Heliyon 10 (2024) e29548

Contents lists available at ScienceDirect

Heliyon

journal homepage: www.cell.com/heliyon

Research article

# POLE -related gene signature predicts prognosis, immune feature, and drug therapy in human endometrioid carcinoma

Wei Qiu [a], Runjie Zhang [b,c,**], Yingchen Qian [a,*]

[a] Department of Pathology, The Affiliated Jiangning Hospital of Nanjing Medical University, No.169, HuShan Road, Nanjing, 211100, China
[b] Hongqiao International Institute of Medicine, Tongren Hospital, Shanghai Jiao Tong University School of Medicine, No.1111, XianXia Road, Shanghai, 200336, China
[c] Obstetrics and Gynecology Department, Tongren Hospital, Shanghai Jiao Tong University School of Medicine, No.1111, XianXia Road, Shanghai, 200336, China

ARTICLE INFO

*Keywords:*
Endometrial carcinoma
A *POLE*-Related risk score signature (PRS)
Prognosis
Immunity
Drug sensitivity

ABSTRACT

The *POLE* subtype of Endometrial carcinoma (EC) is linked to a favourable prognosis in the molecular classification. We proposed to ascertain the potential connection between the POLE subtype and improved prognosis. In order to forecast the prognosis, least absolute shrinkage and selection operator (LASSO) Cox regression analysis and weighted gene co-expression network analysis (WGCNA) were employed, and a *POLE*-related risk signature (PRS) model was developed and validated. Single-sample gene set enrichment analysis (ssGSEA) with the "GSVA" package was employed to analyse immunity characteristics. Drug susceptibility studies were conducted to compare the half-maximal inhibitory concentration ($IC_{50}$) of medicines between high- and low-risk groups. The PRS model was generated employing the LASSO Cox regression coefficients of the *ELF1*, *MMADHC*, *and AL021707.6* genes. Our study demonstrated that the risk score was linked to tumour stage, grade, and survival. Furthermore, the low-risk group possessed elevated levels of gene expression connected with immunological checkpoints and HLA. Our outcomes emerged that the PRS model might have value in identifying patients with a good prognosis and in facilitating personalised treatment in the clinic.

## 1. Introduction

Endometrial carcinoma is a cancerous tumour that develops in the female reproductive system, and based on a report on cancer statistics, in 2023, there has been a steady rise in the occurrence and death rates of endometrial cancer [1]. EC is mostly treated with surgery that is often performed in conjunction with adjuvant therapy, including radiotherapy, chemotherapy, and endocrine therapy [2]. There is still controversy regarding the ideal choice for adjuvant therapy, which depends on the characteristics of cancer patients,

*Abbreviations:* POLE, polymerase epsilon; LASSO, least absolute shrinkage and selection operator; EC, Endometrial carcinoma; WGCNA, Weighted gene co-expression network analysis; MSI, microsatellite stability unstable; CNV, copy number variation; TME, tumor microenvironment; GDC, Genomic Data Commons; GS, gene Significance; OS, overall survival; MM, Module Membership; TMB, tumor mutation burden; GSEA, gene set enrichment analysis; PRS, *POLE*-related risk signature; $IC_{50}$, the half-maximal inhibitory concentration; TCGA, The Cancer Genome Atlas; HRG, high-risk group; LRG, low-risk group; ICP, immune checkpoint.

* Corresponding author.
** Corresponding author. Hongqiao International Institute of Medicine, China.
*E-mail addresses:* zrj@shsmu.edu.cn (R. Zhang), qianyingchen@jnyy10.wecom.work (Y. Qian).

https://doi.org/10.1016/j.heliyon.2024.e29548
Received 16 July 2023; Received in revised form 7 April 2024; Accepted 9 April 2024
Available online 16 April 2024

such as the status of metastasis to the lymph nodes [3]. Moreover, the increase in mortality observed in recent decades has also suggested the limitations of pathological assessment, tumour grade and stage, and lymphovascular space invasion in selecting clinical treatments [4]. Thus, it is vital to develop and create clinical treatment guidelines for the selection of postsurgical adjuvant therapy and to identify new therapeutic targets to improve prognosis.

The Cancer Genome Atlas (TCGA) has introduced four new prognostic classifications: microsatellite stability unstable (MSI) hypermutated, DNA polymerase epsilon ultramutated (POLEmut), copy-number low, and high [5]. Clinical prognosis has been verified to be significantly correlated with these subtypes [6]. For example, it is predicted that the clinical prognosis of EC could be improved and that the incidence of both overtreatment and undertreatment could be reduced by reclassifying patients with EC according to the TCGA classification system [5,7]. As one TCGA subtype, the POLEmut subtype has better outcomes than the other three molecular subtypes [8]. However, the mechanism underlying the good prognosis in patients with the POLEmut subtype remains unclear.

In tumorigenesis-related research, mounting data indicates that the tumour microenvironment (TME) significantly contributes to the growth and medication resistance of tumour cells [9], especially immune cells, as vital constituents of the TME, also contribute to an essential function in tumour survival and development [10]. Previous investigations revealed that the POLE gene has the activity of DNA polymerase and the activity of possesses 3′-5′ exonuclease [11], which is also linked to the presence of immune cells around EC cells [12]. Nevertheless, few investigations have presented the characteristics of EC immune cells, particularly in the POLE subtype.

Weighted gene co-expression network analysis (WGCNA) and least absolute shrinkage and selection operator (LASSO) Cox regression analysis are crucial bioinformatics methods at present [13–15]. These analyses may help researchers identify and study modules and reveal the key genes involved in different diseases [16–18]. Therefore, we established a *POLE*-related risk score (PRS) model employing WGCNA and LASSO Cox regression analysis and assessed its capacity to anticipate the EC prognosis and to identify differences in response to immuno and chemotherapy between risk groups.

## 2. Materials and methods

### 2.1. Data collection

The transcription database was gathered from TCGA UCEC cohort (https://www.cancer.gov/types/uterine). The data on copy number variation (CNV) and somatic mutation were gathered from the NCI Genomic Data Commons (GDC) site. Prior to analysis, the raw data was normalised by employing fragments per kilobase of transcripts per million mapped reads (FPKM) values. Moreover, we excluded specimens with missing information on corresponding patient survival and drug treatment. Then, Perl scripts were used to determine the difference in tumour mutation burden (TMB) according to the total human exon length per sample. A total of 525 cancer

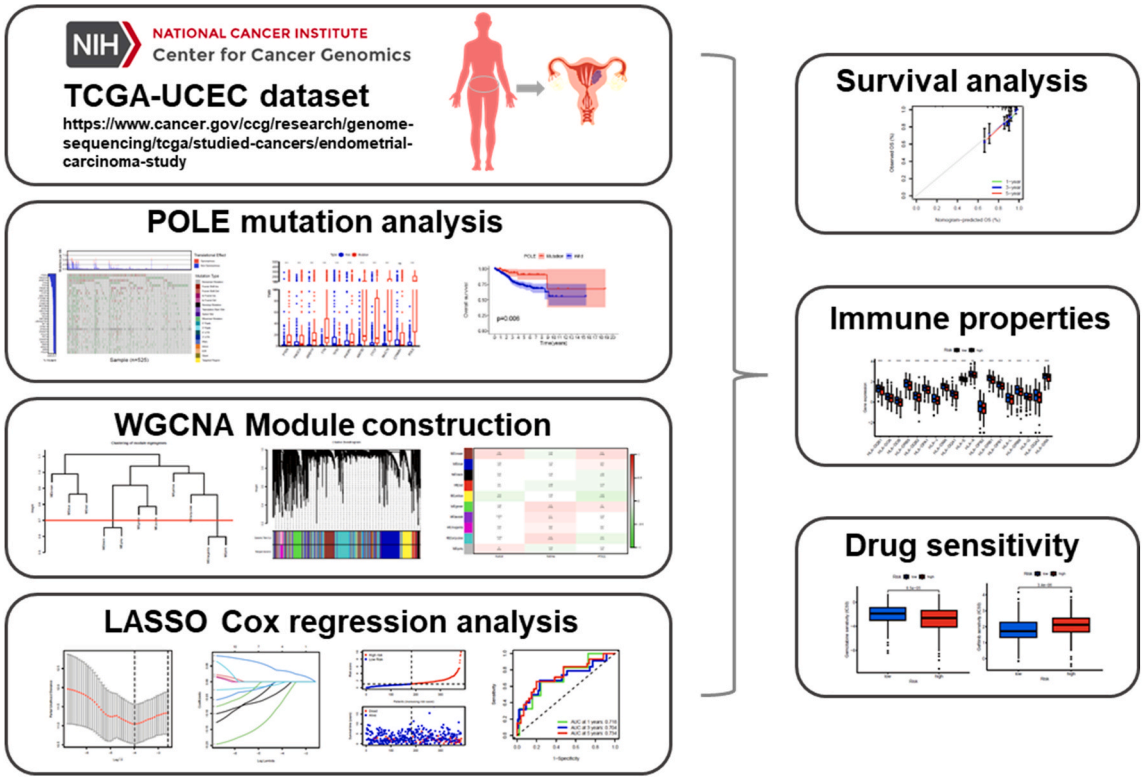



**Fig. 1.** The investigation flowchart.

specimens were visualised as waterfall plots employing the Maftools software [19]. The overall flowchart of this study and the clinical information are manifested in Fig. 1 and Table 1.

### 2.2. WGCNA

After data cleaning, the RNA-seq data of TCGA UCEC was analysed using WGCNA (version 1.61) [20]. Soft thresholding ($\beta = 6$) was selected by scale-free fit index (y = 0.9) to construct the coexpression network. After deleting outlier samples, consensus data were employed to analyse the connection between clinical characteristics and gene expression. Subsequently, the acquisition of new modules was achieved through the merging of modules that had a dissimilarity value of less than 0.7 and choosing the module trees with a tangent of 0.25. Afterwards, a correlation study was executed to ascertain the association between module participation and POLE. Finally, the gene significance (GS) and module membership (MM) of the turquoise module were computed to estimate the connections between the genes and clinical features.

### 2.3. Construction of a prognostic PRS signature

LASSO regression was performed to identify genes related to the POLEmut subtype and correlated with overall survival (OS) using the "glmnet" package. The partial likelihood deviance nomography was used to select the penalisation parameter (λ) as the optimal and minimal criteria. The median risk score was employed to divide the entire EC cohort into high-risk group (HRG) and low-risk group (LRG) (risk score = Exp$i$ × $bi$; Exp: The model gene expression level; $b$: model gene coefficient) [21]. The heatmap analyses were conducted utilising the "pheatmap" package in order to illustrate the distinction between the genes linked to POLE and the pattern of distribution that exists between clinicopathological characteristics and risk groups. The prediction capability of the risk score model was ascertained by employing the area under the time-dependent receiver-operating characteristic curve (AUROC). Additionally, the risk curve was constructed to ascertain the variability in survival status between various risk groups. Completion of the criteria is detailed in Supplementary Tables S1–S3.

### 2.4. Construction of the prognostic model

Univariate and multivariate analyses were employed to ascertain whether the risk score could independently impact the survival of EC patients in the TCGA UCEC cohort. The "regplot" package was employed to combine stage, grade, and risk signature in order to forecast survival. The predicted precision of the nomogram was assessed by means of a calibration curve.

### 2.5. Immune cell infiltration analysis

On the basis of the findings of the single-sample gene set enrichment analysis (ssGSEA), the "ESTIMATE" programme was deployed to ascertain the stromal, immune, and estimate scores. The activity scores of immune-linked pathways were computed employing ssGSEA with the "GSVA" package. The analysis of immune infiltration and functioning was conducted across various risk groups utilising the stat_compare_means method. The immune checkpoint (ICP)- and human leukocyte antigen (HLA)-correlated gene expression was further compared across the different risk groups.

### 2.6. Drug sensitivity between different risk groups

The Genomics of Drug Sensitivity in Cancer database was deployed to anticipate the drug efficacy. All drugs (A.443654, A.770041, ABT.263, ABT.888, AG.014699, AICAR, AKT.inhibitor.VIII, AMG.706, AP.24534, AS601245, ATRA, AUY922, Axitinib, AZ628, AZD.0530, AZD.2281, AZD6244, AZD6482, AZD7762, AZD8055, BAY.61.3606, Bexarotene, BI.2536, BIBW2992, Bicalutamide, BI. D1870, BIRB.0796, Bleomycin, BMS.509744, BMS.536924, BMS.708163, BMS.754807, Bortezomib, Bosutinib, Bryostatin.1, BX.795, Camptothecin, CCT007093, CCT018159, CEP.701, CGP.082996, CGP.60474, CHIR.99021, CI.1040, Cisplatin, CMK, Cyclopamine, Cytarabine, Dasatinib, DMOG, Docetaxel, Doxorubicin, EHT.1864, Elesclomol, Embelin, Epothilone. B, Erlotinib, Etoposide, FH535, FTI.277, GDC.0449, GDC0941, Gefitinib, Gemcitabine, GNF.2, GSK269962A, GSK.650394, GW.441756, GW843682X, Imatinib,

**Table 1**
Clinical characteristics.

| | Alive | Dead |
|---|---|---|
| Grade | | |
| G1 (n = 102) | 102 | 0 |
| G2 (n = 130) | 120 | 10 |
| G3 (n = 364) | 326 | 38 |
| **Stage** | | |
| I (n = 369) | 354 | 15 |
| II (n = 58) | 54 | 4 |
| III (n = 135) | 116 | 19 |
| IV (n = 34) | 24 | 10 |

IPA.3, JNJ.26854165, JNK.9L, JNK.Inhibitor.VIII, JW.7.52.1, KIN001.135, KU.55933, Lapatinib, Lenalidomide, LFM.A13, Metformin, Methotrexate, MG.132, Midostaurin, Mitomycin. C, MK.2206, MS.275, Nilotinib, NSC.87877, NU.7441, Nutlin.3a, NVP.BEZ235, NVP.TAE684, Obatoclax.Mesylate, OSI.906, PAC.1, Paclitaxel, Parthenolide, Pazopanib, PD.0325901, PD.0332991, PD.173074, PF.02341066, PF.4708671, PF.562271, PHA.665752, PLX4720, Pyrimethamine, QS11, Rapamycin, RDEA119, RO.3306, Roscovitine, Salubrinal, SB.216763, SB590885, Shikonin, SL.0101.1, Sorafenib, S.Trityl.L.cysteine, Sunitinib, Temsirolimus, Thapsigargin, Tipifarnib, TW.37, Vinblastine, Vinorelbine, Vorinostat, VX.680, VX.702, WH.4.023, WO2009093972, WZ.1.84, X17.AAG, X681640, XMD8.85, Z.LLNle.CHO, ZM.447439)were analysed, and the half-maximal inhibitory concentration ($IC_{50}$) was compared using the "pRRophetic" and "ggplot2" packages to anticipate further changes in drug sensitivity between various risk groups of EC patients.

### 2.7. Statistical analysis

The statistical analyses and graphs in this investigation were generated employing R software R.4.1.1 (https://www.r-project.org/). In contrast, the Wilcoxon rank-sum test was employed to compare variables that did not follow a normal distribution, whereas the *t*-test was applied to the various risk groups. A significance level of $P < 0.05$ was deemed significant.

## 3. Results

### 3.1. Tumor mutation analysis in EC samples

We investigated the variation between mutation and wild-type groups of TMB in EC tissues using the TCGA UCEC database (Fig. 2a

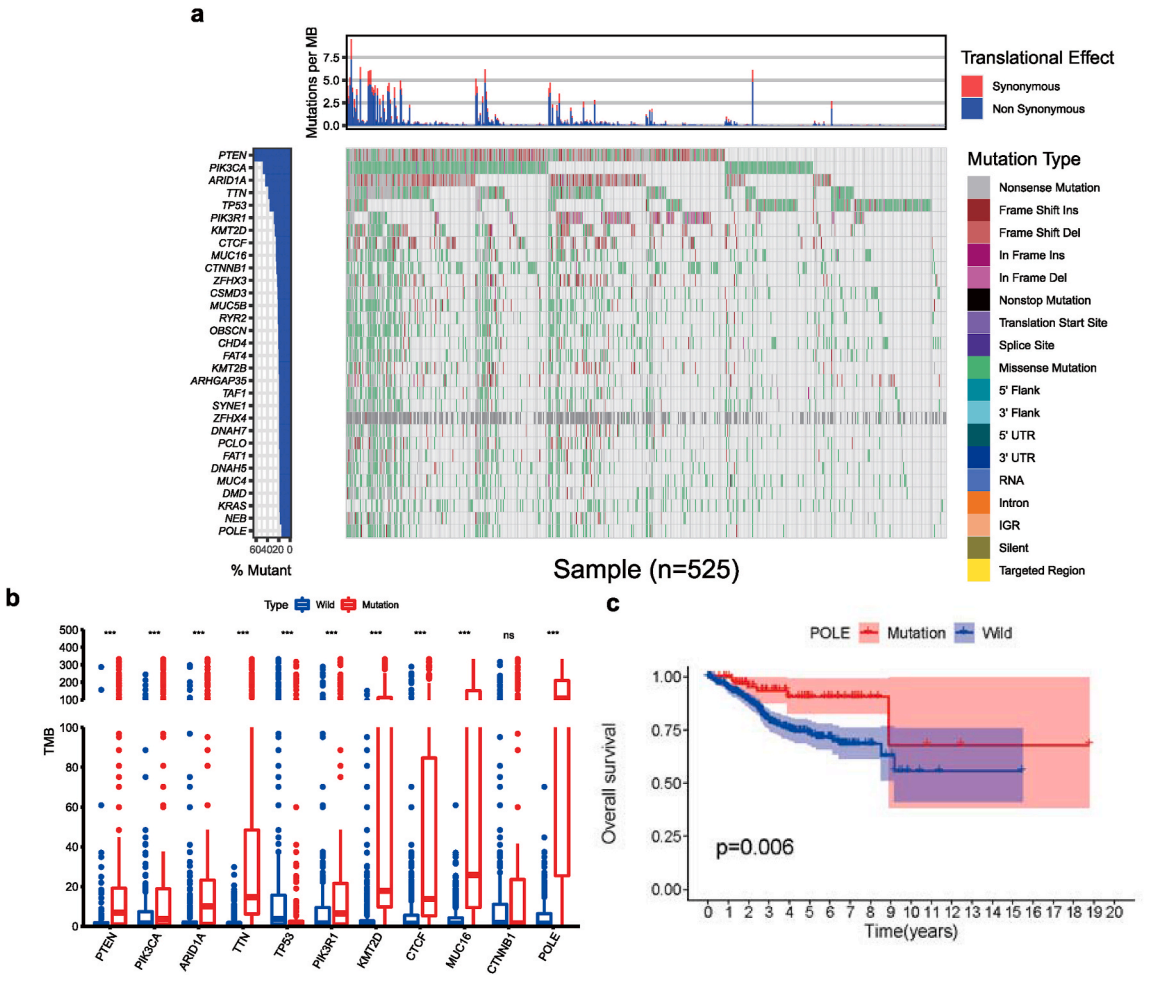



**Fig. 2.** Tumor mutation analysis in EC specimens
In the TCGA UCEC database, **a** Waterfall plot was displayed to analyse Tumor mutation. **b** The histogram plot was used to show the Tumor mutation burden (TMB). **c** K-M survival curve was used to analyse the Overall survival (OS) between POLEmut group and POLEwild group. *$P < 0.05$, **$P < 0.01$.

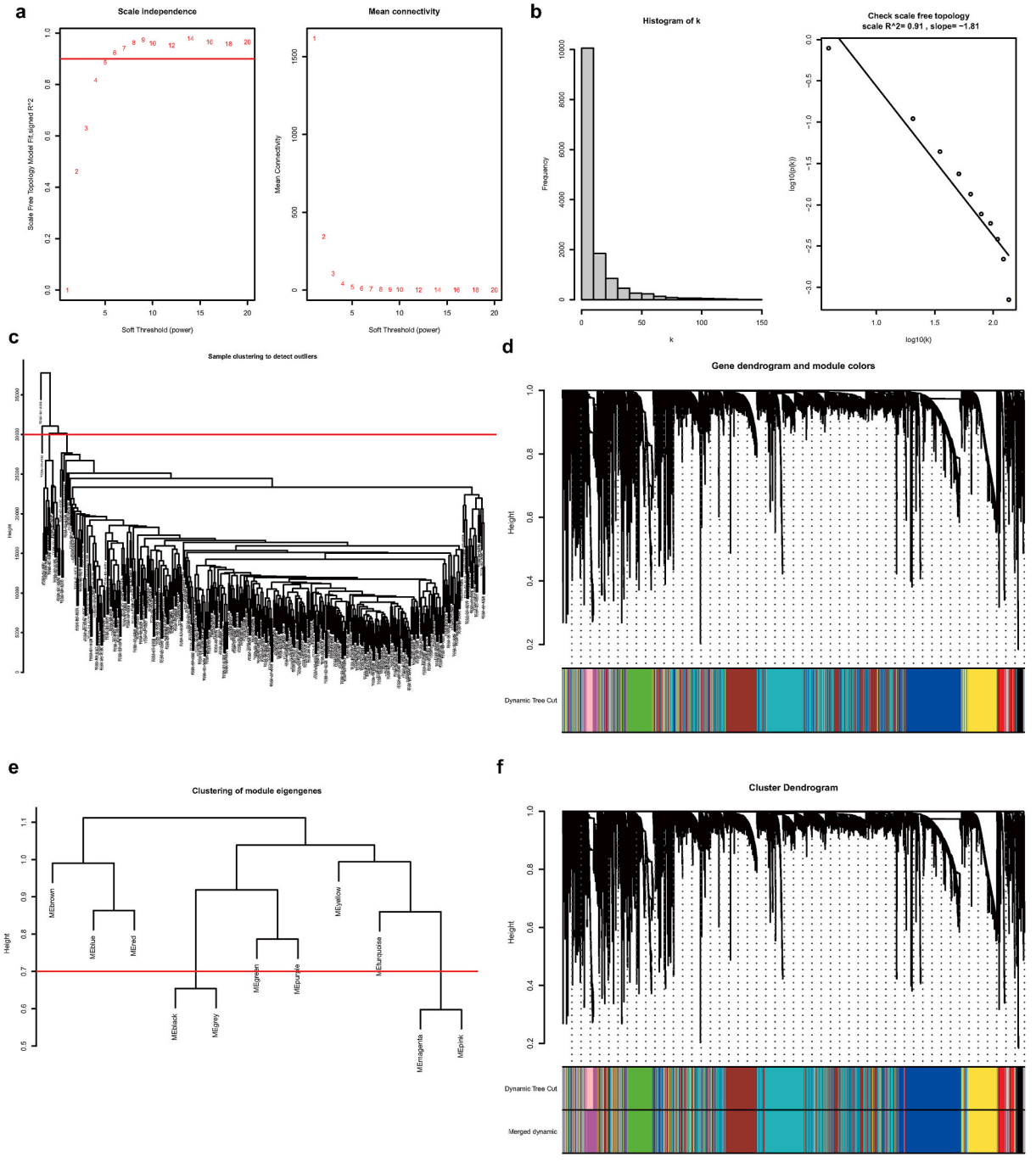



**Fig. 3.** Construction of WGCNA network module

**a** Soft thresholds selection by scale independence and mean connectivity in WGCNA network analysis. **b** The histogram and the correlation coefficient were performed to show k and the correlation between k and p (k). **c-d** Dynamic Tree Cut algorithm was performed, and cluster dendrogram of the genes was plotted in the EC samples. **e-f** Clustering of module eigengenes and cluster dendrogram was performed to show the merge modules in EC samples. **g** The heatmap showed the correlation between merge modules and the survival state, *POLE*. **h** The correlation plot was performed to analyse the connection between GS and MM in the turquoise module. $*P < 0.05$, $**P < 0.01$.

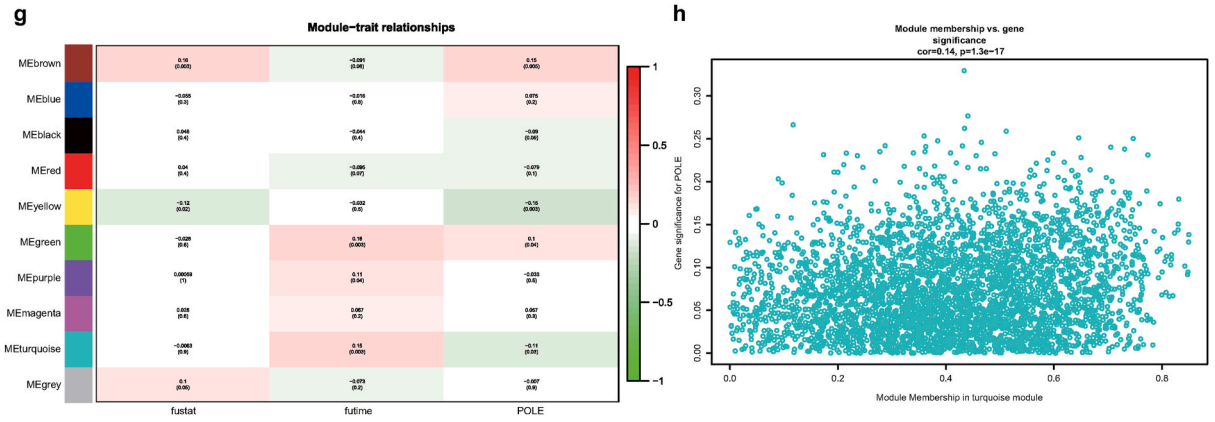



**Fig. 3.** (*continued*).

and b). K-M analysis manifested that the POLEmut group possessed longer OS than the POLEwild group ($P = 0.006$; Fig. 2c).

### 3.2. WGCNA module construction

All genes with information in the TCGA UCEC dataset were applied for the WGCNA analysis to obtain the *POLE*-associated genes. The soft thresholds ($\beta = 6$) used for strengthening the adjacency matrix were determined using scale-free topology with $R^2 > 0.9$ (Fig. 3a and b). Subsequently, topological overlap matrix-based dissimilarity analysis and the dynamic tree-cut algorithm were used to construct a dendrogram of the genes in the EC samples (Fig. 3c and d). Next, hierarchical clustering was conducted, revealing the presence of identical gene expression patterns within the same branch (Fig. 3e). After identifying gene modules and combining similar gene modules, we obtained 10 coexpression modules (Fig. 3f). The correlation coefficient results demonstrated that the turquoise and green modules related to the *POLE* gene exhibited a correlation with survival time. (Fig. 3g). Moreover, the connection between GS and MM in the turquoise module was very significant (Fig. 3h).

### 3.3. Creation and verification of the PRS signature

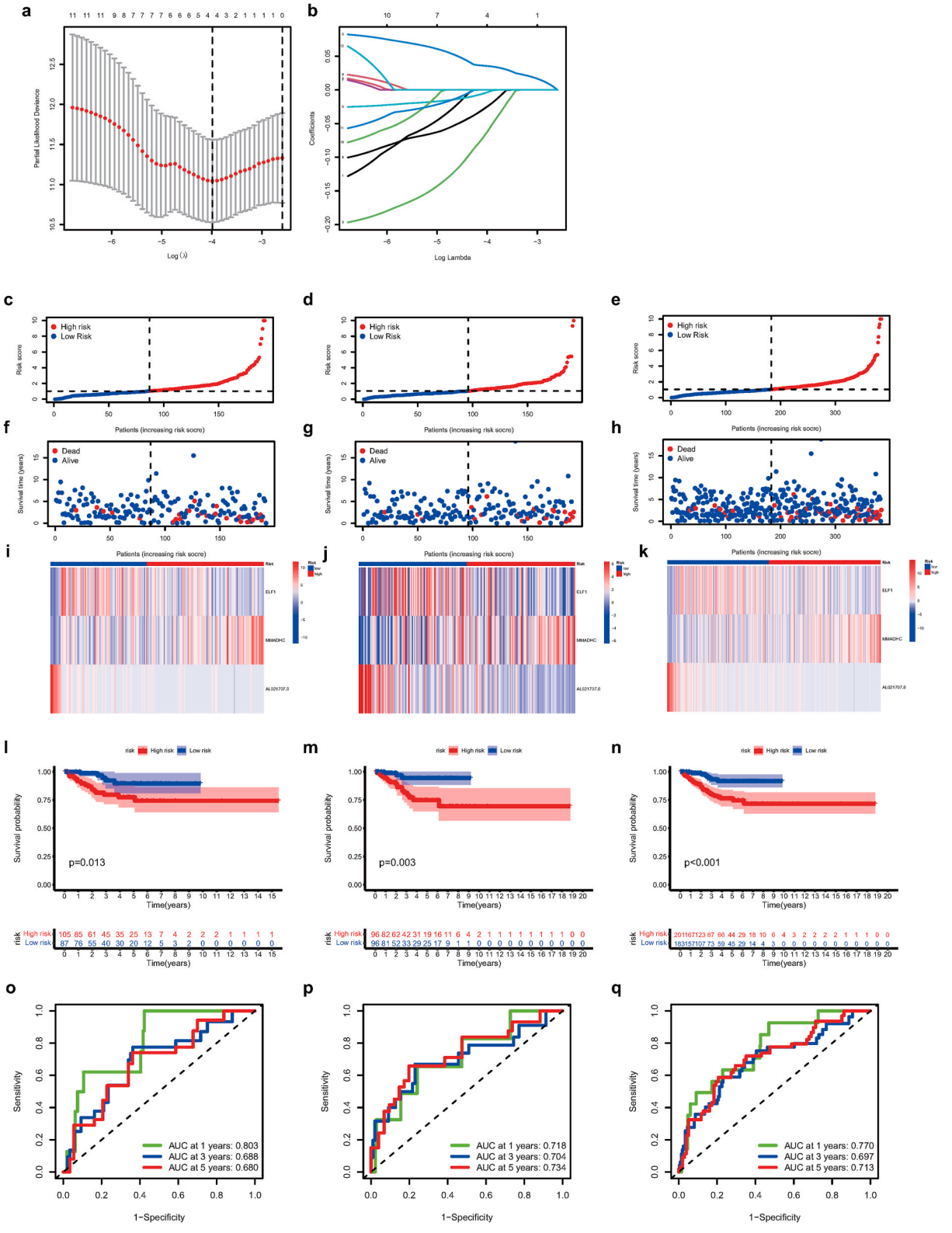
We employed LASSO Cox regression analyses and identified three candidate genes related to *POLE* to ensure the stability of the model (Fig. 4a and b). Risk curve analysis was also conducted in the test, training, and total sets. The plots revealed that the LRG had a lower risk coefficient and mortality than those in the HRG (Fig. 4c–h). The heatmap revealed the change in the *POLE*-associated gene expression between the HRG and LRG (Fig. 4i–k). The K-M analysis manifested that patients with EC in the HRG had a less overall survival (OS) contrasted with those in the LRG (Fig. 4l-n). In the testing group (Fig. 4o), the AUC was 0.803 for the 1-year OS, 0.688 for the 3-year OS, and 0.680 for the 5-year OS. The AUC values for the 1-, 3-, and 5-year OS in the training set were 0.718, 0.704, and 0.734, respectively. As shown in Fig. 4p-q, the AUC values for the entire set were 0.770, 0.697, and 0.713, respectively.

### 3.4. Creation and determination of the predictive nomogram of survive

We executed univariate and multivariate Cox regression analyses to see if the PRS is an independent prognostic factor. The outcomes manifested that the risk score had independent implications for forecasting the prognosis in the TCGA-UCEC cohort (Fig. 5a and b). Moreover, we established the nomenclature and the values of the nomogram plot of prediction and actually revealed that our nomogram is better reliable and accurate in anticipating both short- and long-term survival (Fig. 5c). Furthermore, We developed a nomogram employing the clinical features of PRS, clinical stage, and grade to forecast the short- and long-term survival of EC patients (Fig. 5d).

### 3.5. The immune properties of each PRS signature

Next, we assessed the differences in tumour immune microenvironment features and the immune-associated pathways activation between risk groups. We employed the ESTIMATE algorithm and manifested that the immune, estimate, and stromal scores in the LRG were greater than in the HRG (Fig. 6a). These outcomes manifest that distinct variations in TME cell infiltration patterns were observed between the two risk groups. Thus, an immune infiltration study was performed, and the outcomes manifested variations in the proportions of 15 various kinds of immune cells in the cancer immunological microenvironment between the two risk groups (Fig. 6b). Subsequently, the CIBERSORT analysis plot indicated that the LRG has a greater number of infiltrating immune cells than in the HRG (Fig. 6c). Additionally, immune-associated function analysis demonstrated that most function-related scores, especially HLA and Checkpoint score, were significantly greater in the LRG than in the HRG (Fig. 6d). Then, we conducted a comparative analysis to ascertain the ICP- and HLA-related genes expression levels in the two risk groups. A boxplot exhibited that the checkpoint-associated




*(caption on next page)*

**Fig. 4.** Creation and validation of the PRS signature
**a** The log (lambda) sequence was generated in the LASSO model. **b** The LASSO Cox model was employed to choose three candidates' *POLE*-related genes. **c-e** The risk score distribution was shown in the testing, training, and total sets between the two risk groups. **f-h** The survival time (years) was analysed in the testing, training, and total sets between the two risk groups. **i-k** The heatmap was shown the expression of three candidates *POLE*-related genes in the testing, training, and total sets between the two risk groups. **l-m** K-M survival curve was analysed in the testing, training, and total sets. **o-q** ROC analysis was conducted in the testing, training, and total sets encompassing 1-, 3-, and 5- years of EC patients.

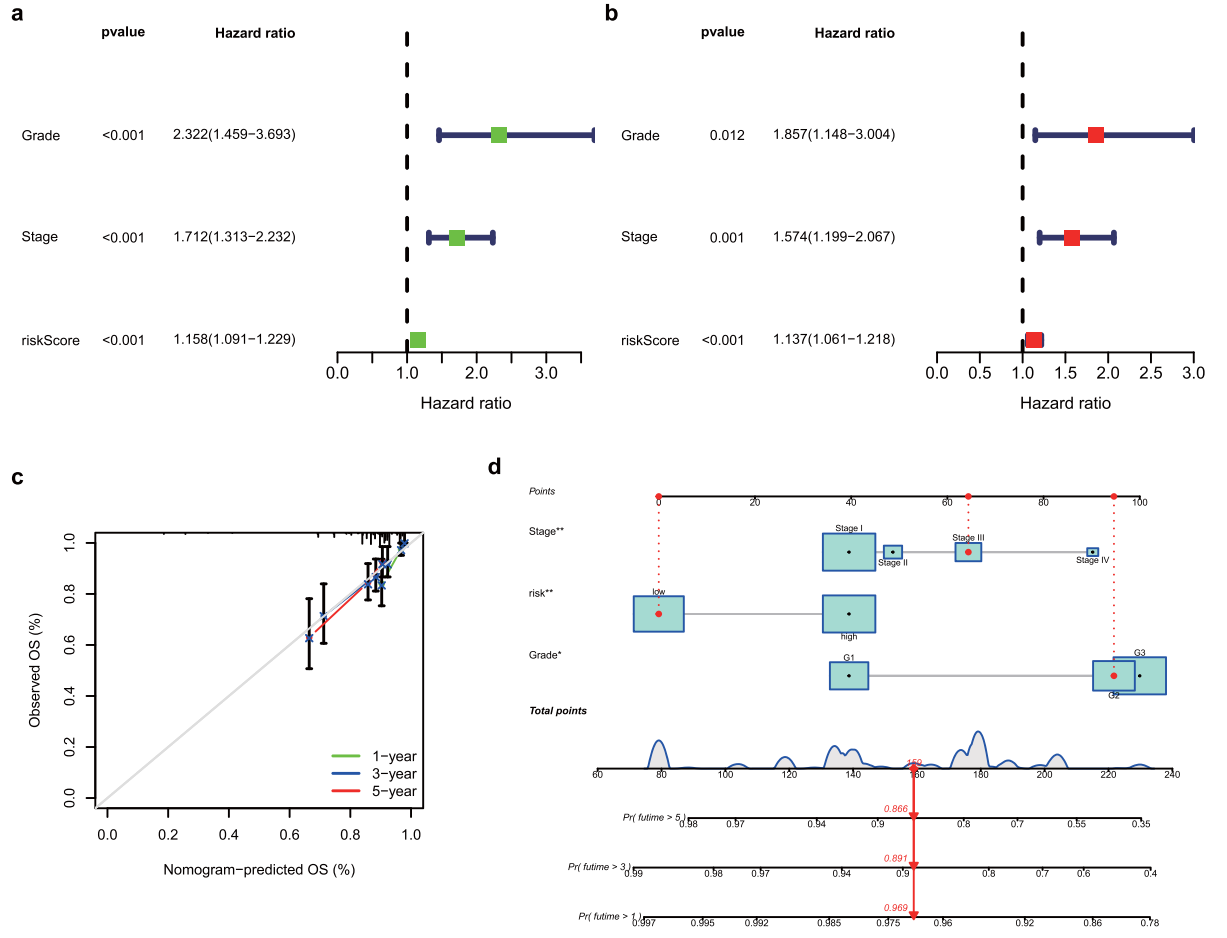



**Fig. 5.** Creation and estimation of the predictive nomogram of survive **a-b** univariate and multivariate Cox regression analyses for the clinical characteristics and risk score. **c** Calibration curves of the nomogram in prediction of the 1-, 3-, and 5-year survival probability. **d** The Nomogram with the clinical characteristics was created to anticipate 1-, 3-, and 5- years of survival.

genes (*PDCD1*, *TNFRSF9*, *VTCN1*, *VSIR LAIR1*, *NRP1*, and *TNFRSF14*) and HLA had higher expression levels in the LRG than in the HRG (Fig. 6e and f). The above outcomes presented that the LRG may exhibit a more immunogenic phenotype than the HRG.

### *3.6. Drug sensitivity between different risk groups*

Furthermore, we evaluated the variation in drug sensitivity across various risk groups, and the outcomes indicated that Temsirolimus, Gefitinib, Lapatinib, Lenalidomide, and Metformin had hindered $IC_{50}$ values in the LRG than in the HRG. Additionally, in the LRG, another 22 drugs exhibited higher $IC_{50}$ values than in the HRG (Fig. 7a-o and Supplementary Fig. S1). These results manifested that patients with different risk groups get advantages from the clinical drug therapy differently.

## 4. Discussion

EC, a female reproductive tract cancer, has recently become a younger trend [1]. Previously, surgery as the primary treatment was manifested to be linked to a good prognosis in early-stage EC patients. However, adjuvant therapy, including radiotherapy,

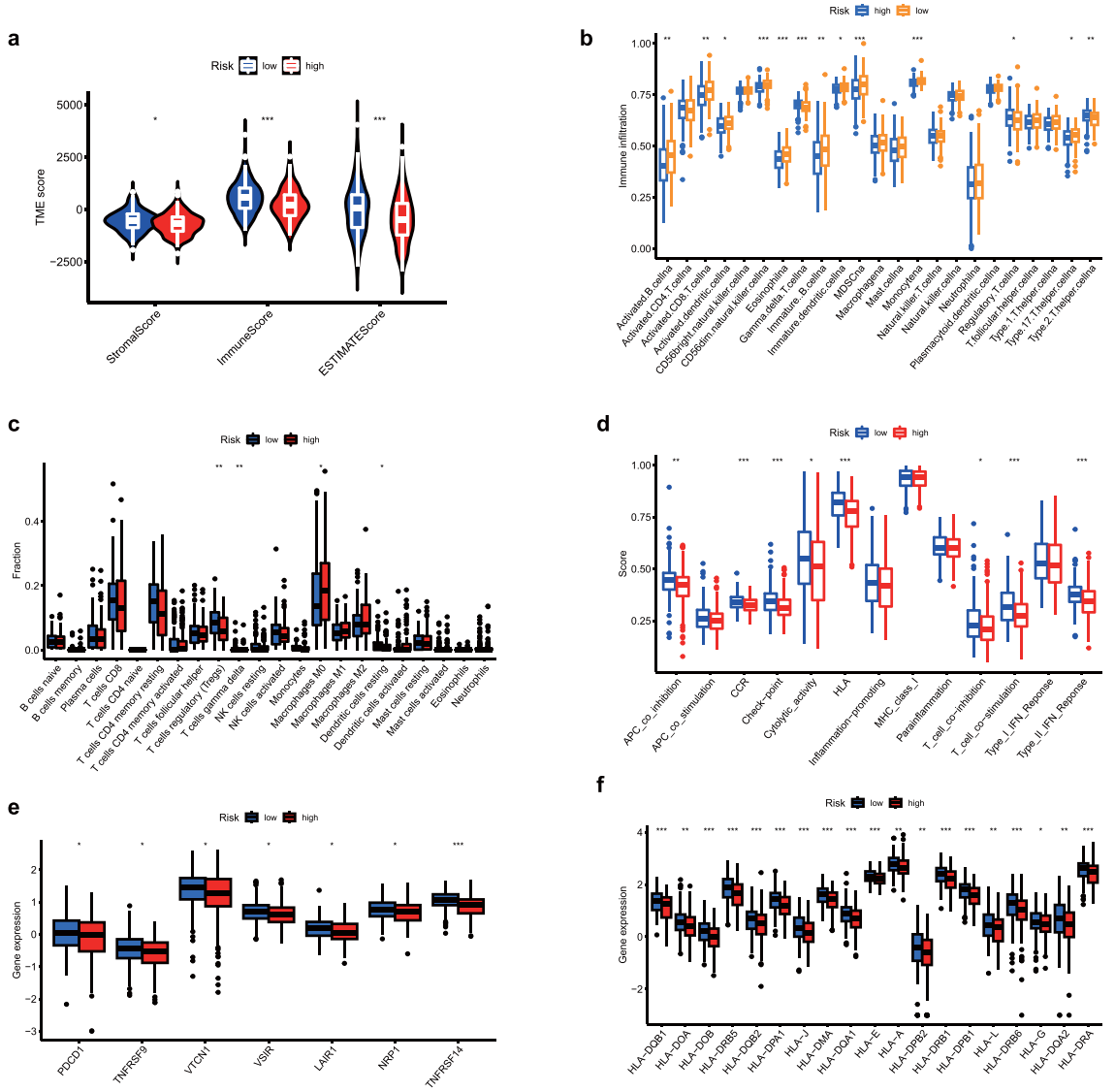



**Fig. 6.** The Immune properties of each PRS signature **a** The stromal, immune, and estimate scores were analysed between various groups. **b** The infiltration of immune cells expression levels was analysed between different groups. **c** CIBERSORT analysis plot was performed between various risk groups. **d** The immune-associated function analysis was illustrated between different groups. **e** The variation in immune checkpoint gene expression was analysed between various risk groups. **f** The variation in HLA expression was analysed between various risk groups. $^{*}P < 0.05$, $^{**}P < 0.01$, $^{***}P < 0.001$.

chemotherapy, or a combination of both, and endocrine therapy are required for EC patients with recurrence and metastasis, as these types of EC have a poor prognosis [22]. An elevating number of researchers have discovered the importance of prognosis prediction before postsurgical adjuvant therapy, and its importance is further highlighted by the increasing EC mortality rate [4]. Therefore, according to the TCGA classification, The patients with EC were categorised into four distinct molecular categories based on their prognosis: copy-number-low/TP53-wild-type, POLE/ultramutated (POLE), copy-number-high/TP53-mutant, and microsatellite-instable/hypermutated (MSI) [23]. The introduction of the TCGA molecular categorisation enhanced the precision of risk prediction in EC patients. Moreover, the POLEmut subtype, one of the four TCGA subtypes, also presented a better prognosis than other subtypes after the clinical treatment [24]. However, there is little knowledge of the connection between the POLEmut subtypes and a good prognosis.

With the ongoing immune therapy, the application of immunotherapy-related drugs in the clinical treatment of several refractory solid malignancies has gradually increased [25]. For example, ICP inhibitors that have been assessed in several studies also exhibited promise in treating different solid and refractory malignancies [26,27]. According to numerous studies, ICP therapy shows potential as a personalised therapeutic option for EC patients [28]. In the POLEmut subtype, patients with EC exhibit high numbers of

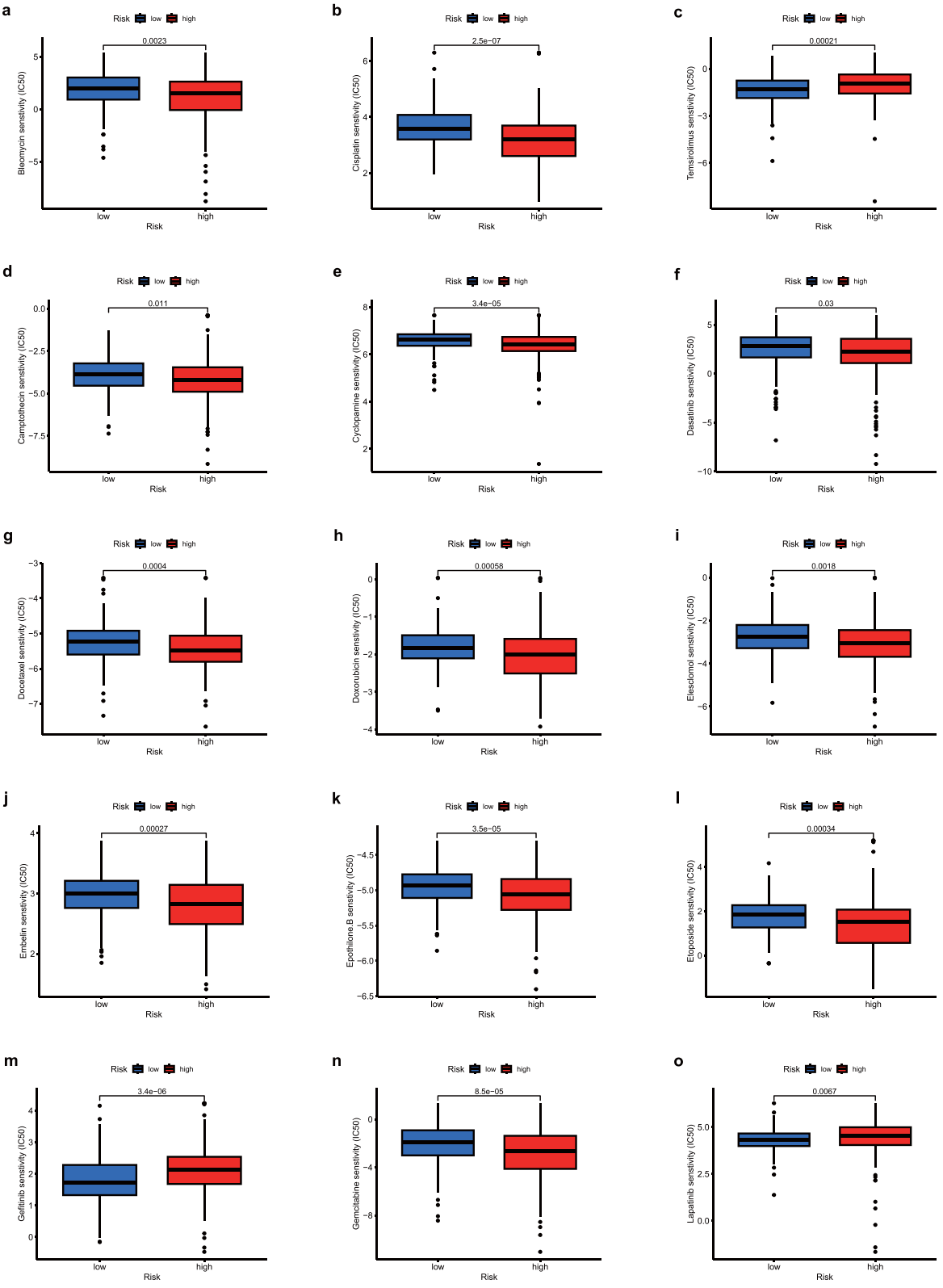



**Fig. 7.** Drug sensitivity between different risk groups **a-o** $IC_{50}$ values of different drugs were compared between different groups.

tumour-infiltrating lymphocytes and PD-1 and -L1 overexpression [29,30] because a high mutational load caused by POLEmut promotes increased expression of neoantigen and the immune system activation [25,31]. The above observations demonstrated that EC patients, especially the POLEmut subtype, may benefit from immunotherapy, suggesting that further investigation of the POLEmut subtype may provide additional useful information that could facilitate personalised treatment in the future.

Hence, we explored the *POLE*-related modules using WGCNA and established a PRS signature model using LASSO Cox regression analysis to validate prognostic risk in the TCGA UCEC datasets. Then, we employed the heatmap analysis to analyse the candidate gene expression levels linked to *POLE* (*ELF1, MMADHC,* and *AL021707.6*). Previous studies have indicated that ELF1, an ETS family member, participates in the tumorigenesis of various cancers and in regulating immunity to defend against pathogens, malignancies, and self-tolerance [32–35]. However, the other two genes were less reported in cancer progression or immunity. Moreover, we explored the candidate gene expression levels connected with the POLEmut subtype according to clinical stage and grade in this section (Supplementary Fig. S2). These results may suggest a potential correlation between candidate genes and clinical characteristics.

With this model, we also observed that the HRG had a significantly greater survival rate than the LRG. The LRG manifested elevated stromal score, raised immunological score, and greater extent of immune cell infiltration. In addition, the CIBERSORT analysis validated a strong association between a low-risk score and reduced levels of activated dendritic cells, together with elevated levels of regulatory T cells (Tregs) (Supplementary Fig. S3). Dendritic cells have been shown to be crucial in the antitumor response by attracting and facilitating the activation of antitumor T lymphocytes in the tumour microenvironment, as indicated by several studies [36]. Meanwhile, dendritic cells may also regulate the T cell differentiation priming towards Treg cells under the pressure of antitumor immunity [37,38]. Moreover, Treg cells participate in tumour development and act as a suppressor in antitumor immunity [39]. These findings may explain the poor prognosis in the LRG. Simultaneously, we observed that the LRG had a higher checkpoint score. The ICP gene expression levels were greater in the LRG contrasted with the HRG. ICP genes with high expression levels probably cultivate a microenvironment of immune suppression and support tumour immune escape [40]. Numerous clinical therapeutic analyses have also confirmed that ICP inhibitors are essential in tumour immunotherapy, including EC patients [41]. Hence, our established model may offer a promising prediction of immunotherapy in clinical treatment of ICP inhibitors.

We proposed to estimate the potential value of the PRS model in forecasting prognosis, immunotherapy efficacy, and drug sensitivity in EC patients. Our outcomes demonstrated the high potential value of this model. However, our study still has some limitations, such as the lack of multiple database analyses and clinical confirmation, which we will address in future studies to clarify the practical value of this model.

## 5. Conclusions

In brief, we created and verified a PRS model to accurately anticipate the OS of EC patients, and we proved its high predictive accuracy. Moreover, we identified candidate genes related to *POLE* (*ELF1, MMADHC,* and *AL021707.6*) via bioinformatics analyses. Furthermore, we evaluated the variation in immune infiltration and drug sensitivity between PRS risk groups. All these results provide information on the relationship between the POLEmut subtype and a good prognosis and provide a PRS model as a new technique for directing personalised clinical treatment.

## Data availability statement

The datasets produced and examined in the present investigation may be found in the TCGA UCEC repository, a free public database available to all researchers [https://www.cancer.gov/types/uterine]. We express our gratitude to the TCGA database for donating their platforms and to the authors for uploading their valuable datasets.

## Funding

The funding for this investigation was provided by the Nanjing Medical University Foundation of Jiangsu Province, China [Grant No. NMUB20210148] and the Affiliated Jiangning Hospital of Nanjing Medical University [Grant No. JNYYZXKY202114].

## Ethical approval

TCGA belong to public databases. Our work uses open-source data. Hence, the Research Ethics Committee at Jiangning Hospital has determined that ethical clearance is unnecessary.

## CRediT authorship contribution statement

**Wei Qiu:** Writing – original draft, Visualization, Formal analysis. **Runjie Zhang:** Validation, Resources, Data curation, Conceptualization. **Yingchen Qian:** Writing – review & editing, Supervision, Funding acquisition.

## Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to

influence the work reported in this paper.

## Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.heliyon.2024.e29548.

## References


[1] R.L. Siegel, K.D. Miller, N.S. Wagle, A. Jemal, Cancer statistics, 2023, CA Cancer J Clin 73 (2023) 17–48, https://doi.org/10.3322/caac.21763.

[2] N. Colombo, C. Creutzberg, F. Amant, T. Bosse, A. Gonzalez-Martin, J. Ledermann, C. Marth, R. Nout, D. Querleu, M.R. Mirza, C. Sessa, Group E-E-EECCW, ESMO-ESGO-ESTRO consensus Conference on endometrial cancer: diagnosis, treatment and follow-up, Ann. Oncol. 27 (2016) 16–41, https://doi.org/10.1093/annonc/mdv484.

[3] R.A. Brooks, G.F. Fleming, R.R. Lastra, N.K. Lee, J.W. Moroney, C.H. Son, K. Tatebe, J.L. Veneris, Current recommendations and recent progress in endometrial cancer, CA Cancer J Clin 69 (2019) 258–279, https://doi.org/10.3322/caac.21561.

[4] A. Travaglino, A. Raffone, A. Gencarelli, A. Mollo, M. Guida, L. Insabato, A. Santoro, G.F. Zannoni, F. Zullo, TCGA classification of endometrial cancer: the Place of carcinosarcoma, Pathol. Oncol. Res. 26 (2020) 2067–2073, https://doi.org/10.1007/s12253-020-00829-9.

[5] N. Cancer Genome Atlas Research, C. Kandoth, N. Schultz, A.D. Cherniack, R. Akbani, Y. Liu, H. Shen, A.G. Robertson, I. Pashtan, R. Shen, C.C. Benz, C. Yau, P. W. Laird, L. Ding, W. Zhang, G.B. Mills, R. Kucherlapati, E.R. Mardis, D.A. Levine, Integrated genomic characterization of endometrial carcinoma, Nature 497 (2013) 67–73, https://doi.org/10.1038/nature12113.

[6] A. Raffone, A. Travaglino, M. Mascolo, L. Carbone, M. Guida, L. Insabato, F. Zullo, TCGA molecular groups of endometrial cancer: pooled data about prognosis, Gynecol. Oncol. 155 (2019) 374–383, https://doi.org/10.1016/j.ygyno.2019.08.019.

[7] E. Stelloo, R.A. Nout, E.M. Osse, I.J. Jurgenliemk-Schulz, J.J. Jobsen, L.C. Lutgens, E.M. van der Steen-Banasik, H.W. Nijman, H. Putter, T. Bosse, C. L. Creutzberg, V.T. Smit, Improved risk assessment by integrating molecular and clinicopathological factors in early-stage endometrial cancer-combined analysis of the PORTEC cohorts, Clin. Cancer Res. 22 (2016) 4215–4224, https://doi.org/10.1158/1078-0432.CCR-15-2878.

[8] Q. Wu, N. Zhang, X. Xie, The clinicopathological characteristics of POLE-mutated/ultramutated endometrial carcinoma and prognostic value of POLE status: a meta-analysis based on 49 articles incorporating 12,120 patients, BMC Cancer 22 (2022) 1157, https://doi.org/10.1186/s12885-022-10267-2.

[9] Y. Xiao, D. Yu, Tumor microenvironment as a therapeutic target in cancer, Pharmacol. Ther. 221 (2021) 107753, https://doi.org/10.1016/j.pharmthera.2020.107753.

[10] D.C. Hinshaw, L.A. Shevde, The tumor microenvironment innately modulates cancer progression, Cancer Res. 79 (2019) 4557–4566, https://doi.org/10.1158/0008-5472.CAN-18-3962.

[11] V.S. Park, Z.F. Pursell, POLE proofreading defects: contributions to mutagenesis and cancer, DNA Repair 76 (2019) 50–59, https://doi.org/10.1016/j.dnarep.2019.02.007.

[12] A. Raffone, A. Travaglino, D. Raimondo, M.P. Boccellino, M. Maletta, G. Borghese, P. Casadio, L. Insabato, A. Mollo, F. Zullo, R. Seracchioli, Tumor-infiltrating lymphocytes and POLE mutation in endometrial carcinoma, Gynecol. Oncol. 161 (2021) 621–628, https://doi.org/10.1016/j.ygyno.2021.02.030.

[13] X. Guo, H. Xiao, S. Guo, L. Dong, J. Chen, Identification of breast cancer mechanism based on weighted gene coexpression network analysis, Cancer Gene Ther. 24 (2017) 333–341, https://doi.org/10.1038/cgt.2017.23.

[14] J. Liu, Y. Ma, W. Xie, X. Li, Y. Wang, Z. Xu, Y. Bai, P. Yin, Q. Wu, Lasso-based machine learning algorithm for predicting postoperative lung complications in elderly: a single-center retrospective study from China, Clin. Interv. Aging 18 (2023) 597–606, https://doi.org/10.2147/CIA.S406735.

[15] J. Ribbing, J. Nyberg, O. Caster, E.N. Jonsson, The lasso–a novel method for predictive covariate model building in nonlinear mixed effects models, J. Pharmacokinet. Pharmacodyn. 34 (2007) 485–517, https://doi.org/10.1007/s10928-007-9057-1.

[16] S. Afshar, T. Leili, P. Amini, I. Dinu, Introducing novel key genes and transcription factors associated with rectal cancer response to chemoradiation through co-expression network analysis, Heliyon 9 (2023) e18869, https://doi.org/10.1016/j.heliyon.2023.e18869.

[17] J. Zhang, S. Feng, M. Chen, W. Zhang, X. Zhang, S. Wang, X. Gan, Y. Zheng, G. Wang, Identification of potential crucial genes shared in psoriasis and ulcerative colitis by machine learning and integrated bioinformatics, Skin Res. Technol. 30 (2024) e13574, https://doi.org/10.1111/srt.13574.

[18] Y. Tian, K. Tao, S. Li, X. Chen, R. Wang, M. Zhang, Z. Zhai, Identification of m6A-related biomarkers in systemic lupus erythematosus: a bioinformation-based analysis, J. Inflamm. Res. 17 (2024) 507–526, https://doi.org/10.2147/JIR.S439779.

[19] A. Mayakonda, D.C. Lin, Y. Assenov, C. Plass, H.P. Koeffler, Maftools: efficient and comprehensive analysis of somatic variants in cancer, Genome Res. 28 (2018) 1747–1756, https://doi.org/10.1101/gr.239244.118.

[20] P. Langfelder, S. Horvath, WGCNA: an R package for weighted correlation network analysis, BMC Bioinf. 9 (2008) 559, https://doi.org/10.1186/1471-2105-9-559.

[21] S. Liu, Y. Zheng, S. Li, Y. Du, X. Liu, H. Tang, X. Meng, Q. Zheng, Integrative landscape analysis of prognostic model biomarkers and immunogenomics of disulfidptosis-related genes in breast cancer based on LASSO and WGCNA analyses, J. Cancer Res. Clin. Oncol. 149 (2023) 16851–16867, https://doi.org/10.1007/s00432-023-05372-z.

[22] J. Zhang, Z. Wang, R. Zhao, L. An, X. Zhou, Y. Zhao, H. Wang, An integrated autophagy-related gene signature predicts prognosis in human endometrial Cancer, BMC Cancer 20 (2020) 1030, https://doi.org/10.1186/s12885-020-07535-4.

[23] M. Alexa, A. Hasenburg, M.J. Battista, The TCGA molecular classification of endometrial cancer and its possible impact on adjuvant treatment decisions, Cancers 13 (2021), https://doi.org/10.3390/cancers13061478.

[24] L. Casey, N. Singh, POLE, MMR, and MSI testing in endometrial cancer: proceedings of the ISGyP companion society session at the USCAP 2020 annual meeting, Int. J. Gynecol. Pathol. 40 (2021) 5–16, https://doi.org/10.1097/PGP.0000000000000710.

[25] L. Musacchio, S.M. Boccia, G. Caruso, G. Santangelo, M. Fischetti, F. Tomao, G. Perniola, I. Palaia, L. Muzii, S. Pignata, P. Benedetti Panici, V. Di Donato, Immune checkpoint inhibitors: a promising choice for endometrial cancer patients? J. Clin. Med. 9 (2020) https://doi.org/10.3390/jcm9061721.

[26] Y. Wu, W. Chen, Z.P. Xu, W. Gu, PD-L1 distribution and perspective for cancer immunotherapy-blockade, knockdown, or inhibition, Front. Immunol. 10 (2019) 2022, https://doi.org/10.3389/fimmu.2019.02022.

[27] H.O. Alsaab, S. Sau, R. Alzhrani, K. Tatiparti, K. Bhise, S.K. Kashaw, A.K. Iyer, PD-1 and PD-L1 checkpoint signaling inhibition for cancer immunotherapy: mechanism, combinations, and clinical outcome, Front. Pharmacol. 8 (2017) 561, https://doi.org/10.3389/fphar.2017.00561.

[28] F. De Felice, C. Marchetti, V. Tombolini, P.B. Panici, Immune check-point in endometrial cancer, Int. J. Clin. Oncol. 24 (2019) 910–916, https://doi.org/10.1007/s10147-019-01437-7.

[29] R. Murali, R.A. Soslow, B. Weigelt, Classification of endometrial carcinoma: more than two types, Lancet Oncol. 15 (2014) e268–e278, https://doi.org/10.1016/S1470-2045(13)70591-6.

[30] S. Bellone, F. Centritto, J. Black, C. Schwab, D. English, E. Cocco, S. Lopez, E. Bonazzoli, F. Predolini, F. Ferrari, D.A. Silasi, E. Ratner, M. Azodi, P.E. Schwartz, A.D. Santin, Polymerase epsilon (POLE) ultra-mutated tumors induce robust tumor-specific CD4+ T cell responses in endometrial cancer patients, Gynecol. Oncol. 138 (2015) 11–17, https://doi.org/10.1016/j.ygyno.2015.04.027.

[31] S. Inaguma, J. Lasota, Z. Wang, A. Felisiak-Golabek, H. Ikeda, M. Miettinen, Clinicopathologic profile, immunophenotype, and genotype of CD274 (PD-L1)-positive colorectal carcinomas, Mod. Pathol. 30 (2017) 278–285, https://doi.org/10.1038/modpathol.2016.185.

[32] M. Hu, H. Li, H. Xie, M. Fan, J. Wang, N. Zhang, J. Ma, S. Che, ELF1 transcription factor enhances the progression of glioma via ATF5 promoter, ACS Chem. Neurosci. 12 (2021) 1252–1261, https://doi.org/10.1021/acschemneuro.1c00070.
[33] X.H. Xiao, S.Y. He, ELF1 activated long non-coding RNA CASC2 inhibits cisplatin resistance of non-small cell lung cancer via the miR-18a/IRF-2 signaling pathway, Eur. Rev. Med. Pharmacol. Sci. 24 (2020) 3130–3142, https://doi.org/10.26355/eurrev_202003_20680.
[34] D. Han, X. Li, Y. Cheng, Transcription factor ELF1 modulates cisplatin sensitivity in prostate cancer by targeting MEIS homeobox 2, Chem. Res. Toxicol. 36 (2023) 360–368, https://doi.org/10.1021/acs.chemrestox.2c00233.
[35] S. Gallant, G. Gilkeson, ETS transcription factors and regulation of immunity, Arch. Immunol. Ther. Exp. 54 (2006) 149–163, https://doi.org/10.1007/s00005-006-0017-z.
[36] S.K. Wculek, F.J. Cueto, A.M. Mujal, I. Melero, M.F. Krummel, D. Sancho, Dendritic cells in cancer immunology and immunotherapy, Nat. Rev. Immunol. 20 (2020) 7–24, https://doi.org/10.1038/s41577-019-0210-z.
[37] M. Clement, G. Fornasa, K. Guedj, S. Ben Mkaddem, A.T. Gaston, J. Khallou-Laschet, M. Morvan, A. Nicoletti, G. Caligiuri, CD31 is a key coinhibitory receptor in the development of immunogenic dendritic cells, Proc Natl Acad Sci U S A 111 (2014) E1101–E1110, https://doi.org/10.1073/pnas.1314505111.
[38] D.H. Munn, A.L. Mellor, Ido in the tumor microenvironment: inflammation, counter-regulation, and tolerance, Trends Immunol. 37 (2016) 193–207, https://doi.org/10.1016/j.it.2016.01.002.
[39] Y. Ohue, H. Nishikawa, Regulatory T (Treg) cells in cancer: can Treg cells be a new therapeutic target? Cancer Sci. 110 (2019) 2080–2089, https://doi.org/10.1111/cas.14069.
[40] T. Wang, L. Dai, S. Shen, Y. Yang, M. Yang, X. Yang, Y. Qiu, W. Wang, Comprehensive molecular analyses of a macrophage-related gene signature with regard to prognosis, immune features, and biomarkers for immunotherapy in hepatocellular carcinoma based on WGCNA and the LASSO algorithm, Front. Immunol. 13 (2022) 843408, https://doi.org/10.3389/fimmu.2022.843408.
[41] X. Wu, O. Snir, D. Rottmann, S. Wong, N. Buza, P. Hui, Minimal microsatellite shift in microsatellite instability high endometrial cancer: a significant pitfall in diagnostic interpretation, Mod. Pathol. 32 (2019) 650–658, https://doi.org/10.1038/s41379-018-0179-3.