Prognostic gene expression signature for high-grade serous ovarian cancer


1Division of Biostatistics, Department of Preventive Medicine, Keck School of Medicine, University of Southern California, Los Angeles, USA; 2School of Women’s and Children’s Health, Faculty of Medicine, University of New South Wales (UNSW) Sydney, Australia; 3CRUK Manchester Institute, The University of Manchester, Manchester, UK; 4Department of Health Science Research, Division of Epidemiology, Mayo Clinic, Rochester, USA; 5British Columbia’s Ovarian Cancer (OVCARE) Program, BC Cancer, Vancouver General Hospital, and the University of British Columbia, Vancouver; 6Department of Pathology and Laboratory Medicine, University of British Columbia, Vancouver, Canada; 7Department of Obstetrics and Gynecology, University of British Columbia, Vancouver, Canada; 8Peter MacCallum Cancer Center, Melbourne, Australia; 9Center for Cancer Prevention and Translational Genomics, Samuel Oschin Comprehensive Cancer Institute, Cedars-Sinai Medical Center, Los Angeles; 10Department of Pathology, Norris Comprehensive Cancer Center, Keck School of Medicine, University of Southern California, USA; 11Anatomical Pathology, Royal Women’s Hospital, Parkville, Australia; 12Department of Pathology, Duke University Hospital, Durham; 13Department of Health Science Research, Division of Biomedical Statistics and Informatics, Mayo Clinic, Rochester, USA; 14Department of Women’s Health, Tuebingen University Hospital, Tuebingen, Germany; 15The Jackson Laboratory for Genomic Medicine, Farmington, USA; 16Department of Genetics and Computational Biology, QiMER Berghofer Medical Research Institute, Brisbane, Queensland; 17Centre for Cancer Research, The Westmead Institute for Medical Research, The University of Sydney, Sydney; 18Sir Peter MacCallum Department of Oncology, The University of Melbourne, Parkville, Australia; 19Department of Obstetrics and Gynecology, Division of Gynecologic Oncology, Royal Alexandra Hospital, Edmonton; 20Alberta Health Services-Cancer Care, Calgary; 21Department of Pathology and Laboratory Medicine, University of Calgary, Foothills Medical Center, Calgary, Canada; 22Division of Cancer and Ovarian Cancer Action Research Centre, Department Surgery & Cancer, Imperial College London, London; 23Institute of Cancer Sciences, University of Glasgow; 24Cancer Research UK Cambridge Institute, University of Cambridge, Cambridge, UK; 25Medical Oncology Service, Hospital Sant Pau, Barcelona; 26HM Hospitales D Centro Integral Oncológico HM Clara Campal, Madrid; 27Medical Oncology Service, Hospital Universitario Funçao Alcorcon, Alcorcón, Spain; 28Division of Gynecologic Oncology, NorthShore University HealthSystem, University of Chicago, Evanston; 29Program in Epidemiology, Division of Public Health Sciences, Fred Hutchinson Cancer Research Center, Seattle; 30Department of Epidemiology, University of Washington, Seattle; 31Department of Systems Pharmacology and Translational Therapeutics, Perelman School of Medicine, University of Pennsylvania, Philadelphia; 32Department of Genetics, Geisel School of Medicine at Dartmouth, Hanover, USA; 33Division of Cancer Epidemiology, German Cancer Research Center (DKFZ), Heidelberg; 34Department of Pathology, Institute of Pathology, University Hospital Heidelberg, Heidelberg; 35National Center for Tumor Diseases, University Hospital Heidelberg and German Cancer Research Center, Heidelberg, Germany; 36David Geffen School of Medicine, Department of Obstetrics and Gynecology, University of California at Los Angeles, Los Angeles; 37Women’s Cancer Program at the Samuel Oschin Comprehensive Cancer Institute, Cedars-Sinai Medical Center, Los Angeles, USA; 38Department of Genetics and Pathology, Pomeranian Medical University, Szczecin; 39Department of Gynecological Surgery and Gynecological Oncology of Adults and Adolescents, Pomeranian Medical University, Szczecin, Poland; 40Department of Pathology, University of Michigan School of Public Health, Ann Arbor; 41Department of Preventive Medicine, Keck School of Medicine, University of Southern California Norris Comprehensive Cancer Center, Los Angeles; 42Department of Epidemiology and Biostatistics, Memorial Sloan-Kettering Cancer Center, New York; 43Department of Preventive Medicine, Keck School of Medicine, University of Southern California, Los Angeles, USA; 44Centre for Cancer Genetic Epidemiology, Department of Oncology, University of Cambridge, Cambridge; 45Department of Pathology, Addenbrooke’s Hospital NHS Trust, Cambridge; 46MRC Clinical Trials Unit at UCL, Institute of Clinical Trials & Methodology, University College London, London; 47Department of Women’s Cancer, Institute for Women’s Health, University College London, London; 48Department of Pathology, Barts Health National Health Service Trust, London, UK; 49Department of Gynaecological Oncology, Westmead Hospital, Sydney; 50Pathology West ICPMR Westmead, Westmead Hospital, The University of Sydney, Sydney; 51University of Western Sydney at Westmead Hospital, Sydney; 52The Crown Princess Mary Cancer Centre Westmead, Sydney-West Cancer Network, Westmead Hospital, Sydney, Australia; 53Department of Laboratory Medicine and Pathology, Division of Anatomic Pathology, Mayo Clinic, Rochester; 54Department of Oncology, Mayo Clinic,
Background: Median overall survival (OS) for women with high-grade serous ovarian cancer (HGSOC) is ~4 years, yet survival varies widely between patients. There are no well-established, gene expression signatures associated with prognosis. The aim of this study was to develop a robust prognostic signature for OS in patients with HGSOC.

Patients and methods: Expression of 513 genes, selected from a meta-analysis of 1455 tumours and other candidates, was measured using NanoString technology from formalin-fixed paraffin-embedded tumour tissue collected from 3769 women with HGSOC from multiple studies. Elastic net regularization for survival analysis was applied to develop a prognostic model for 5-year OS, trained on 2702 tumours from 15 studies and evaluated on an independent set of 1067 tumours from six studies.

Results: Expression levels of 276 genes were associated with OS (false discovery rate < 0.05) in covariate-adjusted single-gene analyses. The top five genes were TAP1, ZFHX4, CXCL9, FBN1 and PTPER3 (P < 0.001). The best performing prognostic signature included 101 genes enriched in pathways with treatment implications. Each gain of one standard deviation in the gene expression score conferred a greater than twofold increase in risk of death [hazard ratio (HR) 2.35, 95% confidence interval (CI) 2.02–2.71; P < 0.001]. Median survival [HR (95% CI)] by gene expression score quintile was 9.5 (8.3 to --), 5.4 (4.6–7.0), 3.8 (3.3–4.6), 3.2 (2.9–3.7) and 2.3 (2.1–2.6) years.

Conclusion: The OTTA-SPOT (Ovarian Tumor Tissue Analysis consortium - Stratified Prognosis of Ovarian Tumours) gene expression signature may improve risk stratification in clinical trials by identifying patients who are least likely to achieve 5-year survival. The identified novel genes associated with the outcome may also yield opportunities for the development of targeted therapeutic approaches.

Key words: formalin-fixed paraffin-embedded, gene expression, high-grade serous ovarian cancer, overall survival, prognosis

INTRODUCTION

Epithelial ovarian cancer (EOC) causes ~125 000 deaths globally every year, and long-term survival rates have changed little in the past three decades. Approximately 70% of women with EOC are diagnosed with advanced stage disease (stages III/IV), and fewer than 50% will survive more than 5 years. There are five major EOC histotypes: high-grade serous, low-grade serous, endometrioid, clear cell and mucinous. High-grade serous ovarian cancer (HGSOC) comprises about two-thirds of cases, is responsible for most deaths and is characterized by profound genomic and clinical heterogeneity.

The most informative prognostic factors for HGSOC are International Federation of Gynecology and Obstetrics (FIGO) stage, residual disease following debulking surgery, BRCA1 or BRCA2 germline mutation and tumour-infiltrating lymphocyte scores. Patients with HGSOC who carry a loss-of-function germline mutation in BRCA1 or BRCA2 have an increased sensitivity to platinum-based chemotherapy and PARP inhibitor treatment and a medium-term survival advantage. However, the frequent development of drug-resistant disease limits the effectiveness of current therapies. Gene expression data have been used to define four tumour molecular subtypes of HGSOC (C1/mesenchymal, C2/immune, C4/differentiated and C5/proliferative).
Using transcriptome-wide data from fresh frozen tissues, The Cancer Genome Atlas (TCGA) project used 215 tumours to identify an overall survival (OS) expression signature of 193 genes that has been validated on three other HGSOC gene expression datasets.12

Despite these findings, gene expression biomarkers have not been implemented clinically owing to several important shortcomings. The majority of the individual markers comprising the 193 gene signature were not statistically significant across all studies, suggesting that the signature may not be robust. The sample sizes in other discovery efforts have been too small for robust statistical inference.12 In addition, previous studies used fresh frozen samples, resulting in logistic and cost barriers to examining large clinically relevant datasets, and translation to the clinical setting.

The aim of this study was to identify a robust and clinically applicable prognostic HGSOC profile that can be applied to formalin-fixed paraffin-embedded (FFPE) tumour tissue.

**PATIENTS AND METHODS**

Twenty studies provided pretreatment FFPE tumour samples from 4071 women diagnosed with HGSOC (supplementary Table S1, available at Annals of Oncology online). All HGSOC cases with available tissue were included. During this period, patients with HGSOC were treated with chemotherapy (carboplatin and paclitaxel) after primary debulking surgery. Study protocols were approved by the respective Institutional Review Board/Ethics Approval Committee for each site (supplementary Table S1, available at Annals of Oncology online).

A schematic of the overall study design is shown in Figure 1. There were four main components: gene selection, gene expression assay, development of prognostic gene signature in a training set and validation of prognostic signature in an independent validation set.

**Gene selection**

Candidate prognostic genes were identified by carrying out an individual participant meta-analysis of six transcriptome-wide microarray studies,11–16 which included tumour samples from 1455 participants. Association of gene expression with OS was evaluated by Cox proportional hazards regression adjusted for molecular subtype (supplementary Table S2, available at Annals of Oncology online). In total, 200 genes from the meta-analysis, most achieving a permutation-based false discovery rate (FDR)17 of <0.05, and an additional 313 candidate genes based on the literature and unpublished data were selected (supplementary Tables S3 and S4, and supplementary Figure S1, available at Annals of Oncology online; for more details see supplementary Material, available at Annals of Oncology online). Five genes, RPL19, ACTB, PGK1, SDHA and POLR1B, were included as house-keeping genes for normalization.

**Gene expression assay in tumour samples from study participants**

FFPE tumour samples were processed with the NanoString nCounter technology at three different locations: Vancouver, Los Angeles and Melbourne. A control set of 48 FFPE tumour samples was run at each location and the average intraclass correlation coefficient was 0.987. Approximately 2% of the samples were run in duplicate and the average Spearman’s correlation coefficient \( r_s \) was 0.995. Single-patient classification methods were used with reference samples to control for batch effects.18 The data in this publication have been deposited in NCBI’s Gene Expression Omnibus19; GEO Series accession number GSE132342 (https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE132342). A total of 3329 samples passed quality control of which 3769 had survival data and assessable gene expression for 513 genes. Data can be found in NCBI GEO: Accession numbers GSE132342 and GPL26748.

**Overall survival analysis of individual genes**

Samples that contributed to the meta-analysis dataset \( n = 211 \) were removed from subsequent selected analyses to ensure independence of study samples between the gene selection and final survival analysis. Time-to-event analyses were carried out for OS with right censoring at 10 years and left truncation of prevalent cases. Associations between log-transformed normalized gene expression and survival time were tested using likelihood ratio tests with Cox proportional hazards models adjusted for age, race and stage, and stratified by study. Patients with missing race or stage information were assigned to ‘unknown’ categories. Age was modelled using a B-spline with a knot at the median age, which yielded a better fit than using knots at quartiles or categorical variables. Stage was dichotomized into early (FIGO stage I/II) and advanced (FIGO stage III/IV). Genes were scaled to have a standard deviation of one, so hazard ratios (HRs) correspond to a change of one standard deviation. A Benjamini–Hochberg FDR of <0.05 was used to identify notable associations. Because the expression of genes can be correlated, an analysis of correlated genes was performed using data from TCGA. Advanced stage ovarian cancer usually has disease spread throughout the abdomen, and therefore sensitivity analyses were performed to assess effects of the anatomical location of tumour samples included in the study by removing observations corresponding to samples known to be extraovarian \( ( n = 437 ) \).

**Prognostic signature development and validation**

Studies were initially randomized to training set \( N = 14 \) and validation set \( N = 6 \). The TRI study was randomized to the validation set, but, because 107 samples were part of the meta-analysis data used for gene selection, the study was split, so those 107 samples were included in the model training dataset. Thus 2702 samples from 15 studies were used for model training and 1067 samples from 6 studies were used for validation (supplementary Table S1, available at Annals of Oncology online). In the training set, four
Figure 1. Schematic of study design.

*The TRI study was split across the training and validation sets due to 107 samples overlapping with the meta-analysis. GWAS, genome-wide association studies; HGSOC, high-grade serous ovarian cancer.
modelling approaches (stepwise regression, elastic net regularized regression, boosting and random survival forests) were applied to construct competing gene expression–based biomarkers. Each was evaluated in the training data using 10-fold cross-validation for its prognostic value for OS at 2 and 5 years of follow-up using an area under the curve (AUC) measure derived from receiver operator characteristic analysis (see supplementary Material, available at Annals of Oncology online, for additional details). The best performing method, elastic net regularized regression, was applied to the full training set to determine the final gene signature and scoring method, which was then evaluated using the independent testing set. All models were constrained to include age and stage, where age was modelled as categorical based on quartiles of the training dataset with groups aged <53, 53–59, 60–66, and ≥67. Stage was modelled as described earlier for the OS individual gene analysis.

RESULTS

Association of expression of individual genes with OS in HGSOC

In a gene-by-gene analysis of the full dataset adjusted for age, race and stage, and stratified by study, 276 of the 513 selected genes were associated with OS (FDR < 0.05). Of these, 138 were selected from the meta-analysis of six published microarray studies (supplementary Table S2, available at Annals of Oncology online)11–16 and 144 from candidate gene approaches (supplementary Tables S5 and S6, available at Annals of Oncology online). HRs for one standard deviation change in gene expression ranged from 0.84 to 1.19, with multiple genes exhibiting associations at very stringent significance levels (e.g. 19 genes with $P < 1 \times 10^{-8}$, supplementary Tables S5 and S6, available at Annals of Oncology online). The five most significant genes were TAP1, ZFHX4, CXCL9, FBN1 and PTGER3 (Table 1). We did not find extensive evidence of high co-expression between these five genes and genes measured in TCGA project (supplementary Table S7, available at Annals of Oncology online). In sensitivity analyses we found that excluding samples from omentum and other extra-ovarian sites did not substantially affect the results (supplementary Tables S8 and S9, available at Annals of Oncology online).

Development of a novel prognostic gene signature

The four predictive modelling approaches that were evaluated in the training data using 10-fold cross-validation yielded median AUCs that ranged from 0.69 to 0.73 for 2-year OS and 0.69 to 0.74 for 5-year survival (supplementary Figure S2, available at Annals of Oncology online) with better prediction of 5-year OS than of 2-year OS. The elastic net approach yielded the highest median AUC for both 2- and 5-year OS and was selected for final development of the signature. Using the model on the full training dataset resulted in a prognostic signature of 101 genes in addition to age and stage (supplementary Table S10, available at Annals of Oncology online). Of these, 66 genes were associated with OS (FDR < 0.05) in the single gene models. There was no obvious subset of signature genes that performed as well or nearly as well as the full 101 gene signature (supplementary Figure S3, available at Annals of Oncology online).

Performance of the signature including age and stage was AUC 0.69 [95% confidence interval (CI) 0.65–0.73] and AUC 0.75 [95% CI 0.72–0.78] for 2- and 5-year OS, respectively (Figures 2 and 3; supplementary Figure S4, available at Annals of Oncology online). This was substantially better than age and stage alone with AUC 0.61 [95% CI 0.57–0.65] and AUC 0.62 (95% CI 0.59–0.67) for 2- and 5-year OS, respectively, particularly for the 5-year OS outcome with non-overlapping 95% CI. One standard deviation change in the gene expression score was associated with an HR of 2.35 [95% CI 2.02–2.71]; $P = 5.1 \times 10^{-31}$, and median survival [HR (CI)] varied substantially across quintiles of the gene expression score [9.5 (8.3 to –), 5.4 (4.6–7.0), 3.8 (3.3–4.6), 3.2 (2.9–3.7) and 2.3 (2.1–2.6) years, respectively, from smallest to largest quintile (Q1–Q5); Table 2].

For a subset of cases, there was clinical and experimental data for known prognostic factors. All samples had molecular subtype classification (Talhouk et al.20); residual disease was known for 1771 cases, primary chemotherapy treatment for 687, germline BRCA mutation status for 904 and nuclear CD8 tumour-infiltrating lymphocyte counts6 for 1111 (supplementary Table S11, available at Annals of Oncology online). When examined by quintile of gene expression score there were differences, as expected, for each of the known prognostic factors, including age and stage that were included in the model (Table 3). However, in sensitivity analyses, applying the signature to specific

<table>
<thead>
<tr>
<th>Gene</th>
<th>HR (95% CI)</th>
<th>$P$</th>
<th>Selection</th>
<th>Correlated gene</th>
<th>rs</th>
</tr>
</thead>
<tbody>
<tr>
<td>TAP1</td>
<td>0.84 (0.80–0.87)</td>
<td>8.3 x 10^{-18}</td>
<td>Meta</td>
<td>PSMB9</td>
<td>0.89</td>
</tr>
<tr>
<td>ZFHX4</td>
<td>1.19 (1.14–1.25)</td>
<td>1.4 x 10^{-15}</td>
<td>Meta</td>
<td>LOC100192378</td>
<td>0.74</td>
</tr>
<tr>
<td>CXCL9</td>
<td>0.85 (0.82–0.88)</td>
<td>1.8 x 10^{-15}</td>
<td>Meta and candidate</td>
<td>CXCR6</td>
<td>0.89</td>
</tr>
<tr>
<td>FBN1</td>
<td>1.18 (1.13–1.24)</td>
<td>4.2 x 10^{-14}</td>
<td>Candidate</td>
<td>SPARC</td>
<td>0.91</td>
</tr>
<tr>
<td>PTGER3</td>
<td>1.18 (1.13–1.24)</td>
<td>1.2 x 10^{-13}</td>
<td>Meta</td>
<td>COL6A1</td>
<td>0.67</td>
</tr>
</tbody>
</table>

CI, confidence interval; HR, hazard ratio.

a Most correlated gene according to Spearman’s rank correlation coefficient, $r_s$, computed in The Cancer Genome Atlas (TCGA) Ovarian Serous Cystadenocarcinoma RNA-seq dataset.

b SPARC was included in this project and was less significant.
patient groups, a robustness of stratification was demonstrated, suggesting that the prognostic power of the signature is not explained by the individual factors, residual disease, treatment, BRCA status or CD8 score (Figure 3 and supplementary Figures S5–S7, available at Annals of Oncology online). The signature score showed modest differences by molecular subtype (supplementary Figure S8, available at Annals of Oncology online), and adjusting for molecular subtype in the Cox analysis resulted in only minor changes to the HR estimates for signature quintiles (Table 2). The signature was shown to be prognostic within a homogenous group of 316 stage IIIC cases with no residual disease, within early stage cases (FIGO IA and IB) and within patients whose samples were collected from the omentum (supplementary Figures S9 and S10, available at Annals of Oncology online). Analysis of the signature score for paired ovarian and omental tissue from 42 of the cases showed a highly significant Pearson’s correlation coefficient, $r = 0.79 \ (P = 5.4 \times 10^{-10}$; supplementary Figure S11, available at Annals of Oncology online).

A gene set enrichment analysis was performed for the 101 genes in the signature, as well as for genes correlated with signature genes achieving $r^2 > 0.75$ (supplementary Table S12, available at Annals of Oncology online). For the correlated gene analysis, the three most significant pathways involved the immune system, including the adaptive immune system and cytokine signalling. A further 10 immune pathways were significantly enriched and included interferon signalling, innate immune system and TCR signalling and antigen presentation pathways. Restricting to the signature genes only, there was also enrichment in the immune system, but the top two pathways were PI-3K (phosphoinositide 3-kinase) cascade and GPCR (G protein–coupled receptor) ligand binding. Four other pathways were related to the cell cycle and mitosis, with the remaining enriched for fibroblast growth factor receptor (FGFR) and epidermal growth factor receptor (ERBB) signalling, and one pathway related to homologous combination repair.

**DISCUSSION**

In a large-scale study of patients with HGSOC, we identified a 101-gene expression signature able to predict clinically relevant differences in OS. Using methods that are both economical and applicable to standard clinical sampling techniques, we showed that the signature performs substantially better than age and stage alone for prognosis of both 2- and 5-year OS. The number of patients and samples included in this study is an order of magnitude greater than previous comparable studies of gene expression and OS in patients with HGSOC.\textsuperscript{12,21,22} Thus, we have been able to more precisely quantify the prognostic value of gene expression.

We report definitive associations between OS and expression of 276 genes. Of the five most significant genes (TAP1, ZFHX4, CXCL9, FBN1 and PTGER3), four have been previously reported to be associated with survival in HGSOC. The top prognostic gene, TAP1, is involved in the antigen-presenting pathway. Expression was reduced in metastatic HGSOC, positively associated with OS,\textsuperscript{23} as observed here, and linked to tumour regression in response to treatment.\textsuperscript{24} Further, hypomethylation of TAP1 was associated with improved time to disease recurrence.\textsuperscript{25}

CXCL9 is a chemokine that mediates the recruitment of T cells to solid tumours.\textsuperscript{26} High expression of intratumoural CXCL9 was associated with higher OS\textsuperscript{27} and higher lymphocytic infiltration, which is also a robust prognostic factor in HGSOC\textsuperscript{11,28} and a feature of the immunoreactive HGSOC molecular subtype.\textsuperscript{11} CXCL9 has also been proposed as a therapeutic target due to evidence that it inhibits angiogenesis and promotes antitumour adaptive immunity.\textsuperscript{28–31} Strikingly, the signature was able to further refine prognostic groups within patients with high tumour-infiltrating lymphocyte counts suggesting that CXCL9 and TAP1 expression may be strong indicators of immune competency in HGSOC.

FBN1 is an extracellular matrix protein previously found to be a biomarker associated with early recurrence in patients with ovarian cancer who are initially sensitive to chemotherapy\textsuperscript{32} and strongly correlated with desmoplasia in HGSOC. The prostaglandin E$_2$ receptor PTGER3 is expressed in ovarian tumour cells and is associated with relapse-free survival.\textsuperscript{13} By contrast, ZFHX4 does not have previous associations with HGSOC.

Associations between the expression of specific genes in tumour tissues and OS in patients with HGSOC may suggest new drug targets and lead to insights into biological variation in treatment response. For example, cases in the Q5 quintile with the poorest outcome had increased expression of IGF2, FGFR1 and MYC, a possible argument for the use of...
IGFR1, FGFR, bromodomain (MYC) or a combination of PARP and CDK4/6 inhibitors (MYC). More immediately, the signature may help clinicians identify patients most in need of intervention, such as patients that could potentially benefit from neoadjuvant chemotherapy (NACT). Alternatively, in clinical trials it could be used to stratify randomization by patients’ risk, thereby reducing heterogeneity within subgroups and increasing heterogeneity between subgroups. The signature will be incorporated into future prospective clinical trials to determine if it can predict response to specific treatments.

Measurement of the signature required standard FFPE tissue used in routine histopathology. In addition, data preprocessing and normalization were conducted on an individual level, thus translatable to a general patient population. That is, 5-year OS prognosis of future patients

<table>
<thead>
<tr>
<th>Quintile</th>
<th>N</th>
<th>Deaths</th>
<th>Median survival</th>
<th>HR (95% CI) Adjusted for age and stage</th>
<th>Adjusted for molecular subtype age and stage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Q1</td>
<td>214</td>
<td>81</td>
<td>9.47 (8.32 to –)</td>
<td>0.44 (0.33–0.58)</td>
<td>0.34 (0.22–0.55)</td>
</tr>
<tr>
<td>Q2</td>
<td>213</td>
<td>117</td>
<td>5.38 (4.63–6.97)</td>
<td>0.73 (0.57–0.93)</td>
<td>0.71 (0.55–0.91)</td>
</tr>
<tr>
<td>Q3</td>
<td>213</td>
<td>145</td>
<td>3.80 (3.34–4.60)</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>Q4</td>
<td>213</td>
<td>158</td>
<td>3.23 (2.85–3.68)</td>
<td>1.56 (1.25–1.96)</td>
<td>1.56 (1.24–1.97)</td>
</tr>
<tr>
<td>Q5</td>
<td>214</td>
<td>179</td>
<td>2.27 (2.09–2.62)</td>
<td>2.23 (1.78–2.78)</td>
<td>2.11 (1.67–2.67)</td>
</tr>
</tbody>
</table>

CI, confidence interval; HR, hazard ratio.

* Median survival (95% CI) in years for patients in the validation set.
samples used in this study were chemo-naïve, as few
signature. classi all of the individual genes used in the molecular subtype
able to differences among molecular subtypes. In addition,
strong performance of the signature is not solely attribut-
the subtype-adjusted meta-analysis, suggesting that the
substantial proportion of signature genes were identi
molecular subtype can improve biomarker performance. A
in future research. Another important question is whether
was beyond the scope of this study but should be pursued

cancer.

Table 3. Clinical data for the 3769 patients that passed quality control and the percentage of patients in each quintile of the gene expression score

<table>
<thead>
<tr>
<th></th>
<th>Total</th>
<th>Q1</th>
<th>Q2</th>
<th>Q3</th>
<th>Q4</th>
<th>Q5</th>
<th>P value</th>
</tr>
</thead>
<tbody>
<tr>
<td>N</td>
<td>3769</td>
<td>754</td>
<td>754</td>
<td>753</td>
<td>754</td>
<td>754</td>
<td></td>
</tr>
<tr>
<td>Median survival (years)</td>
<td>4.1</td>
<td>9.5</td>
<td>5.4</td>
<td>3.8</td>
<td>3.2</td>
<td>2.3</td>
<td></td>
</tr>
<tr>
<td>% 5-year survival</td>
<td>41</td>
<td>75</td>
<td>57</td>
<td>39</td>
<td>25</td>
<td>10</td>
<td></td>
</tr>
<tr>
<td>Age median</td>
<td>63</td>
<td>58</td>
<td>57</td>
<td>61</td>
<td>64</td>
<td>70</td>
<td></td>
</tr>
<tr>
<td>Age quartile Q1</td>
<td>894</td>
<td>30.8</td>
<td>31.3</td>
<td>20.0</td>
<td>13.4</td>
<td>4.5</td>
<td></td>
</tr>
<tr>
<td>Age quartile Q2</td>
<td>838</td>
<td>21.5</td>
<td>20.0</td>
<td>22.9</td>
<td>21.2</td>
<td>14.3</td>
<td></td>
</tr>
<tr>
<td>Age quartile Q3</td>
<td>961</td>
<td>16.0</td>
<td>20.2</td>
<td>21.4</td>
<td>23.6</td>
<td>18.7</td>
<td></td>
</tr>
<tr>
<td>Age quartile Q4</td>
<td>1076</td>
<td>3.5</td>
<td>10.4</td>
<td>16.4</td>
<td>21.3</td>
<td>38.5</td>
<td></td>
</tr>
<tr>
<td>FIGO stage I/II</td>
<td>607</td>
<td>97.4</td>
<td>2.6</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>&lt;1 x 10^−5</td>
</tr>
<tr>
<td>FIGO stage III/IV</td>
<td>3067</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Primary chemo1 1</td>
<td>136</td>
<td>16.2</td>
<td>22.1</td>
<td>23.5</td>
<td>19.1</td>
<td>19.1</td>
<td>0.163</td>
</tr>
<tr>
<td>Primary chemo2 2</td>
<td>190</td>
<td>16.3</td>
<td>20.0</td>
<td>21.6</td>
<td>22.1</td>
<td>20.0</td>
<td></td>
</tr>
<tr>
<td>Primary chemo3 3</td>
<td>361</td>
<td>11.1</td>
<td>16.9</td>
<td>22.4</td>
<td>20.5</td>
<td>29.1</td>
<td></td>
</tr>
<tr>
<td>Residual disease: No</td>
<td>614</td>
<td>32.4</td>
<td>22.1</td>
<td>17.8</td>
<td>15.5</td>
<td>12.2</td>
<td>&lt;1 x 10^−5</td>
</tr>
<tr>
<td>Germline BRCA1 mutation</td>
<td>130</td>
<td>23.8</td>
<td>31.5</td>
<td>26.2</td>
<td>11.5</td>
<td>6.9</td>
<td>2.22 x 10^−7</td>
</tr>
<tr>
<td>Germline BRCA2 mutation</td>
<td>71</td>
<td>28.2</td>
<td>26.8</td>
<td>18.3</td>
<td>18.3</td>
<td>8.5</td>
<td></td>
</tr>
<tr>
<td>Germline no BRCA1/2 mutation</td>
<td>663</td>
<td>19.6</td>
<td>16.7</td>
<td>18.7</td>
<td>20.7</td>
<td>24.3</td>
<td></td>
</tr>
<tr>
<td>CD8 TIL score 0</td>
<td>192</td>
<td>19.8</td>
<td>14.6</td>
<td>12.5</td>
<td>21.4</td>
<td>31.8</td>
<td>2.46 x 10^−4</td>
</tr>
<tr>
<td>CD8 TIL score 1–2</td>
<td>186</td>
<td>18.3</td>
<td>14.0</td>
<td>18.8</td>
<td>21.5</td>
<td>27.4</td>
<td></td>
</tr>
<tr>
<td>CD8 TIL score 3–19</td>
<td>515</td>
<td>19.8</td>
<td>24.1</td>
<td>20.8</td>
<td>17.9</td>
<td>17.5</td>
<td></td>
</tr>
<tr>
<td>Molecular subtype C1,MES</td>
<td>1105</td>
<td>5.4</td>
<td>10.4</td>
<td>20.7</td>
<td>27.4</td>
<td>36.0</td>
<td>&lt;1 x 10^−5</td>
</tr>
<tr>
<td>Molecular subtype C2,IMM</td>
<td>907</td>
<td>23.2</td>
<td>28.8</td>
<td>21.2</td>
<td>16.2</td>
<td>10.7</td>
<td></td>
</tr>
<tr>
<td>Molecular subtype C4,DIF</td>
<td>1144</td>
<td>32.6</td>
<td>25.5</td>
<td>17.9</td>
<td>12.8</td>
<td>11.2</td>
<td></td>
</tr>
<tr>
<td>Molecular subtype C5,PRO</td>
<td>613</td>
<td>18.1</td>
<td>14.0</td>
<td>20.7</td>
<td>25.8</td>
<td>21.4</td>
<td></td>
</tr>
<tr>
<td>FIGO stage IA and IB</td>
<td>111</td>
<td>96.4</td>
<td>3.6</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>&lt;1 x 10^−5</td>
</tr>
<tr>
<td>FIGO stage IIIC residual disease: No</td>
<td>316</td>
<td>6.3</td>
<td>31.0</td>
<td>24.4</td>
<td>20.9</td>
<td>17.4</td>
<td>6.24 x 10^−5</td>
</tr>
</tbody>
</table>

Q1 is the quintile with the best survival and Q5 the worst survival. Samples with missing data are reported in supplementary Table S31, available at Annals of Oncology online. P values for BRCA1/2 mutation status were calculated for BRCA1 or BRCA2 mutation versus no mutation.

FIGO, International Federation of Gynecology and Obstetrics; TIL, tumour-infiltrating lymphocyte.

* Treatment: 1 = known to have received first-line chemotherapy treatment of ≥4 cycles of IV carboplatin AUC 5 or 6 and paclitaxel 135 or 175 mg/m² every 3 weeks. 2 = known to have received first-line chemotherapy treatment of ≥4 cycles of IV carboplatin and paclitaxel three times weekly but at doses presumed to be carboplatin AUC 5 or 6 and paclitaxel 135 or 175 mg/m². 3 = all remaining cases with chemo regimens that do not fit criteria 1 or 2 and include unknown or no chemotherapy.

patients had NACT during the calendar period in which these samples were collected. Because the signature appears to be prognostic in omentum samples, future studies may assess the value in NACT patients, using pretreatment omental biopsies or post-treatment tumour samples. Future work will also address if the signature can predict platinum-refractory patients.

We have developed a robust prognostic signature for HGSOC that can be used to stratify patients and identify those in need of alternative treatments. Gene set enrichment analysis applied to the signature indicates an important role for the immune system in OS and supports further investigation of immune therapy in ovarian cancer. More generally, the identification here of high-confidence prognostic genes may lead to new hypotheses for targeted treatments.

ACKNOWLEDGEMENTS

We thank all the study participants who contributed to this study and all the researchers, clinicians and technical and administrative staff who have made this work possible. This project received technical and data management support from OVCARE’s core units, including the Cheryl Brown Ovarian Cancer Outcomes Unit and the Genetic Pathology
Evaluation Centre, and statistical analysis support from the Biostatistics Core of the Norris Comprehensive Cancer Center. The AOV study recognizes the valuable contributions from Mie Konno, Shuhong Liu, Michelle Darago, Faye Chambers and the staff at the Tom Baker Cancer Centre Translational Laboratories. We thank Olivier Tredan and Pierre Heudel as investigators on the TRIO14 study and Sandrine Berge-Montamat as assistant for clinical research. The Australian Ovarian Cancer Study gratefully acknowledges additional support from Ovarian Cancer Australia and the Peter MacCallum Foundation. The AOCs also acknowledges the cooperation of the participating institutions in Australia and acknowledges the contribution of the study nurses, research assistants and all clinical and scientific collaborators to the study. The complete AOCs Study Group can be found at www.aocstudy.org. We thank all of the women who participated in these research programs.

**FUNDING**

This work was funded by the National Institutes of Health/National Cancer Institute (NCI) Grants to SJR [grant number R01CA172404] and J.A. Doherty and MAR [grant number R01CA168758], the Canadian Institutes for Health Research (Proof-of-Principle I program, no grant number applicable) and the United States Department of Defense Ovarian Cancer Research Program [grant number OC110433]. J. Millstein and SJR received support from National Institutes of Health/National Cancer Institute [grant number P30CA014089] and J. Millstein received support from NIH/National Cancer Institute award number P01CA196569. MSA receives funding from the Janet D. Cottrelle Foundation Scholar’s program managed by the BC Cancer Foundation (no grant number applicable). JG was partially supported by the National Institutes of Health/ National Cancer Institute [grant number P30CA034196]. CW was a Career Enhancement Awardee of the Mayo Clinic SPORE in Ovarian Cancer [grant number P50 CA136393]. MJH received funding from Cancer Australia (1067110), DGH receives support from the Dr Chew Wei Memorial Professorship in Gynecologic Oncology, (no grant number applicable) the Canada Research Chairs program (Research Chair in Molecular and Genomic Pathology, no grant number applicable), and the Janet D. Cottrelle Foundation (no grant number applicable). MW receives funding from the European Union’s Horizon 2020 European Research Council Programme [grant number H2020 BRCA-ERC] under Grant Agreement No. 742432, as well as from the charity The Eve Appeal (https://eveappeal.org.uk/, no grant number applicable) and support from the National Institute for Health Research (NIHR, no grant number applicable) and the University College London Hospitals (UCLH) Biomedical Research Centre (no grant number applicable). GEK is supported by the Miriam and Sheldon Adelson Medical Research Foundation (no grant number applicable). BYK is funded by the American Cancer Society Early Detection Professorship [grant number SIOP-06-258-01-COUN] and the National Center for Advancing Translational Sciences (NCATS) [grant number UL1TR000124]. HRH is supported by the National Institutes of Health/National Cancer Institute [grant number K22 CA193860]. OVCARE (including the VAN study) receives core funding through the BC Cancer Foundation (no grant number applicable) and The VGH–UBC Hospital Foundation (authors AT, BG, DGH and MSA, no grant number applicable). The AOV study is supported by the Canadian Institutes of Health Research [grant number MOP-86727]. The Gynaecological Oncology Biobank at Westmead, a member of the Australasian Biospecimen Network- Oncology group, was funded by the National Health and Medical Research Council Enabling [grant numbers ID 310670, ID 628903] and the Cancer Institute NSW [grant numbers ID 12/RIG/1-17, 15/RIG/1-16]. The Australian Ovarian Cancer Study Group was supported by the U.S. Army Medical Research and Materiel Command [grant number DAMD17-01-1-0729]. The Cancer Council Victoria, Queensland Cancer Fund, The Cancer Council New South Wales, The Cancer Council South Australia, The Cancer Council Tasmania and The Cancer Foundation of Western Australia (Multi-State Applications 191, 211 and 182) and the National Health and Medical Research Council of Australia [grant numbers NHMRC, ID199600, ID400413, ID400281]. BriTROC-1 was funded by Ovarian Cancer Action (to IAM and JDB) [grant number 006] and supported by Cancer Research UK [grant numbers A15973, A15601, A18072, A17197, A19274, A19694] and the National Institute for Health Research Cambridge and Imperial Biomedical Research Centres (no grant number applicable). SEARCH was supported by Cancer Research UK [grant number A16561]. The University of Cambridge receives salary support for PDPP from the NHS Clinical Academic Reserve (no grant number applicable). Samples from the Mayo Clinic were collected and provided with support of the National Institutes of Health/National Cancer Institute (NCI) P50 CA136393 (ELG, GLK, SHK, MES). S. Orsulic was funded by the Department of Defense Award W81XWH-17-1-0144. MRC Clinical Trials Unit at UCL receives funding from The Eve Appeal (The Oak Foundation) with investigators supported by the National Institute for Health Research University College London Hospitals Biomedical Research Centre and MRC core funding (MR_UU_12023).

**DISCLOSURE**

BYK served on Invitae Corporation’s Advisory Board from 2017 to 2018. IAM has acted on the Advisory Boards for AstraZeneca, Clovis Oncology, Tesaro, Carrick Therapeutics and Takeda. His institution receives funding from AstraZeneca. RG is on the Advisory Boards for AstraZeneca, Tesaro, Clovis and Immunogen and does consultancy work for SOTIO. She has received support to attend conferences from AstraZeneca, Roche and Tesaro. Her institution has received research funding from Boehringer Ingelheim and Lilly/Ignyta and she is the national co-ordinating investigator for the UK for trials sponsored by AstraZeneca and Tesaro and site principal investigator for trials sponsored by AstraZeneca, Tesaro, Immunogen, Pfizer, Lilly and Clovis.
PAF has received grants from Novartis, BioNtech and Cepheid as well as personal fees from Novartis, Roche, Pfizer, Celgene, Daiichi-Sankyo, TEVA, Astra Zeneca, Merck Sharp & Dohme, Myelo Therapeutics, MacroGenics, Eisai and Puma during the conduct of the study. JDB has acted on Advisory Boards for AstraZeneca and has received support from GSK to attend conferences. His institution receives funding from AstraZeneca and Aprea. UM has shares in Abcodia Ltd. Sandra Orsulic and Beth Y. Karlan have patents on predictive gene signatures in ovarian cancer (US010253368 and EU2908913). All remaining authors have declared no conflicts of interest.

REFERENCES

APPENDIX 1. AOCS STUDY GROUP

Management Group
D Bowtell1,3,4,5,6, G Chenevix-Trench7, A Green8, P Webb2, A DeFazio8,9, D Gertz10

Project and Data Managers
N Traficante1, S Fereday1, S Moore2, J Hung7, K Harrap2, T Sadkowsky2, N Pandeya2

Research Nurses and Assistants
M Malt2, A Mellon11, R Robertson11, T Vanden Bergh12, M Jones12, P Mackenzie12, J Maidens13, K Nattrass14, YE Chiew1, A Stenlake9, H Sullivan9, B Alexander2, P Ashover2, S Brown7, T Corriss7, L Green7, L Jackman7, K Ferguson7, K Martin7, A Martyn7, B Ranieri7, J White15, V Jayde16, P Mamers17, L Bowes1, L Galletta1, D Giles1, J Hendley1, K Alsop1, T Schmidt18, H Shirley18, C Ball19, C Young19, S Viduka18, Hoa Tran18, Sanela Bilic18, Lydia Glavinas18, Julia Brooks20

Clinical and Scientific Collaborators

1Peter MacCallum Cancer Centre, Melbourne, Australia; 2QIMR Berghofer Medical Research Institute, Brisbane, Australia; 3Department of Pathology, University of Melbourne, Parkville, Australia; 4Sir Peter MacCallum Cancer Centre Department of Oncology, University of Melbourne, Parkville, Australia; 5Department of Biochemistry and Molecular Biology, University of Melbourne, Parkville, Australia; 6Ovarian Cancer Action Research Centre, Department of Surgery and Cancer, Imperial College London, London, UK; 7Centre for Cancer Research, The Westmead Institute for Medical Research, Sydney, Australia; 8The University of Sydney, Sydney, Australia; 9Department of Gynaecological Oncology, Westmead Hospital, Sydney, Australia; 10Melbourne School of Population and Global Health, University of Melbourne, Parkville, Australia; 11John Hunter Hospital, Lookout Road, New Lambton, Australia; 12Royal Hospital for Women, Barker Street, Randwick, Australia; 13Royal North Shore Hospital, St Leonards, Australia; 14Royal Prince Alfred Hospital, Camperdown, Australia; 15Royal Adelaide Hospital, North Terrace, Adelaide, Australia; 16Royal Hobart Hospital, Hobart, Australia; 17Monash Medical Centre, Clayton, Australia; 18Western Australian Research Tissue Network (WARTN), St John of God Pathology, Osborne Park, Australia; 19Women and Infant’s Research Foundation, King Edward Memorial Hospital, Subiaco, Australia; 20St John of God Hospital, Subiaco, Australia; 21Canberra Hospital, Yamba Drive, Garran, Australia; 22Bankstown Cancer Centre, Bankstown Hospital, Bankstown, Australia; 23Northern Haematology & Oncology Group, Integrated Cancer Centre, Wahroonga, Australia; 24Illawarra Shoalhaven Local Health District, Wollongong Hospital, Wollongong, Australia; 25Nepean Hospital, Kingswood, Australia; 26Newcastle Mater Misericordiae Hospital, Waratah, Australia; 27Port Macquarie Base Hospital, Port Macquarie, Australia; 28Prince of Wales Clinical School, University of New South Wales, Australia; 29St George Hospital, Kogarah, Australia; 30St Vincent’s Hospital, Darlinghurst, Australia; 31Wagga Wagga Base Hospital, Wagga Wagga, Australia; 32Crow Princess Mary Cancer Centre, Westmead Hospital, Westmead, Sydney, Australia; 33Department of Pathology, Westmead Clinical School, Westmead Hospital, The University of Sydney, Australia; 34Mater Misericordiae Hospital, Raymond Terrace, South Brisbane, Australia; 35The Royal Brisbane and Women’s Hospital, Butterfield Street, Herston, Australia; 36Wesley Hospital, Auchenflower, Australia; 37Burnside Hospital, Toorak Gardens, Australia; 38Flinders Medical Centre, Bedford Park, Australia; 39Queen Elizabeth Hospital, Woodville South, Australia; 40Freemasons Hospital, East Melbourne, Australia; 41Mercy Hospital for Women, Heidelberg, Australia; 42The Royal Women’s Hospital, Parkville, Australia; 43Border Medical Oncology, Wodonga, Australia; 44Andrew Love Cancer Centre, Geelong, Australia; 45Ballarat Base Hospital, Ballarat, Australia; 46Bendigo Health Care Group, Bendigo, Australia; 47Peninsula Health, Frankston, Australia; 48Mount Hospital, Perth, Australia; 49Faculty of Medicine, Bar-Ilan University, Safed, Israel

The seven people in bold are named authors on the manuscript.

aDeceased.