesis heparin sulphate, B cell receptor signalling pathway, glycosaminoglycan biosynthesis chondroitin sulphate, aldosterone regulated sodium reabsorption, and arachidonic acid metabolism, all showed significantly differential enrichment between high- and low-risk groups of three gynaecologic cancers.

6. Mutation analysis for gynaecologic cancers

We compared TMBs in high- and low-risk groups of three gynaecologic cancers. There was no significant relationship between TMB and risk of gynaecologic cancer patients (data not shown). Mutation type and mutation frequencies of signature genes for signature genes in the three gynaecologic cancers are illustrated in Additional file 10. It could be inferred from the waterfall plots that various types of mutations occurred in signature genes, and missense mutation was the predominant mutation type. RASAL3, DIO3 and IHH were the most frequently mutated signature genes (Additional file 10).

7. Validation of immune and stromal-related risk signatures for pan-gynaecologic cancers

Kaplan-Meier survival curves for eight signature genes (CHST11, HAVCR2, ASGR2, SELP, SLITRK4, ARHGAP15, P2RY14 and RASAL3) of immune and stromal score-related risk signatures for CESC suggested that none of them was a statistically prognostic factor for disease-free survival of 300 CESC patients (data not shown). Two signature genes (BNC1 and CCDC80) were significantly correlated with the inferior disease-free survival of 285 OVCA patients from GSE9891, consistent with their indication of worse survival from univariate and multivariate Cox regression analysis. As for UCEC, IHH expression was obviously higher in 21 survivors of UCEC than in 24 non-survivors of UCEC, corresponding to the prognosis-protective effect of IHH from univariate and multivariate Cox regression analysis (Additional file 11).

8. Protein expression of IHH in UCEC and paracarcinoma endometrium tissues

As illustrated in Fig. 5, IHH presented low expression in cancer cells of UCEC tissues (4.833 ± 2.863). By comparison, the expression of IHH was high in non-cancer endometrium tissues (9.917 ± 2.019) ($P < 0.001$).

Discussion

In the current work, we constructed a set of immune and stromal score-based risk signatures for gynaecologic cancers and explored the molecular mechanism underlying the immune and stromal score-related risk signature across pan-gynaecologic cancers. The prognosis prediction ability of the immune and stromal score-related risk signature withstood the tests of tROC curves and Kaplan-Meier survival curves, proving to be effective in risk stratification. Although previous researchers have worked out immune-related risk signatures for cervical cancer and OVCA [14, 15], risk signatures developed by prior studies were constructed on immune cells or immune genes imported from the ImmPort database for only one gynaecologic cancer. The highlights of the current work embodied in the methods adopted to identify immune and stromal score-related signature genes, the research scope of pan-gynaecologic cancers, and multi-omics analysis for investigating the molecular mechanism of immune and stromal score-related signatures in gynaecologic cancers.

The immune and stromal landscapes of the three gynaecologic cancers varied from each other, and there was also similarity between the immune backgrounds of the three gynaecologic cancers. Six genes, including TRAV9-2, CD3G, TRBC2, LINC00861, MS4A1 and GIMAP7, all demonstrated significant prognostic values across three gynaecologic cancers. We assumed that the six genes might serve as the pivots for connecting the immunogenetic mechanism of the immune and stromal score-related risk signatures in gynaecologic cancers. None of the six genes have been studied in gynaecologic cancers. We for the first time mined out the prognostic values of these genes in pan-gynaecologic cancers and further evaluated the close association between expression of the six genes in gynaecologic cancers and tumour purity or infiltration of immune cells in the three gynaecologic cancers. GSVA analysis for immune-score-related risk signature stratified groups of three gynaecologic cancers also provided essential clues for the crosstalk immunogenetic mechanisms underlying the initiation and progression of the three gynaecologic cancers. Three of the intersected KEGG pathways in pan-gynaecologic cancers, including ECM receptor interaction, glycosaminoglycan biosynthesis heparin sulphate, glycosaminoglycan biosynthesis chondroitin sulphate, and arachidonic acid metabolism, were actively involved in the formation and development of gynaecologic cancers. Brule S et al. reported that glycosaminoglycan/chemokine interaction was a crucial mediator of SDF-1/CXCL12-induced HeLa cell invasion [16]. The arachidonic acid metabolism process plays a critical role in human cancers, and enzymes of arachidonic acid metabolism and its metabolic products have been considered novel preventive and therapeutic targets in cancer [17]. In OVCA, arachidonic acid presented high concentrations in OVCA and emerged as an important signalling pathway in proliferation, migration or drug resistance and OVCA cells [18]. The extracellular matrix (ECM) is a crucial constituent of the cellular microenvironment and mediates normal tissue development and homeostasis [19]. It was observed by Sahoo et al. that suppressing ECM-mediated TGF- β signalling could attenuate endometrial cancer metastasis [19]. Participation of the immune and stromal score-related risk signature genes in the above KEGG pathways might explain the mechanism of risk signatures across pan-gynaecologic cancers. The different mutations that occurred in risk signature genes from SNP analysis in this study also drop hints for interpreting the differential survival conditions in high- and low-risk gynaecologic patients.

Finally, validation of the risk signatures in independent datasets was an indispensable step. Although we did not find appropriate microarrays to examine the prognostic efficiency of the immune and stromal score-related risk signatures in gynaecologic cancers, we at least verified the significant prognostic value of genes such as BNC1, CCDC80 and IHH in gynaecologic cancers. Moreover, the down-expression of IHH in UCEC was confirmed by IHC, which was in alignment with the protective effect of IHH on the prognosis of UCEC patients. IHH is a key constituent of the hedgehog (HH) signalling pathway, and the HH signalling pathway plays critical roles in embryonic development, tissue homeostasis, and cancer [20]. Previous investigations have demonstrated that IHH could function as a tumour suppressor through HH signalling pathways. Florian reported that loss of IHH might increase stromal expression of ErbB family members, thus accelerating the proliferation of colon cancer cells [21]. The research by Sahba et al. suggested that stromal HH pathway activated by IHH inhibited growth and metastasis of lung adenocarcinoma cells by restricting reactive oxygen species [22]. These findings echoed the prognostic

protection effect of IHH in UCEC and the stromal or immune-related biological computing background of IHH in UCEC.

Limitations of the present study will be narrated in the following aspects. We only included three types of gynaecologic cancers with ample prognostic data in this study (CESC, OVCA and UCEC). We would construct immune and stromal score-related risk signatures in other gynaecologic cancers, such as uterine carcinosarcoma, when we collect enough survival information to test the application of the immune and stromal score-related risk signatures in pan-gynaecologic cancers. Second, in vitro and in vivo experiments were warranted to investigate the functional roles of the signature genes on the immune or stromal-related processes of gynaecologic cancers. Third, it was necessary to validate immune and stromal score-related risk signatures built in the present study through sequencing analysis of clinical gynaecologic samples when experimental conditions permit.

Conclusion

In summary, the immune and stromal score-related risk signature constructed in this study might be promising prognostic markers for pan-gynaecologic cancers. Hub genes, including TRAV9-2, CD3G, TRBC2, LINC00861, MS4A1 and GIMAP7, as well as KEGG pathways, such as glycosaminoglycan biosynthesis chondroitin sulphate, ECM receptor interaction, glycosaminoglycan biosynthesis heparin sulphate, and arachidonic acid metabolism, may be the junction point of the immunogenetic mechanism of pan-gynaecologic cancers.

Abbreviations

TME: tumour microenvironment; CESC: cervical squamous carcinoma; UCEC: uteri corpus endometrial cancer; OVCA: ovarian cancer; TCGA: the cancer genome atlas; SNP: single nucleotide polymorphism; TMB: tumour mutation burden; ESTIMATE: estimation of stromal and immune cells in malignant tumours using expression data; GSVA: gene set variation analysis; IHC: immunohistochemistry; FPKM: fragments per kilobase million; GDC: genomic data commons; DEG: differentially expressed genes; GO: gene ontology; BP: biological process; CC: cellular component; MF: molecular function; tROC: time-dependent receiver's operating characteristics;

Declarations

Ethics approval and consent to participate

Permission for the study was obtained from the ethics committee of the First Affiliated Hospital of Guangxi Medical University, and all patients signed informed consent forms.

Consent for publication

Not applicable.

Availability of data and materials

The datasets generated and/or analysed during the current study are available in the TCGA (<https://portal.gdc.cancer.gov/>) and GEO (<https://www.ncbi.nlm.nih.gov/gds>) databases.

Competing interests

The authors declare that they have no competing interests.

Funding

This work was supported by Guangxi Zhuang Autonomous Region Health and Family Planning Commission Self-financed Scientific Research Project (Z20180979); Guangxi Medical High-level Key Talents Training "139" Program (2020); Guangxi Degree and Postgraduate Education Reform and Development Research Projects, China (JGY2019050); Guangxi Educational Science Planning Key Project (2021B167); Guangxi Higher Education Undergraduate Teaching Reform Project (2020JGA146); Guangxi Medical University Education and Teaching Reform Project (2019XJGZ04); Innovation Project of Guangxi Graduate Education; Guangxi Medical University Training Program for Distinguished Young Scholars (2017); Medical Excellence Award Funded by the Creative Research Development Grant from the First Affiliated Hospital of Guangxi Medical University (2016).

Authors' contributions

Li Gao: Formal analysis, Writing- Original draft preparation; Zi-Qian Liang: Data curation, Investigation; Zhu-Xin Wei: Resources, Software; Jun-Hong Chen: Methodology, Visualization; Jing-Jing Zeng: Supervision, Writing - review & editing; Gang Chen: Project administration; Conceptualization, Funding acquisition, Validation.

Acknowledgements

None.

References

- [1] Ma X, Lakshmipriya T and Gopinath SCB. Recent Advances in Identifying Biomarkers and High-Affinity Aptamers for Gynecologic Cancers Diagnosis and Therapy. *J Anal Methods Chem*, 2019; 2019: 5426974.
- [2] Berger AC, Korkut A, Kanchi RS, Hegde AM, Lenoir W, Liu W, et al. A Comprehensive Pan-Cancer Molecular Study of Gynecologic and Breast Cancers. *Cancer Cell*, 2018; 33: 690-705 e699.
- [3] Ferlay J, Colombet M, Soerjomataram I, Mathers C, Parkin DM, Pineros M, et al. Estimating the global cancer incidence and mortality in 2018: GLOBOCAN sources and methods. *Int J Cancer*, 2019; 144: 1941-1953.

- [4] Skanjeti A, Dhompas A, Paschetta C, Tordo J and Giammarile F. Sentinel Node Mapping in Gynecologic Cancers: A Comprehensive Review. *Semin Nucl Med*, 2019; 49: 521-533.
- [5] Hinchcliff E, Paquette C, Roszik J, Kelting S, Stoler MH, Mok SC, et al. Lymphocyte-specific kinase expression is a prognostic indicator in ovarian cancer and correlates with a prominent B cell transcriptional signature. *Cancer Immunol Immunother*, 2019; 68: 1515-1526.
- [6] Ittiarnornlert P and Ruengkachorn I. Neutrophil-lymphocyte ratio as a predictor of oncologic outcomes in stage IVB, persistent, or recurrent cervical cancer patients treated by chemotherapy. *BMC Cancer*, 2019; 19: 51.
- [7] Workel HH, Komdeur FL, Wouters MC, Plat A, Klip HG, Eggink FA, et al. CD103 defines intraepithelial CD8+ PD1+ tumour-infiltrating lymphocytes of prognostic significance in endometrial adenocarcinoma. *Eur J Cancer*, 2016; 60: 1-11.
- [8] Yoshihara K, Shahmoradgoli M, Martinez E, Vegesna R, Kim H, Torres-Garcia W, et al. Inferring tumour purity and stromal and immune cell admixture from expression data. *Nat Commun*, 2013; 4: 2612.
- [9] Law CW, Chen Y, Shi W and Smyth GK. voom: Precision weights unlock linear model analysis tools for RNA-seq read counts. *Genome Biol*, 2014; 15: R29.
- [10] Gao L, Xie ZC, Pang JS, Li TT and Chen G. A novel alternative splicing-based prediction model for uteri corpus endometrial carcinoma. *Aging (Albany NY)*, 2019; 11: 263-283.
- [11] Xu Z, Dai J, Wang D, Lu H, Dai H, Ye H, et al. Assessment of tumor mutation burden calculation from gene panel sequencing data. *Onco Targets Ther*, 2019; 12: 3401-3409.
- [12] Thomas A, Routh ED, Pullikuth A, Jin G, Su J, Chou JW, et al. Tumor mutational burden is a determinant of immune-mediated survival in breast cancer. *Oncoimmunology*, 2018; 7: e1490854.
- [13] Gao L, Pang YY, Guo XY, Zeng JJ, Tang ZQ, Xiong DD, et al. Polo like kinase 1 expression in cervical cancer tissues generated from multiple detection methods. *PeerJ*, 2020; 8: e10458.
- [14] Wang J, Li Z, Gao A, Wen Q and Sun Y. The prognostic landscape of tumor-infiltrating immune cells in cervical cancer. *Biomed Pharmacother*, 2019; 120: 109444.
- [15] Shen S, Wang G, Zhang R, Zhao Y, Yu H, Wei Y, et al. Development and validation of an immune gene-set based Prognostic signature in ovarian cancer. *EBioMedicine*, 2019; 40: 318-326.
- [16] Brule S, Friand V, Sutton A, Baleux F, Gattegno L and Charneau N. Glycosaminoglycans and syndecan-4 are involved in SDF-1/CXCL12-mediated invasion of human epitheloid carcinoma HeLa cells. *Biochim Biophys Acta*, 2009; 1790: 1643-1650.

[17] Yarla NS, Bishayee A, Sethi G, Reddanna P, Kalle AM, Dhananjaya BL, et al. Targeting arachidonic acid pathway by natural products for cancer prevention and therapy. *Semin Cancer Biol*, 2016; 40-41: 48-81.

[18] Kouba S, Ouldamer L, Garcia C, Fontaine D, Chantome A, Vandier C, et al. Lipid metabolism and Calcium signaling in epithelial ovarian cancer. *Cell Calcium*, 2019; 81: 38-50.

[19] Sahoo SS, Quah MY, Nielsen S, Atkins J, Au GG, Cairns MJ, et al. Inhibition of extracellular matrix mediated TGF-beta signalling suppresses endometrial cancer metastasis. *Oncotarget*, 2017; 8: 71400-71417.

[20] Varjosalo M and Taipale J. Hedgehog: functions and mechanisms. *Genes Dev*, 2008; 22: 2454-2472.

[21] Westendorp F, Karpus ON, Koelink PJ, Vermeulen JLM, Meisner S, Koster J, et al. Epithelium-derived Indian Hedgehog restricts stromal expression of ErbB family members that drive colonic tumor cell proliferation. *Oncogene*, 2021;

[22] Kasiri S, Chen B, Wilson AN, Reczek A, Mazambani S, Gadhvi J, et al. Stromal Hedgehog pathway activation by IHH suppresses lung adenocarcinoma growth and metastasis by limiting reactive oxygen species. *Oncogene*, 2020; 39: 3258-3275.

Figures

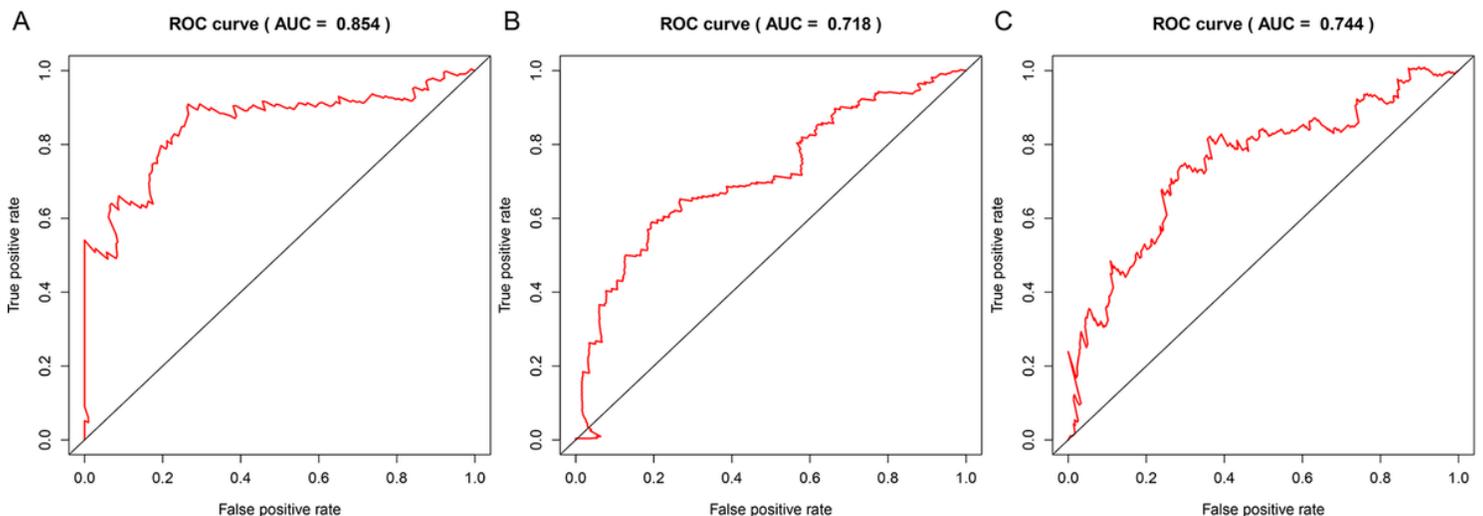


Figure 1

Assessment of the prognostic prediction ability of the immune and stromal-related risk signatures. A. tROC curves at five years for immune and stromal score-related risk signature of CESC; B. tROC curves at five years for immune and stromal score-related risk signature of OVCA; C. tROC curves at five years for immune and stromal score-related risk signature of UCEC.

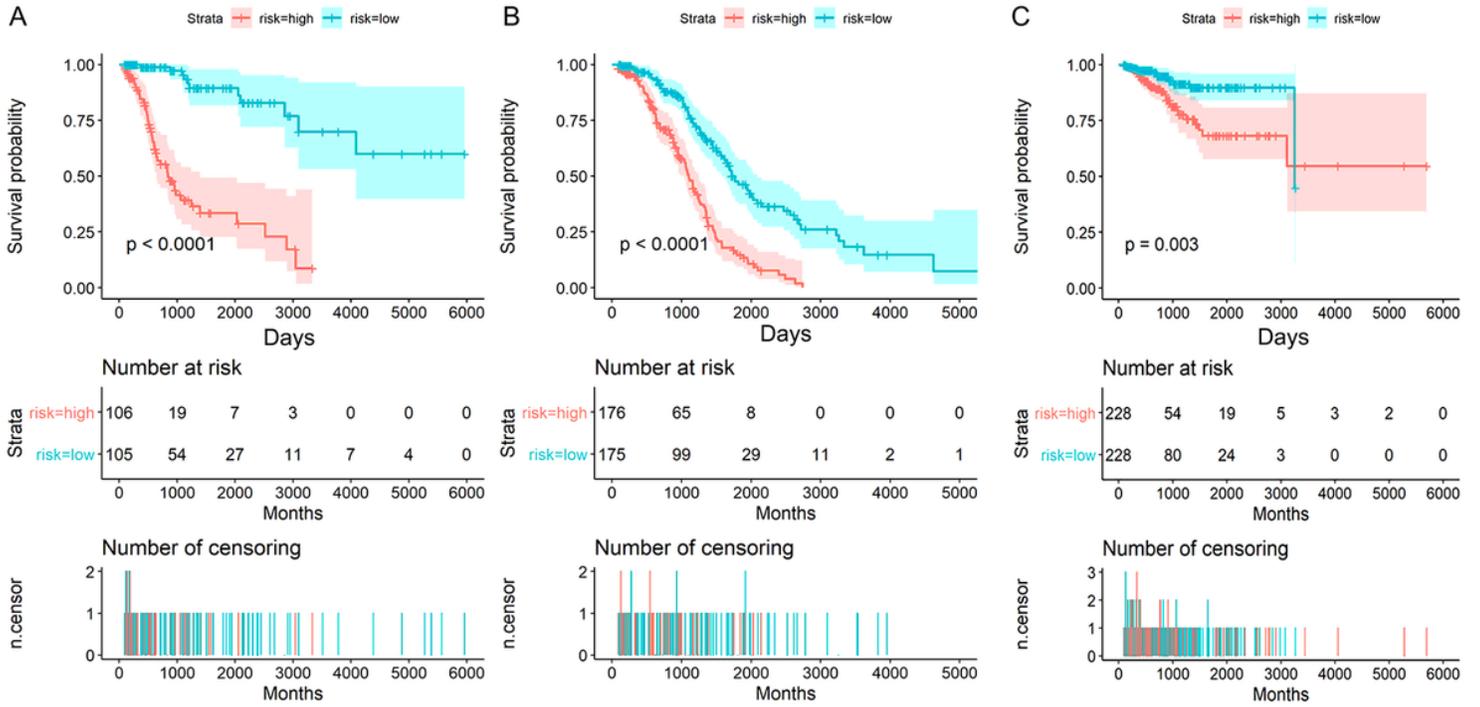


Figure 2

Comparison of survival time between high- and low-risk gynaecologic patients. A. Kaplan-Meier survival curves for immune and stromal score-related risk signature of CESC; B. Kaplan-Meier survival curves for immune and stromal score-related risk signature of OVCA; C. Kaplan-Meier survival curves for immune and stromal score-related risk signature of UCEC.

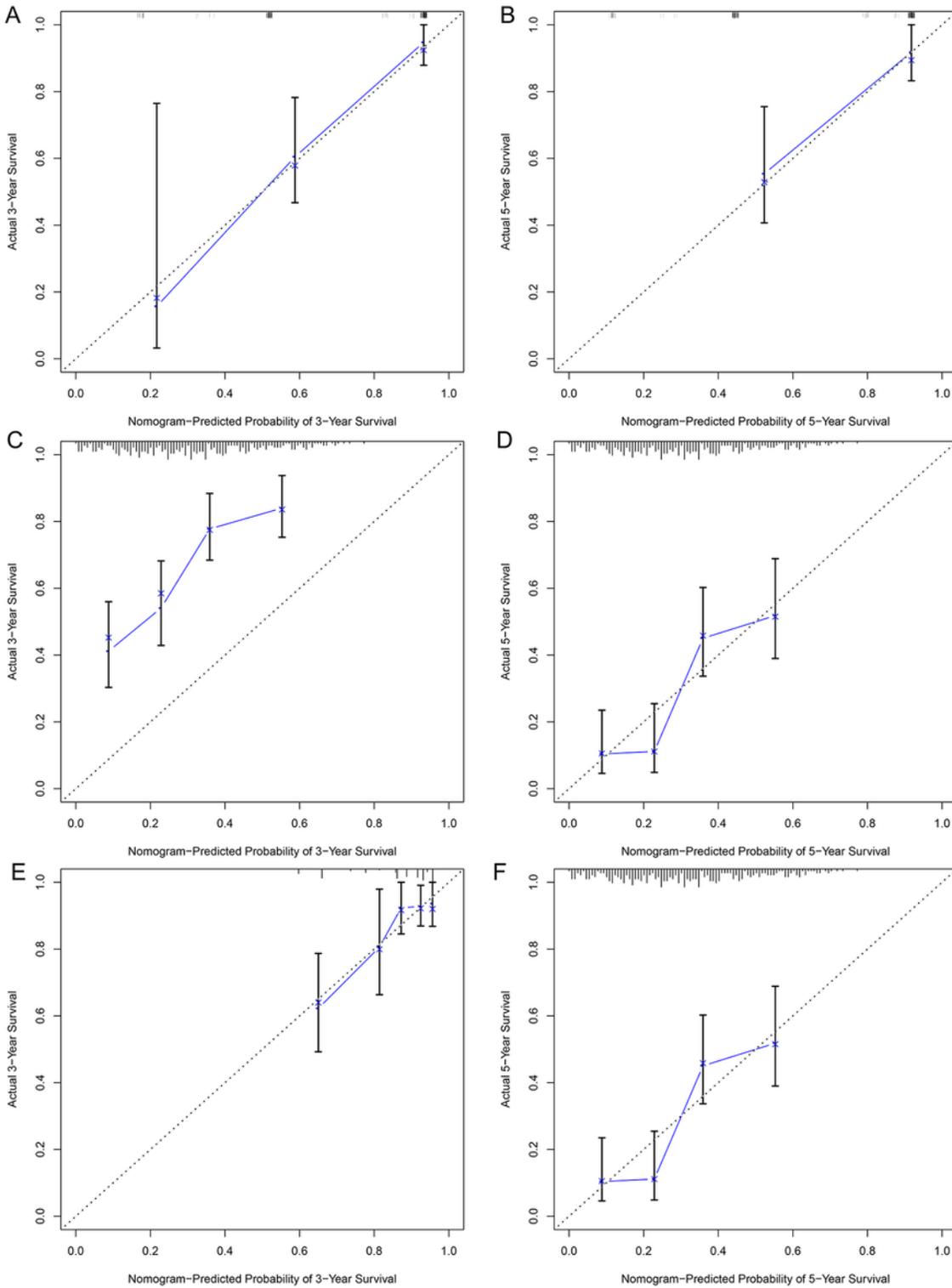


Figure 3

Evaluation of nomograms for gynaecologic cancers. A. Calibration curve of at three years for CESC; B. Calibration curve of at five years for CESC; C. Calibration curve of at three years for OVCA; D. Calibration curve of at five years for OVCA; E. Calibration curve of at three years for UCEC; F. Calibration curve of at five years for UCEC.

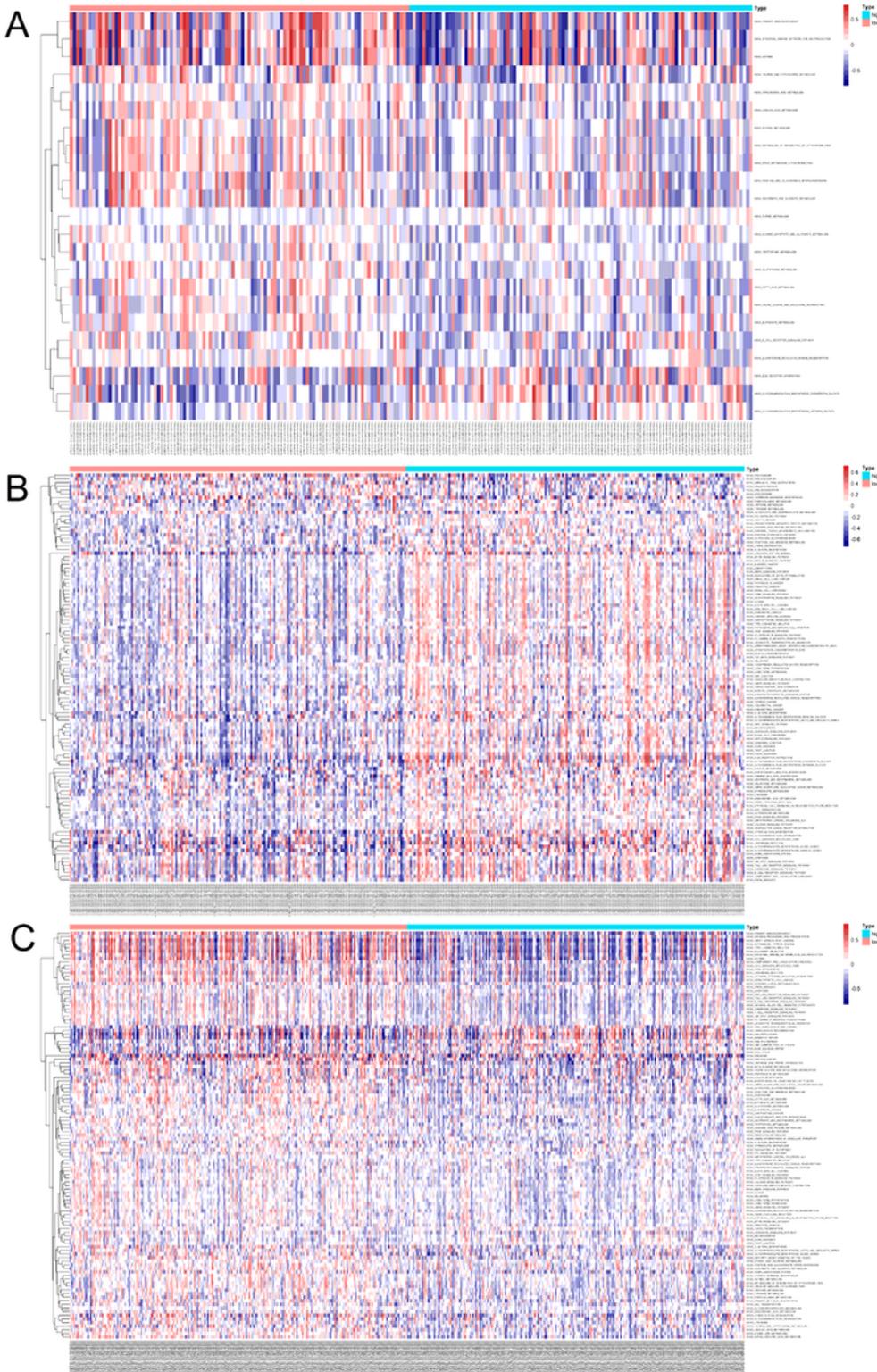


Figure 4

Differentially enriched KEGG pathways between high- or low-risk groups of gynaecologic cancer patients. A. Volcano plot for differentially enriched KEGG pathways in CESC; B. Heatmap for differentially enriched KEGG pathways in CESC; C. Volcano plot for differentially enriched KEGG pathways in OVCA; D. Heatmap for differentially enriched KEGG pathways in OVCA; E. Volcano plot for differentially enriched KEGG pathways in UCEC; F. Heatmap for differentially enriched KEGG pathways in UCEC.

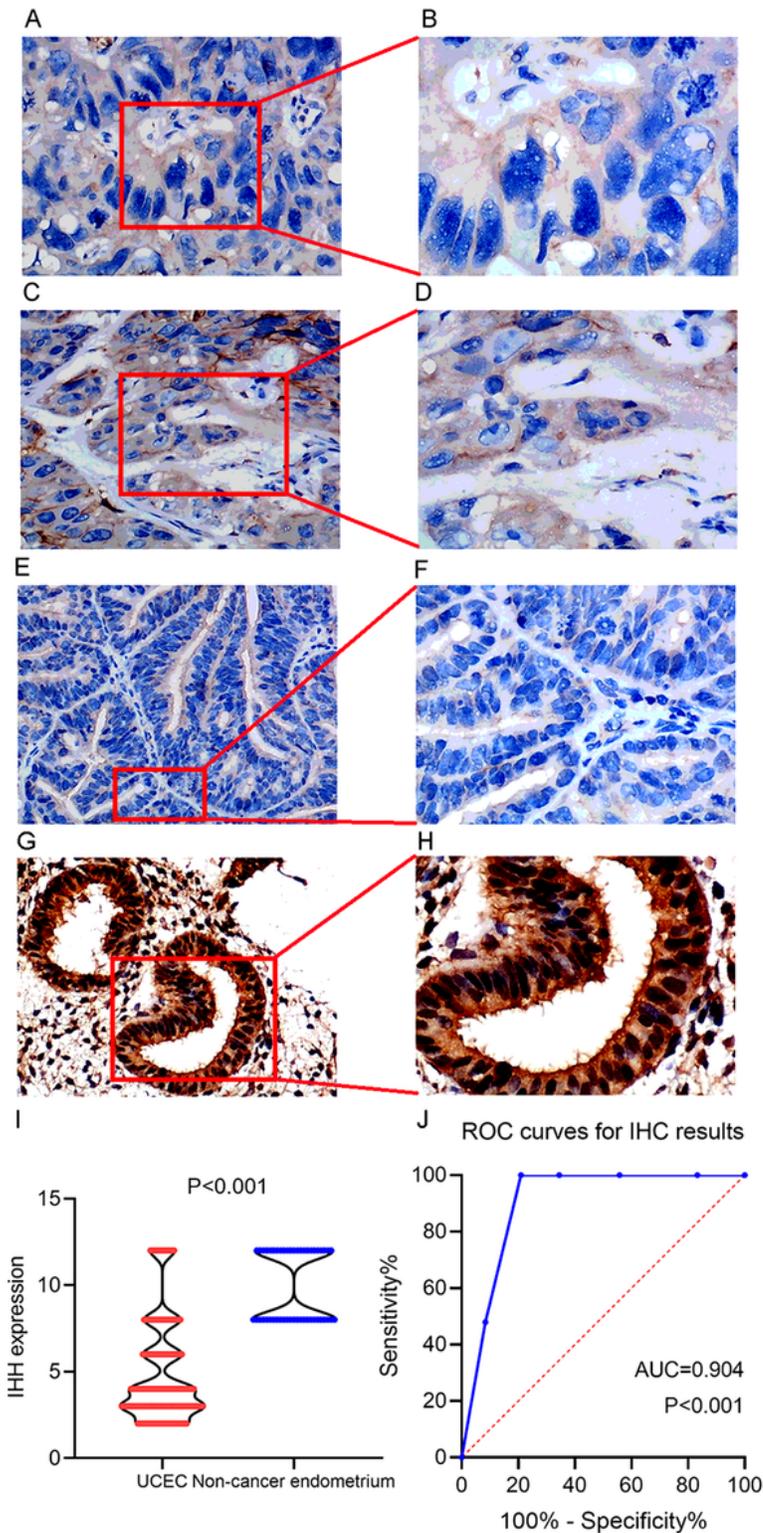


Figure 5

Protein expression of IHH in UCEC and non-cancer endometrium tissues. A. Negative staining of IHH in UCEC tissues (200X) (immune score: 3); B. Negative staining of IHH in UCEC tissues (400X) (immune score: 3); C. Weak staining of IHH in UCEC tissues (200X) (immune score: 4); D. Weak staining of IHH in UCEC tissues (400X) (immune score: 4); E. Negative staining of IHH in UCEC tissues (200X) (immune score: 2); F. Negative staining of IHH in UCEC tissues (400X) (immune score: 2); G. Strong staining of IHH

in Non-cancer endometrium tissues (200X) (immune score: 12); H. Strong staining of IHH in Non-cancer endometrium tissues (400X) (immune score: 12). I. Violin plot of IHH protein expression in UCEC and non-cancer endometrium tissues. J. ROC curves of IHH protein expression in UCEC and non-cancer endometrium tissues.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [Additionalfile1.png](#)
- [Additionalfile2.png](#)
- [Additionalfile3.xlsx](#)
- [Additionalfile4.png](#)
- [Additionalfile5.png](#)
- [Additionalfile6.png](#)
- [Additionalfile7.png](#)
- [Additionalfile8.png](#)
- [Additionalfile9.png](#)
- [Additionalfile10.png](#)
- [Additionalfile11.png](#)