Original article

Prognostic gene expression signature for high-grade serous ovarian cancer

J. Millstein[1], T. Budden[2, 3], E. L. Goode[4], M. S Anglesio[5-7], A. Talhouk[5, 7], M. P. Intermaggio[2], H. S. Leong[8], S. Chen[9], W. Elatre[10], B. Gilks[5, 6], T. Nazeran[5], M. Volchek[11], R. C. Bentley[12], C. Wang[13], D. S. Chiu[5], S. Kommoss[14], S. C. Y. Leung[5], J. Senz[5, 6], A. Lum[5], V. Chow[5], H. Sudderuddin[5], R. Mackenzie[5], J. George[15], AOCS Group[8, 16, 17], S. Fereday[8, 18], J. Hendley[8, 18], N. Traficante[8, 18], H. Steed[19], J. M. Koziak[20], M. Köbel[21], I. A. McNeish[22, 23], T. Goranova[24], D. Ennis[22, 23], G. Macintyre[24], D. Silva[24], T. Ramón y Cajal[25], J. García-Donas[26], S. Hernando Polo[27], G. C. Rodriguez[28], K. L. Cushing-Haugen[29], H. R. Harris[29, 30], C. S. Greene[31], R. A. Zelaya[32], S. Behrens[33], R. T. Fortner[33], P. Sinn[34], E. Herpel[35], J. Lester[36, 37], J. Lubiński[38], O. Oszurek[38], A. Tołoczko[38], C. Cybulski[38], J. Menkiszak[39], C. L. Pearce[40, 41], M. C. Pike[41, 42], C. Tseng[43], J. Alsop[44], V. Rhenius[44], H. Song[44], M. Jimenez-Linan[45], A. Piskorz[24], A. Gentry-Maharaj[46], C. Karpinskyj[46], M. Widschwendter[47], N. Singh[48], C. J. Kennedy[17, 49], R. Sharma[50, 51], P. R. Harnett[17, 52], B. Gao[17, 52], S. E. Johnatty[16], R. Sayer[49], J. Boros[17, 49], S. J. Winham[13], G. L. Keeney[53], S. H. Kaufmann[54, 55], M. C. Larson[13], H. Luk[56], B. Y. Hernandez[56], P. J. Thompson[57], L. R. Wilkens[56], M. E. Carney[58], B. Trabert[59], J. Lissowska[60], L. Brinton[59], M. E. Sherman[61], C. Bodelon[59], S. Hinsley[62], L. A. Lewsley[62], R. Glasspool[63], S. N. Banerjee[64], E. A. Stronach[22], P. Haluska[54], I. Ray-Coquard[65], S. Mahner[66], B. Winterhoff[67], D. Slamon[68], D. A. Levine[69, 70], L. E. Kelemen[71], J. Benitez[72, 73], J. Chang-Claude[33, 74], J. Gronwald[38], A. H. Wu[43], U. Menon[46], M. T. Goodman[57], J. M. Schildkraut[75], N. Wentzensen[59], R. Brown[76], A. Berchuck[77], G. Chenevix-Trench[16], A. deFazio[17, 49], S. A. Gayther[78], M. J. García[73, 79], M. Henderson[80], M. A. Rossing[29, 30], A. Beeghly-Fadiel[81], P. A. Fasching[68, 82], S. Orsulic[37], B. Y. Karlan[36, 37], G. E. Konecny[68], D. G. Huntsman[5-7, 83], D. D. Bowtell[8, 18], J. D. Brenton[24], J. A. Doherty[84], P. D. P. Pharoah[44, 85], S. J. Ramus[2, 86, 87] *

1. Division of Biostatistics, Department of Preventive Medicine, Keck School of Medicine. University of Southern California. Los Angeles: USA.

2. School of Women's and Children's Health, Faculty of Medicine. University of NSW Sydney. Sydney, New South Wales: Australia; 2052.

3. CRUK Manchester Institute. The University of Manchester. Manchester: UK.

4. Department of Health Science Research, Division of Epidemiology. Mayo Clinic. Rochester, MN: USA; 55905.

5. British Columbia's Ovarian Cancer Research (OVCARE) Program. BC Cancer, Vancouver General Hospital, and University of British Columbia. Vancouver, BC: Canada; V5Z 4E6.

6. Department of Pathology and Laboratory Medicine. University of British Columbia. Vancouver, BC: Canada; V5Z 4E6.

7. Department of Obstetrics and Gynecology. University of British Columbia. Vancouver, BC: Canada.

8. Peter MacCallum Cancer Center. Melbourne, Victoria: Australia; 3000.

9. Center for Cancer Prevention and Translational Genomics, Samuel Oschin Comprehensive Cancer Institute. Cedars-Sinai Medical Center. Los Angeles, CA: USA; 90048.

10. Department of Pathology, Norris Comprehensive Cancer Center, Keck School of Medicine. University of Southern California. Los Angeles: USA.

11. Anatomical Pathology. Royal Women's Hospital. Parkville, Victoria: Australia.

12. Department of Pathology. Duke University Hospital. Durham, NC: USA; 27710.

13. Department of Health Science Research, Division of Biomedical Statistics and Informatics. Mayo Clinic. Rochester, MN: USA; 55905.

14. Department of Women's Health. Tuebingen University Hospital. Tuebingen: Germany; 72076.

15. The Jackson Laboratory for Genomic Medicine. Farmington, CT: USA; CT-06032.

16. Department of Genetics and Computational Biology. QIMR Berghofer Medical Research Institute. Brisbane, Queensland: Australia; 4006.

17. Centre for Cancer Research, The Westmead Institute for Medical Research. The University of Sydney. Sydney, New South Wales: Australia; 2145.

18. Sir Peter MacCallum Department of Oncology. The University of Melbourne. Parkville, Victoria: Australia; 3000.

19. Department of Obstetrics and Gynecology, Division of Gynecologic Oncology. Royal Alexandra Hospital. Edmonton, Alberta: Canada; T5H 3V9.

20. Alberta Health Services-Cancer Care. Calgary, AB: Canada.

21. Department of Pathology and Laboratory Medicine. University of Calgary, Foothills Medical Center. Calgary, AB: Canada; T2N 2T9.

22. Division of Cancer and Ovarian Cancer Action Research Centre, Department Surgery & Cancer. Imperial College London. London: UK; W12 0NN.

23. Institute of Cancer Sciences. University of Glasgow. Glasgow: UK; G61 1QH.

24. Cancer Research UK Cambridge Institute. University of Cambridge. Cambridge: UK; CB2 0RE.

25. Medical Oncology Service. Hospital Sant Pau. Barcelona: Spain.

26. HM Hospitales Ð Centro Integral Oncol—gico HM Clara Campal. Madrid: Spain.

27. Medical Oncology Service. Hospital Universitario Funcacion Alcorcon. Alcorc—n: Spain.

28. Division of Gynecologic Oncology. NorthShore University HealthSystem, University of Chicago. Evanston, IL: USA; 60201.

29. Program in Epidemiology, Division of Public Health Sciences. Fred Hutchinson Cancer Research Center. Seattle, WA: USA; 98109.

30. Department of Epidemiology. University of Washington. Seattle, WA: USA; 98195.

31. Department of Systems Pharmacology and Translational Therapeutics, Perelman School of Medicine. University of Pennsylvania. Philadelphia: USA; PA 19103.

32. Department of Genetics. Geisel School of Medicine at Dartmouth. Hanover, NH: USA.

33. Division of Cancer Epidemiology. German Cancer Research Center (DKFZ). Heidelberg: Germany; 69120.

34. Department of Pathology, Institute of Pathology. University Hospital Heidelberg. Heidelberg: Germany; 69120.

35. National Center for Tumor Diseases. University Hospital and German Cancer Research Center. Heidelberg: Germany; 69120.

36. David Geffen School of Medicine, Department of Obstetrics and Gynecology. University of California at Los Angeles. Los Angeles, CA: USA; 90095.

37. Women's Cancer Program at the Samuel Oschin Comprehensive Cancer Institute. Cedars-Sinai Medical Center. Los Angeles, CA: USA; 90048.

38. Department of Genetics and Pathology. Pomeranian Medical University. Szczecin: Poland; 71-252.

39. Department of Gynecological Surgery and Gynecological Oncology of Adults and Adolescents. Pomeranian Medical University. Szczecin: Poland; 70-111.

40. Department of Epidemiology. University of Michigan School of Public Health. Ann Arbor, MI: USA; 48109.

41. Department of Preventive Medicine, Keck School of Medicine. University of Southern California Norris Comprehensive Cancer Center. Los Angeles, CA: USA; 90033.

42. Department of Epidemiology and Biostatistics. Memorial Sloan-Kettering Cancer Center. New York, NY: USA; 10065.

43. Department of Preventive Medicine, Keck School of Medicine. University of Southern California. Los Angeles, CA: USA; 90033.

44. Centre for Cancer Genetic Epidemiology, Department of Oncology. University of Cambridge. Cambridge: UK; CB1 8RN.

45. Department of Pathology. Addenbrooke's Hospital NHS Trust. Cambridge: UK.

46. MRC Clinical Trials Unit at UCL, Institute of Clinical Trials & Methodology. University College London. London: UK; WC1V 6LJ.

47. Department of Women's Cancer, Institute for Women's Health. University College London. London: UK; W1T 7DN.

48. Department of Pathology. Barts Health National Health Service Trust. London: UK; E1 8PR.

49. Department of Gynaecological Oncology. Westmead Hospital. Sydney, New South Wales: Australia; 2145.

50. Pathology West ICPMR Westmead, Westmead Hospital. The University of Sydney. Sydney, New South Wales: Australia; 2145.

51. University of Western Sydney at Westmead Hospital. Sydney, New South Wales: Australia; 2145.

52. The Crown Princess Mary Cancer Centre Westmead, Sydney-West Cancer Network. Westmead Hospital. Sydney, New South Wales: Australia; 2145.

53. Department of Laboratory Medicine and Pathology, Division of Anatomic Pathology. Mayo Clinic. Rochester, MN: USA; 55905.

54. Department of Oncology. Mayo Clinic. Rochester, MN: USA; 55905.

55. Department of Molecular Pharmacology and Experimental Therapeutics. Mayo Clinic. Rochester, MN: USA; 55905.

56. Cancer Epidemiology Program. University of Hawaii Cancer Center. Honolulu, HI: USA; 96813.

57. Samuel Oschin Comprehensive Cancer Institute, Cancer Prevention and Genetics Program. Cedars-Sinai Medical Center. Los Angeles, CA: USA; 90048.

58. John A. Burns School of Medicine, Department of Obstetrics and Gynecology. University of Hawaii. Honolulu, HI: USA.

59. Division of Cancer Epidemiology and Genetics. National Cancer Institute. Bethesda, MD: USA; 20892.

60. Department of Cancer Epidemiology and Prevention. M Sklodowska-Curie Cancer Center, Oncology Institute. Warsaw: Poland; 02-034.

61. Department of Health Sciences Research. Mayo Clinic College of Medicine. Jacksonville, FL: USA; 32224.

62. Cancer Research UK Clinical Trials Unit, Institute of Cancer Sciences. University of Glasgow. Glasgow: UK; G12 0YN.

63. Department of Medical Oncology. Beatson West of Scotland Cancer Centre and University of Glasgow. Glasgow: UK; G12 0YN.

64. Gynaecology Unit. Royal Marsden Hospital. London: UK; SW3 6JJ.

65. Centre Leon Berard and University Claude Bernard Lyon 1. Lyon: France; 69373.

66. Department of Gynecology and Obstetrics. Ludwig Maximilian University of Munich. Munich: Germany; 80336.

67. Department of Obstetrics, Gynecology and Women's Health. University of Minnesota. Minneapolis, MN: USA; 55455.

68. David Geffen School of Medicine, Department of Medicine Division of Hematology and Oncology. University of California at Los Angeles. Los Angeles, CA: USA; 90095.

69. Gynecology Service, Department of Surgery. Memorial Sloan Kettering Cancer Center. New York, NY: USA; 10065.

70. Gynecologic Oncology, Laura and Isaac Pearlmutter Cancer Center. NYU Langone Medical Center. New York, NY: USA; 10016.

71. Hollings Cancer Center and Department of Public Health Sciences. Medical University of South Carolina. Charleston, SC: USA; 29425.

72. Centro de Investigaci—n en Red de Enfermedades Raras (CIBERER). Madrid: Spain; 28029.

73. Human Cancer Genetics Programme. Spanish National Cancer Research Centre (CNIO). Madrid: Spain; 28029.

74. Cancer Epidemiology Group, University Cancer Center Hamburg (UCCH). University Medical Center Hamburg-Eppendorf. Hamburg: Germany; 20246.

75. Department of Public Health Sciences. University of Virginia. Charlottesville, VA: USA; 22908.

76. Division of Cancer and Ovarian Cancer Action Research Centre, Department of Surgery and Cancer. Imperial College London. London: UK; W12 0NN.

77. Department of Gynecologic Oncology. Duke University Hospital. Durham, NC: USA; 27710.

78. Center for Bioinformatics and Functional Genomics and the Cedars Sinai Genomics Core. Cedars-Sinai Medical Center. Los Angeles, CA: USA; 90048.

79. Biomedical Network on Rare Diseases (CIBERER). Madrid: Spain; 28029.

80. Children's Cancer Institute, Lowy Cancer Research Centre. University of NSW Sydney. Sydney, New South Wales: Australia; 2052.

81. Division of Epidemiology, Department of Medicine, Vanderbilt Epidemiology Center, Vanderbilt-Ingram Cancer Center. Vanderbilt University School of Medicine. Nashville, TN: USA; 37232.

82. Department of Gynecology and Obstetrics, Comprehensive Cancer Center ER-EMN. University Hospital Erlangen, Friedrich-Alexander-University Erlangen-Nuremberg. Erlangen: Germany; 91054.

83. Department of Molecular Oncology. BC Cancer Research Centre. Vancouver, BC: Canada; V5Z 4E6.

84. Huntsman Cancer Institute, Department of Population Health Sciences. University of Utah. Salt Lake City, UT: USA; 84112.

85. Centre for Cancer Genetic Epidemiology, Department of Public Health and Primary Care. University of Cambridge. Cambridge: UK; CB1 8RN.

86. Adult Cancer Program, Lowy Cancer Research Centre. University of NSW Sydney. Sydney, New South Wales: Australia; 2052.

* **Corresponding author:**

Prof Susan Ramus

School of Women's and Children's Health, Faculty of Medicine

University of NSW Sydney, Sydney, New South Wales, Australia

(02) 9385 1720, s.ramus@unsw.edu.au

Running head: prognostic signature for high grade serous ovarian cancer

Abstract

Background

Median overall survival (OS) for women with high-grade serous ovarian cancer (HGSOC) is approximately four 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 overall survival in HGSOC patients.

Patients and methods

Expression of 513 genes, selected from a meta-analysis of 1455 tumours and other candidates, were measured using NanoString technology from formalin-fixed, paraffin-embedded (FFPE) tumour tissue from 3,769 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 fifteen 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 (FDR) < 0.05] in covariate-adjusted single gene analyses. The top five genes were *TAP1, ZFHX4, CXCL9, FBN1,* and *PTGER3* ($P \ll 0.001$). 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 (GES) conferred a greater than two-fold increase in risk of death [HR = 2.35 (2.02, 2.71); $P \ll 0.001$]. Median survival by GES quintile was 9.5 (8.3, --), 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: high grade serous ovarian cancer, gene expression, prognosis, overall survival, formalin fixed paraffin embedded

Highlights

- A gene expression signature for high-grade serous ovarian cancer prognostic for two and five-year overall survival (OS).
- The 101 gene expression signature performs substantially better than age and stage alone.
- Median survival by quintile was 9.5, 5.4, 3.8, 3.2 and 2.3 years.
- The top five genes associated with OS were *TAP1, ZFHX4, CXCL9, FBN1,* and *PTGER3* ($P \ll 0.001$).

37 ## Introduction

38 Epithelial ovarian cancer (EOC) causes approximately 125,000 deaths globally every year, and long-term
39 survival rates have changed little in the past three decades[88]. Approximately 70% of women with EOC are
40 diagnosed with advanced stage disease (stages III/IV), and fewer than 50% will survive more than 5 years[89].
41 There are five major EOC histotypes – high-grade serous; low-grade serous; endometrioid; clear cell and
42 mucinous[90]. High-grade serous ovarian cancer (HGSOC) comprises about two-thirds of cases, is responsible
43 for most deaths and is characterized by profound genomic and clinical heterogeneity.

44 The most informative prognostic factors for HGSOC are International Federation of Gynecology and Obstetrics
45 (FIGO) stage, residual disease following debulking surgery[91], *BRCA1* or *BRCA2* germline mutation[92, 93] and
46 tumour-infiltrating lymphocyte scores[94, 95]. Patients with HGSOC who carry a loss-of-function germline
47 mutation in *BRCA1* or *BRCA2* have an increased sensitivity to platinum-based chemotherapy and PARP
48 inhibitor treatment[96, 97] and a medium-term survival advantage[92]. However, the frequent development
49 of drug resistant disease[93] limits the effectiveness of current therapies.

50 Gene-expression data have been used to define four tumour molecular subtypes of HGSOC (C1/mesenchymal,
51 C2/immune, C4/differentiated and C5/proliferative)[98, 99]. Using transcriptome-wide data from fresh frozen
52 tissues, The Cancer Genome Atlas (TCGA) project used 215 tumours to identify an overall survival (OS)
53 expression signature of 193 genes that has been validated on three other HGSOC gene expression data
54 sets[99].

55 Despite these findings, gene expression biomarkers have not been implemented clinically owing to several
56 important shortcomings. The majority of the individual markers comprising the 193 gene signature were not
57 statistically significant across all studies, suggesting that the signature may not be robust. The sample sizes in
58 other discovery efforts have been too small for robust statistical inference [99]. Also, previous studies used

59 fresh frozen samples, resulting in logistic and cost barriers to examining large clinically relevant data sets, and
60 translation to the clinical setting.

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

63 Patients and methods

64 Twenty studies provided pre-treatment, FFPE tumour samples from 4,071 women diagnosed with HGSOC
65 (Supplemental Table S1). All HGSOC cases with available tissue were included. During this time period HGSOC
66 patients were treated with chemotherapy (carboplatin and paclitaxel) after primary debulking surgery. Study
67 protocols were approved by the respective Institutional Review Board / ethics approval committee for each
68 site (Supplemental Table S1).

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

72 Gene selection

73 Candidate prognostic genes were identified by carrying out an individual participant meta-analysis of six
74 transcriptome-wide microarray studies[98-103], which included tumour samples from 1,455 participants.
75 Gene expression association with overall survival was evaluated by Cox proportional hazards regression
76 adjusted for molecular subtype (Supplemental Table S2). In total, 200 genes from the meta-analysis, most
77 achieving a permutation-based FDR[104] of less than 0.05, and an additional 313 candidate genes based on the
78 literature and unpublished data were selected (Supplemental Tables S3 and S4, Figure S1; for more details see
79 Supplemental Material). Five genes, RPL19, ACTB, PGK1, SDHA, and POLR1B, were included as house-keeping
80 genes for normalization.

81 **Gene expression assay in study participants samples**

82 FFPE tumour samples were processed with the NanoString nCounter technology at 3 different locations,
83 Vancouver, Los Angeles and Melbourne. A control set of 48 FFPE tumour samples were run at each location
84 and the average intraclass correlation coefficient (ICC) was 0.987. Approximately 2 percent of the samples
85 were run in duplicate and the average Spearman correlation r^2 was 0.995. Single-patient classification methods
86 were used with reference samples to control for batch effects[105]. The data in this publication have been
87 deposited in NCBI's Gene Expression Omnibus[106]; GEO Series accession number GSE132342
88 (https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE132342). Three thousand eight hundred and
89 twenty-nine samples passed quality control of which 3,769 had survival data and assessable gene expression
90 for 513 genes. Data can be found in NCBI GEO: Accession numbers GSE132342 and GPL26748.

91 **Overall survival analysis of individual genes**

92 Samples that contributed to the meta-analysis data set (n=211) were removed from subsequent selected
93 analyses to enforce independence of study samples between the gene selection and final survival analysis.
94 Time-to-event analyses were carried out for OS with right-censoring at 10 years and left-truncation of
95 prevalent cases. Associations between log-transformed normalized gene expression and survival time were
96 tested using likelihood ratio tests with Cox proportional hazards models adjusted for age, race, and stage, and
97 stratified by study. Patients with missing race or stage information were assigned to 'unknown' categories. Age
98 was modeled using a B-spline with a knot at the median age, which yielded a better fit than using knots at
99 quartiles or categorical variables. Stage was dichotomized into early (International Federation of Gynecology
100 and Obstetrics [FIGO] stage I/II) and advanced (FIGO stage III/IV). Genes were scaled to have a standard
101 deviation of one, so hazard ratios correspond to a change of one standard deviation. A Benjamini-Hochberg
102 (BH) false discovery rate (FDR) of less than 0.05 was used to identify notable associations. Since the expression
103 of genes can be correlated, an analysis of correlated genes was performed using data from TCGA. Advanced

104 stage ovarian cancer usually has disease spread throughout the abdomen, therefore sensitivity analyses were
105 performed to assess effects of the anatomical location of tumor samples included in the study by removing
106 observations corresponding to samples known to be extraovarian (n = 437).

107 **Prognostic signature development and validation**

108 Studies were initially randomized to training set (*N* = 14) and validation set (*N* = 6). The TRI study was
109 randomized to the validation set, but, because 107 of the samples were part of the meta-analysis data used for
110 gene selection, the study was split, so those 107 samples were included in the model training data set. Thus
111 2,702 samples from 15 studies were used for model training and 1,067 samples from 6 studies were used for
112 validation (Supplemental Table S1). In the training set, four modelling approaches (stepwise regression, elastic
113 net regularized regression, boosting and random survival forests) were applied to construct competing gene
114 expression-based biomarkers. Each was evaluated in the training data using 10-fold cross-validation for its
115 prognostic value for OS at two and five years of follow-up using an area under the curve (AUC) measure
116 derived from receiver operator characteristic (ROC) analysis (see Supplemental Material for additional details).
117 The best performing method, elastic net regularized regression, was applied to the full training set to
118 determine the final gene signature and scoring method, which was then evaluated using the independent
119 testing set. All models were constrained to include age and stage, where age was modelled as categorical
120 based on quartiles of the training dataset with groups: less than 53 years old, 53 to 59, 60 to 66, and 67 or
121 greater. Stage was modelled as described above for the OS individual gene analysis.

122 ## Results

123 **Association of expression of individual genes with OS in HGSOC.**

124 In a gene-by-gene analysis of the full data set adjusted for age, race, and stage, and stratified by study, 276 of
125 the 513 selected genes were associated with OS (FDR < 0.05). Of these, 138 were selected from the meta-

126 analysis of six published microarray studies (Supplemental Table S2)[98-103] and 144 from candidate gene
127 approaches (Supplemental Tables S5 and S6). Hazard ratios (HR) for one standard deviation change in gene
128 expression ranged from 0.84–1.19, with multiple genes exhibiting associations at very stringent significance
129 levels (e.g., 19 genes with $P < 1 \times 10^{-8}$; Supplemental Tables S5 and S6). The five most significant genes were
130 *TAP1*, *ZFHX4*, *CXCL9*, *FBN1* and *PTGER3* (Table 1). We did not find extensive evidence of high co-expression
131 between these five genes and genes measured in TCGA project (Supplemental Table S7). In sensitivity analyses
132 we found that excluding samples from omentum and other extra-ovarian sites did not substantially affect the
133 results (Supplemental Tables S8 and S9).

134 **Development of a novel prognostic gene signature**

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

143 Performance of the signature including age and stage was AUC = 0.69 (95% CI 0.65-0.73) and 0.75 (95% CI 0.72-
144 0.78) for 2-yr and 5-yr OS, respectively (Figure 2, Figure 3, Supplemental Figure S4). This was substantially
145 better than age and stage alone with AUC = 0.61 (95% CI 0.57-0.65) and 0.62 (95% CI 0.59- 0.67) for 2-yr and 5-
146 yr OS, respectively), particularly for the 5-yr OS outcome with non-overlapping 95% CI. One standard deviation
147 change in the gene expression score was associated with a hazard ratio of 2.35 [95% CI = (2.02, 2.71); P =
148 5.1×10^{-31}], and median survival varied substantially across quintiles of the gene expression score [9.5 (8.3, ---),

149 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;
150 Table 2].

151 For a subset of cases, there was clinical and experimental data for known prognostic factors. All samples had
152 molecular subtype classification (Talhouk et al. submitted), residual disease was known for 1,771 cases,
153 primary treatment for 687, germline *BRCA* mutation status for 904, and nuclear CD8 TIL counts[95] for 1,111
154 (Supplemental Table S11). When examined by quintile of gene expression score there were differences, as
155 expected, for each of the known prognostic factors, including age and stage that were included in the model
156 (Table 3). However, in sensitivity analyses, applying the signature to specific patient groups, a robustness of
157 stratification was demonstrated, suggesting that the prognostic power of the signature is not explained by the
158 individual factors, residual disease, treatment, *BRCA* status, or CD8 score (Figure 3, Supplemental Figures S5-
159 S7). The signature score showed modest differences by molecular subtype (Supplemental Figure S8), and
160 adjusting for molecular subtype in the Cox analysis resulted in only minor changes to the HR estimates for
161 signature quintiles (Table 2). The signature was shown to be prognostic within a homogenous group of 316
162 stage 3C cases with no residual disease, within early stage cases (FIGO 1a and 1b), and within patients whose
163 samples were collected from the omentum (Supplemental Figures S9-S10). Analysis of the signature score for
164 paired ovary and omental tissue from 42 of the cases showed a highly significant Pearson correlation
165 coefficient, $r = 0.79$ ($p = 5.4 \times 10^{-10}$) (Supplemental Figure S11).

166 A geneset enrichment analysis was performed for the 101 genes in the signature, as well as for genes
167 correlated with signature genes achieving $r2 > 0.75$ (Supplemental Table S12). For the correlated gene analysis,
168 the three most significant pathways involved the immune system, including the adaptive immune system and
169 cytokine signalling. A further ten immune pathways were significantly enriched and included interferon
170 signalling, innate immune system, and TCR signalling and antigen presentation pathways. Restricting to the
171 signature genes only, there was also enrichment in the immune system, but the top two pathways were PI-3K
172 cascade and GPCR ligand binding. Four other pathways were related to the cell cycle and mitosis, with the

173 remaining enriched for fibroblast growth factor receptor (FGFR) and epidermal growth factor receptor (ERRB)
174 signalling, and one pathway related to homologous combination repair.

175 Discussion

176 In a large-scale study of HGSOC patients, we identified a 101 gene expression signature able to predict
177 clinically relevant differences in OS. Using methods that are both economical and applicable to standard
178 clinical sampling techniques, we showed that the signature performs substantially better than age and stage
179 alone for prognosis of both two and five-year OS. The number of patients and samples included in this study is
180 an order of magnitude greater than previous comparable studies of gene expression and OS in HGSOC
181 patients[99, 107, 108]. Thus, we have been able to more precisely quantify the prognostic value of gene
182 expression.

183 We report definitive associations between OS and expression of 276 genes. Of the five most significant genes
184 (*TAP1*, *ZFHX4*, *CXCL9*, *FBN1*, and *PTGER3*), four have been previously reported to be associated with survival in
185 HGSOC. The top prognostic gene, *TAP1*, is involved in the antigen presenting pathway. Expression was reduced
186 in metastatic HGSOC, positively associated with OS[109] as observed here, and linked to tumour regression in
187 response to treatment[110]. Also, hypomethylation of *TAP1* was associated with improved time to disease
188 recurrence[111]. *CXCL9* is a chemokine that mediates the recruitment of T-cells to solid tumours[112]. High
189 expression of intratumoural *CXCL9* was associated with higher OS[113] and higher lymphocytic infiltration,
190 which is also a robust prognostic factor in HGSOC[95, 98, 114] and a feature of the immunoreactive HGSOC
191 molecular subtype[98]. *CXCL9* has also been proposed as a therapeutic target due to evidence that it inhibits
192 angiogenesis and promotes antitumour adaptive immunity[115-117]. Strikingly, the signature was able to
193 further refine prognostic groups within patients with high TIL counts suggesting that *CXCL9* and *TAP1*
194 expression may be strong indicators of immune competency in HGSOC.

195 *FBN1* is an extracellular matrix (ECM) protein previously found to be a biomarker associated with early
196 recurrence in ovarian cancer patients who are initially sensitive to chemotherapy[118] and strongly correlated
197 with desmoplasia in HGSOC. The prostaglandin E2 receptor *PTGER3* is expressed in ovarian tumour cells and is
198 associated with relapse-free survival[119]. In contrast, *ZFHX4* does not have previous associations with HGSOC.

199 Associations between the expression of specific genes in tumour tissues and OS in HGSOC patients may
200 suggest new drug targets and lead to insights into biological variation in treatment response. For example,
201 cases in the Q5 quintile with the poorest outcome had increased expression of *IGF2*, *FGFR1*, and *MYC*, a
202 possible argument for the use of *IGFR1*, *FGFR*, Bromodomain (*MYC*), or a combination of PARP and CDK4/6
203 inhibitors (*MYC*) [33]. More immediately, the signature may help clinicians identify patients most in need of
204 intervention, patients that could potentially benefit from neo-adjuvant chemotherapy (NACT). Alternatively, in
205 clinical trials it could be used to stratify randomization by patients' risk, thereby reducing heterogeneity within
206 subgroups and increasing heterogeneity between subgroups. The signature will be incorporated into future
207 prospective clinical trials to determine if it can predict response to specific treatments.

208 Measurement of the signature required standard FFPE tissue used in routine histopathology. Also, data
209 preprocessing and normalization were conducted on an individual level, thus translatable to a general patient
210 population. That is, 5-year OS prognosis of future patients can be evaluated against the patient population
211 reported here by i) following the same steps described here for generating the normalized gene expression
212 data, 2) computing an individual signature score, and 3) assigning an HR based on the score or comparing it to
213 the reported quintiles (Supplemental Material). NanoString gene expression is highly reproducible as seen by
214 our quality control metrics (Supplemental Material) and the FDA approval of the ProSigna test for breast
215 cancer.

216 The question of heterogeneity by ancestry or ethnicity was beyond the scope of this study but should be
217 pursued in future research. Another important question is whether molecular subtype can improve biomarker

218 performance. A substantial proportion of signature genes were identified by the subtype adjusted meta-
219 analysis, suggesting that the strong performance of the signature is not solely attributable to differences
220 among molecular subtypes. Additionally, all of the individual genes used in the molecular subtype classification
221 were included in development of the signature.

222 Although the cases received chemotherapy, the FFPE samples used in this study were chemo-naïve, as few
223 patients had NACT during the calendar period in which these samples were collected. Because the signature
224 appears to be prognostic in omentum samples, future studies may assess the value in NACT patients, using
225 pre-treatment omental biopsies or post treatment tumour samples. Future work will also address if the
226 signature can predict platinum-refractory patients.

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

232

233 **Acknowledgements**

234 We thank all the study participants who contributed to this study and all the researchers, clinicians and
235 technical and administrative staff who have made possible this work. This project received technical and data
236 management support from OVCARE's core units, including the Cheryl Brown Ovarian Cancer Outcomes Unit
237 and the Genetic Pathology Evaluation Centre, and statistical analysis support from the Biostatistics Core of the
238 Norris Comprehensive Cancer Center. The AOV study recognizes the valuable contributions from Mie Konno,
239 Shuhong Liu, Michelle Darago, Faye Chambers and the staff at the Tom Baker Cancer Centre Translational

240 Laboratories. We thank Olivier Tredan and Pierre Heudel as investigators on the TRIO14 study and Sandrine
241 Berge-Montamat as assistant for clinical research. The Australian Ovarian Cancer Study gratefully
242 acknowledges additional support from Ovarian Cancer Australia and the Peter MacCallum Foundation. The
243 AOCS also acknowledges the cooperation of the participating institutions in Australia and acknowledges the
244 contribution of the study nurses, research assistants and all clinical and scientific collaborators to the study.
245 The complete AOCS Study Group can be found at www.aocstudy.org. We would like to thank all of the women
246 who participated in these research programs.

247 **FUNDING**

248 This work was funded by the National Cancer Institute (NCI) Grants R01CA172404 (to SJR) and R01CA168758
249 (to JAD and MAR), the Canadian Institutes for Health Research (Proof-of-Principle I program) and the United
250 States Department of Defense Ovarian Cancer Research Program (OC110433). M. Milstein and S.J. Ramus
251 received support from NIH/National Cancer Institute award number P30CA014089. M.S. Anglesio receives
252 funding from the Janet D. Cottrelle Foundation Scholar's program managed by the BC Cancer Foundation. J.
253 George was partially supported by the NIH/National Cancer Institute award number P30CA034196. C.Wang
254 was a Career Enhancement Awardee of the Mayo Clinic SPORE in Ovarian Cancer (P50 CA136393). D.G.
255 Huntsman receives support from the Dr. Chew Wei Memorial Professorship in Gynecologic Oncology, the
256 Canada Research Chairs program (Research Chair in Molecular and Genomic Pathology), and the Janet D.
257 Cottrelle Foundation. M. Widschwendter receives funding from the European Union's Horizon 2020 European
258 Research Council Programme, H2020 BRCA-ERC under Grant Agreement No. 742432 as well as the charity, The
259 Eve Appeal (https://eveappeal.org.uk/), and support of the National Institute for Health Research (NIHR) and
260 the University College London Hospitals (UCLH) Biomedical Research Centre. G.E. Konecny is supported by the
261 Miriam and Sheldon Adelson Medical Research Foundation. B.Y. Karlan is funded by the American Cancer
262 Society Early Detection Professorship (SIOP-06-258-01-COUN) and the National Center for Advancing
263 Translational Sciences (NCATS), Grant UL1TR000124. H.R. Harris is supported by the NIH/National Cancer

264 Institute award number K22 CA193860. OVCARE (including the VAN study) receives core funding through the
265 BC Cancer Foundation and The VGH+UBC Hospital Foundation (authors AT, BG, DGH, and MSA). The AOV study
266 is supported by the Canadian Institutes of Health Research (MOP-86727). The Gynaecological Oncology
267 Biobank at Westmead, a member of the Australasian Biospecimen Network-Oncology group, was funded by
268 the National Health and Medical Research Council Enabling Grants ID 310670 & ID 628903 and the Cancer
269 Institute NSW Grants ID 12/RIG/1-17 & 15/RIG/1-16. The Australian Ovarian Cancer Study Group was
270 supported by the U.S. Army Medical Research and Materiel Command under DAMD17-01-1-0729, The Cancer
271 Council Victoria, Queensland Cancer Fund, The Cancer Council New South Wales, The Cancer Council South
272 Australia, The Cancer Council Tasmania and The Cancer Foundation of Western Australia (Multi-State
273 Applications 191, 211 and 182) and the National Health and Medical Research Council of Australia (NHMRC;
274 ID199600; ID400413 and ID400281). BriTROC-1 was funded by Ovarian Cancer Action (to IAM and JDB, grant
275 number 006) and supported by Cancer Research UK (grant numbers A15973, A15601, A18072, A17197,
276 A19274 and A19694) and the National Institute for Health Research Cambridge and Imperial Biomedical
277 Research Centres. SEARCH was supported by Cancer Research UK (A16561). The University of Cambridge
278 receives salary support for PDPP from the NHS Clinical Academic Reserve (no grant number applicable).
279 Samples from the Mayo Clinic were collected and provided with support of P50 CA136393 (ELG, GLK, SHK,
280 MES).

281 Disclosure

282 Beth Y. Karlan served on Invitae Corporation's Advisory Board from 2017 to 2018. Iain McNeish has acted on
283 Advisory Boards for AstraZeneca, Clovis Oncology, Tesaro, Carrick Therapeutics and Takeda. His institution
284 receives funding from AstraZeneca. Ros Glasspool in on the Advisory Boards for AstraZeneca, Tesaro, Clovis
285 and Immunogen and does consultancy work for SOTIO. She has received support to attend conferences from
286 AstraZeneca, Roche and Tesaro. Her institution has received research funding from Boehringer Ingelheim and
287 Lilly/Ignyta and she is the national co-ordinating investigator for the UK for trials sponsored by AstraZeneca

288 and Tesaro and site principal investigator for trials sponsored by AstraZeneca, Tesaro, Immunogen, Pfizer, Lilly
289 and Clovis. Peter Fasching has received grants from Novartis, Biontech and Cepheid as well as personal fees
290 from Novartis, Roche, Pfizer, Celgene, Daiichi-Sankyo, TEVA, Astra Zeneca, Merck Sharp & Dohme, Myelo
291 Therapeutics, Macrogenics, Eisai and Puma during the conduct of the study. Usha Menon has shares in
292 Abcodia Ltd. All remaining authors have declared no conflicts of interest.
293

294 ## References

295 1. Vaughan S, Coward JI, Bast RC, Jr. et al. Rethinking ovarian cancer: recommendations for
296 improving outcomes. Nat Rev Cancer 2011; 11: 719-725.

297 2. Torre LA, Trabert B, DeSantis CE et al. Ovarian cancer statistics, 2018. CA Cancer J Clin 2018;
298 68: 284-296.

299 3. Bowtell DD, Bohm S, Ahmed AA et al. Rethinking ovarian cancer II: reducing mortality from
300 high-grade serous ovarian cancer. Nat Rev Cancer 2015; 15: 668-679.

301 4. du Bois A, Reuss A, Pujade-Lauraine E et al. Role of surgical outcome as prognostic factor in
302 advanced epithelial ovarian cancer: a combined exploratory analysis of 3 prospectively randomized
303 phase 3 multicenter trials: by the Arbeitsgemeinschaft Gynaekologische Onkologie Studiengruppe
304 Ovarialkarzinom (AGO-OVAR) and the Groupe d'Investigateurs Nationaux Pour les Etudes des Cancers
305 de l'Ovaire (GINECO). Cancer 2009; 115: 1234-1244.

306 5. Bolton KL, Chenevix-Trench G, Goh C et al. Association between BRCA1 and BRCA2 mutations
307 and survival in women with invasive epithelial ovarian cancer. JAMA 2012; 307: 382-390.

308 6. Candido-dos-Reis FJ, Song H, Goode EL et al. Germline mutation in BRCA1 or BRCA2 and ten-
309 year survival for women diagnosed with epithelial ovarian cancer. Clin Cancer Res 2015; 21: 652-657.

310 7. Goode EL, Block MS, Kalli KR et al. Dose-Response Association of CD8+ Tumor-Infiltrating
311 Lymphocytes and Survival Time in High-Grade Serous Ovarian Cancer. JAMA Oncol 2017; 3: e173290.

312 8. Zhang L, Conejo-Garcia JR, Katsaros D et al. Intratumoral T cells, recurrence, and survival in
313 epithelial ovarian cancer. N Engl J Med 2003; 348: 203-213.

314 9. Pujade-Lauraine E, Ledermann JA, Selle F et al. Olaparib tablets as maintenance therapy in
315 patients with platinum-sensitive, relapsed ovarian cancer and a BRCA1/2 mutation (SOLO2/ENGOT-
316 Ov21): a double-blind, randomised, placebo-controlled, phase 3 trial. Lancet Oncol 2017; 18: 1274-
317 1284.

318 10. Moore K, Colombo N, Scambia G et al. Maintenance Olaparib in Patients with Newly
319 Diagnosed Advanced Ovarian Cancer. N Engl J Med 2018; 379: 2495-2505.

320 11. Tothill RW, Tinker AV, George J et al. Novel molecular subtypes of serous and endometrioid
321 ovarian cancer linked to clinical outcome. Clin Cancer Res 2008; 14: 5198-5208.

322 12. Cancer Genome Atlas Research N. Integrated genomic analyses of ovarian carcinoma. Nature
323 2011; 474: 609-615.

324 13. Bonome T, Levine DA, Shih J et al. A gene signature predicting for survival in suboptimally
325 debulked patients with ovarian cancer. Cancer Res 2008; 68: 5478-5486.

326 14. Karlan BY, Dering J, Walsh C et al. POSTN/TGFBI-associated stromal signature predicts poor
327 prognosis in serous epithelial ovarian cancer. Gynecol Oncol 2014; 132: 334-342.

328 15. Konecny GE, Haluska P, Janicke F et al. A phase II, multicenter, randomized, double-blind,
329 placebo-controlled trial of ganitumab or placebo in combination with carboplatin/paclitaxel as front-
330 line therapy for optimally debulked primary ovarian cancer: The TRIO14 trial. Journal of Clinical
331 Oncology 2014; 32: 5529.

332 16. Konecny GE, Wang C, Hamidi H et al. Prognostic and therapeutic relevance of molecular
333 subtypes in high-grade serous ovarian cancer. J Natl Cancer Inst 2014; 106.

334 17. Millstein J, Volfson D. Computationally efficient permutation-based confidence interval
335 estimation for tail-area FDR. Front Genet 2013; 4: 179.

336 18. Talhouk A, Kommoss S, Mackenzie R et al. Single-Patient Molecular Testing with NanoString
337 nCounter Data Using a Reference-Based Strategy for Batch Effect Correction. PLoS One 2016; 11:
338 e0153844.

339 19. Edgar R, Domrachev M, Lash AE. Gene Expression Omnibus: NCBI gene expression and
340 hybridization array data repository. Nucleic Acids Res 2002; 30: 207-210.

341 20. Jin C, Xue Y, Li Y et al. A 2-Protein Signature Predicting Clinical Outcome in High-Grade Serous
342 Ovarian Cancer. Int J Gynecol Cancer 2018; 28: 51-58.

343 21. Mankoo PK, Shen R, Schultz N et al. Time to recurrence and survival in serous ovarian tumors
344 predicted from integrated genomic profiles. PLoS One 2011; 6: e24709.

345 22. Nymoen DA, Hetland Falkenthal TE, Holth A et al. Expression and clinical role of
346 chemoresponse-associated genes in ovarian serous carcinoma. Gynecol Oncol 2015; 139: 30-39.

347 23. Jimenez-Sanchez A, Memon D, Pourpe S et al. Heterogeneous Tumor-Immune
348 Microenvironments among Differentially Growing Metastases in an Ovarian Cancer Patient. Cell 2017;
349 170: 927-938.e920.

350 24. Wang C, Cicek MS, Charbonneau B et al. Tumor hypomethylation at 6p21.3 associates with
351 longer time to recurrence of high-grade serous epithelial ovarian cancer. Cancer Res 2014; 74: 3084-
352 3091.

353 25. Gorbachev AV, Kobayashi H, Kudo D et al. CXC chemokine ligand 9/monokine induced by IFN-
354 gamma production by tumor cells is critical for T cell-mediated suppression of cutaneous tumors. J
355 Immunol 2007; 178: 2278-2286.

356 26. Bronger H, Singer J, Windmuller C et al. CXCL9 and CXCL10 predict survival and are regulated
357 by cyclooxygenase inhibition in advanced serous ovarian cancer. Br J Cancer 2016; 115: 553-563.

358 27. Dose-Response Association of CD8+ Tumor-Infiltrating Lymphocytes and Survival Time in High-
359 Grade Serous Ovarian Cancer. JAMA Oncol 2017; 3: e173290.

360 28. Tokunaga R, Zhang W, Naseem M et al. CXCL9, CXCL10, CXCL11/CXCR3 axis for immune
361 activation - A target for novel cancer therapy. Cancer Treat Rev 2018; 63: 40-47.

362 29. Xiao P, Guo Y, Zhang H et al. Myeloid-restricted ablation of Shp2 restrains melanoma growth
363 by amplifying the reciprocal promotion of CXCL9 and IFN-gamma production in tumor
364 microenvironment. Oncogene 2018.

365 30. Zhang R, Tian L, Chen LJ et al. Combination of MIG (CXCL9) chemokine gene therapy with low-
366 dose cisplatin improves therapeutic efficacy against murine carcinoma. Gene Ther 2006; 13: 1263-
367 1271.

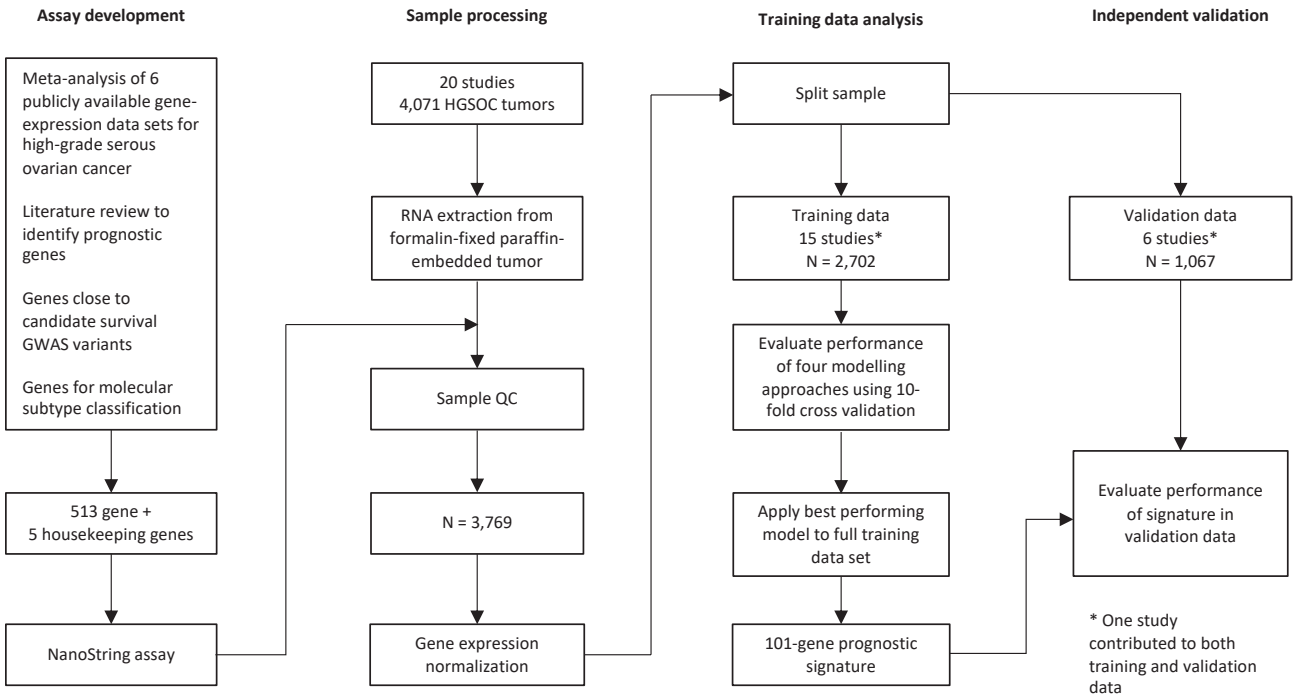
368 31. Zhang W, Ota T, Shridhar V et al. Network-based survival analysis reveals subnetwork
369 signatures for predicting outcomes of ovarian cancer treatment. PLoS Comput Biol 2013; 9:
370 e1002975.

371 32. Reinartz S, Finkernagel F, Adhikary T et al. A transcriptome-based global map of signaling
372 pathways in the ovarian cancer microenvironment associated with clinical outcome. Genome Biol
373 2016; 17: 108.

374 33. Konecny GE. Combining PARP and CDK4/6 inhibitors in MYC driven ovarian cancer.
375 EBioMedicine 2019; 43:9-10.
376

Figure Legends

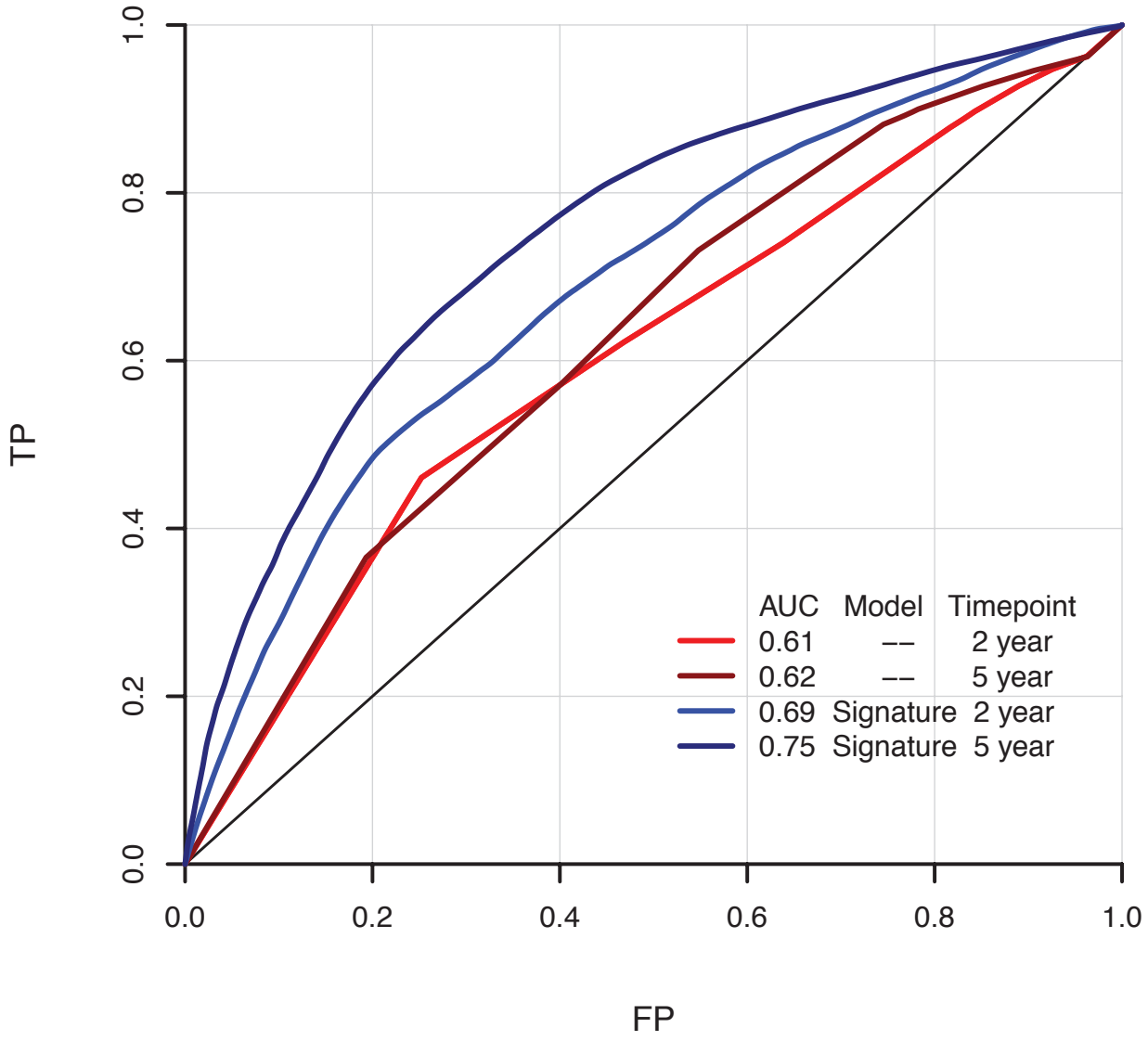
382 **Figure 1.** Schematic of study design. * The TRI study was split across the training and validation sets due to 107
383 samples overlapping with the meta-analysis.



384

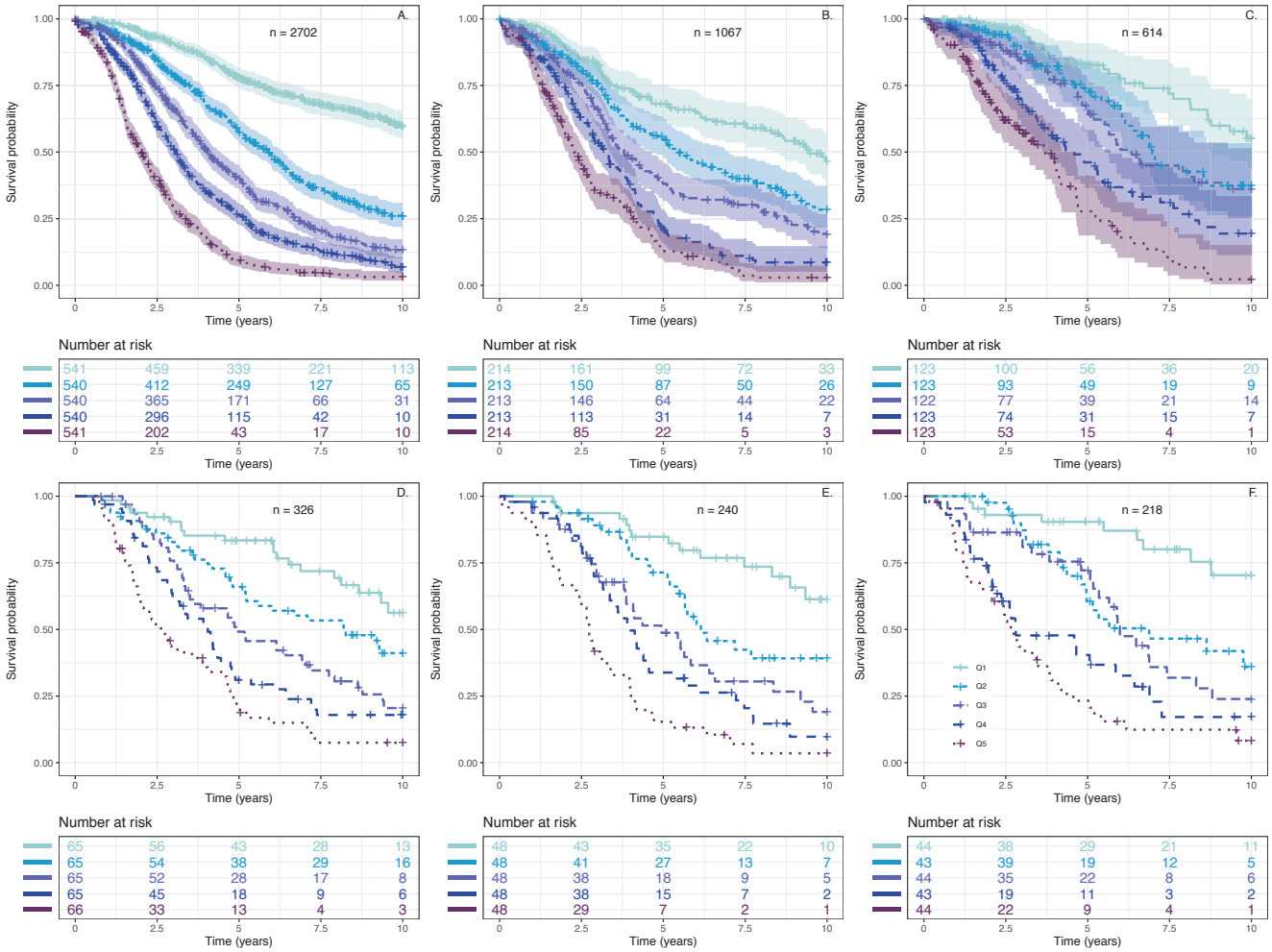
385

386 **Figure 2**. ROC curves for prognostic performance of the gene expression signature in independent HGSOC
387 patients (testing data). There was no overlap between studies or patient data used to develop models (training
388 data) and compute ROC curves and AUC values shown here (testing data). All models included age and stage as
389 described in Methods. TP denotes the true positive rate (sensitivity) and FP denotes the false positive rate (1 –
390 specificity).



391
392

Figure 3. KM curves of overall survival for patients A) in the training and B) testing sets. Patients were assigned to quintiles (Q1-Q5) of the signature score including age and stage. Shaded areas indicate 95 percent confidence regions, only included for plots representing larger sample sizes. Due to limited sample size, the following plots represent all such patients in the entire data set, training or testing, C) no macroscopic residual disease after debulking surgery, D) primary treatment ≥ 4 cycles of IV carboplatin AUC 5 or 6 & paclitaxel 135 or 175 mg/m² every 3 weeks (actual dose known or presumed), E) *BRCA1* or *BRCA2* germline mutation, and F) CD8 > 19.



402 **Table 1**. Hazard ratios and 95% CIs for top 5 prognostic genes in covariate-adjusted single gene analyses.

Gene	HR (95% CI)	*P*	Selection	Correlated gene*	r_s
TAP1	0.84 (0.80, 0.87)	8.3×10^{-18}	Meta	PSMB9	0.89
ZFHX4	1.19 (1.14, 1.25)	1.4×10^{-15}	Meta	LOC100192378	0.74
CXCL9	0.85 (0.82, 0.88)	1.8×10^{-15}	Meta and candidate	CXCR6	0.89
FBN1	1.18 (1.13, 1.24)	4.2×10^{-14}	Candidate	SPARC^	0.91
PTGER3	1.18 (1.13, 1.24)	1.2×10^{-13}	Meta	COL8A1	0.67

403 *Most correlated gene according to Spearman's rank correlation coefficient, r_s, computed in The Cancer
404 Genome Atlas (TCGA) Ovarian Serous Cystadenocarcinoma RNA-seq data set.

405 ^ *SPARC* was included in this project and was less significant.

406

407

408 **Table 2**. Hazard ratios and 95% CIs for quintiles of the gene expression signature score in validation data.

Quintile	*N*	Deaths	Median Survival*	HR (95% CI)	Adjusted for Age and Stage HR (95% CI)	Adjusted for M. Subtype Age and Stage HR (95% CI)
Q1	214	81	9.47 (8.32, ------)	0.44 (0.33, 0.58)	0.34 (0.22, 0.55)	0.37 (0.23, 0.59)
Q2	213	117	5.38 (4.63, 6.97)	0.73 (0.57, 0.93)	0.71 (0.55, 0.91)	0.74 (0.58, 0.96)
Q3	213	145	3.80 (3.34, 4.60)			
Q4	213	158	3.23 (2.85, 3.68)	1.56 (1.25, 1.96)	1.56 (1.24, 1.97)	1.56 (1.24, 1.96)
Q5	214	179	2.27 (2.09, 2.62)	2.23 (1.78, 2.78)	2.11 (1.67, 2.67)	2.07 (1.63, 2.61)

409 *Median survival (95% CI) in years for patients in the validation set.

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

	Total	Q1	Q2	Q3	Q4	Q5	p-value
N	3769	754	754	753	754	754	
median survival (years)	4.1	9.5	5.4	3.8	3.2	2.3	
% 5-year survival	41	75	57	39	25	10	
Age median	63	58	57	61	64	70	
Age range	25-89	39-78	25-86	36-82	27-89	39-86	
Age quartile q1	894	30.8	31.3	20.0	13.4	4.5	$<1x10^{-50}$
Age quartile q2	838	21.5	20.0	22.9	21.2	14.3	
Age quartile q3	961	16.0	20.2	21.4	23.6	18.7	
Age quartile q4	1076	13.5	10.4	16.4	21.3	38.5	
FIGO stage I / II	607	97.4	2.6	0.0	0.0	0.0	$<1x10^{-50}$
FIGO stage III/IV	3067	3.8	23.0	24.1	24.4	24.6	
Primary chemo* 1	136	16.2	22.1	23.5	19.1	19.1	0.163
Primary chemo* 2	190	16.3	20.0	21.6	22.1	20.0	
Primary chemo* 3	361	11.1	16.9	22.4	20.5	29.1	
Residual disease No	614	32.4	22.1	17.8	15.5	12.2	$<1x10^{-50}$
Residual disease Yes	1157	6.0	19.2	24.1	24.5	26.2	
germline BRCA1 mutation	130	23.8	31.5	26.2	11.5	6.9	$2.22x10^{-7}$
germline BRCA2 mutation	71	28.2	26.8	18.3	18.3	8.5	
germline no mutation	663	19.6	16.7	18.7	20.7	24.3	
CD8 TIL score 0	192	19.8	14.6	12.5	21.4	31.8	$2.46x10^{-14}$
CD8 TIL score 1-2	186	18.3	14.0	18.8	21.5	27.4	
CD8 TIL score 3-19	515	19.8	24.1	20.8	17.9	17.5	
CD8 TIL score >20	218	34.4	31.2	16.5	11.5	6.4	
Molecular subtype C1.MES	1105	5.4	10.4	20.7	27.4	36.0	$<1x10^{-50}$
Molecular subtype C2.IMM	907	23.2	28.8	21.2	16.2	10.7	
Molecular subtype C4.DIF	1144	32.6	25.5	17.9	12.8	11.2	
Molecular subtype C5.PRO	613	18.1	14.0	20.7	25.8	21.4	
FIGO stage 1A & 1B	111	96.4	3.6	0.0	0.0	0.0	$<1x10^{-50}$
FIGO stage 3C	1979	3.1	23.7	24.6	24.1	24.6	$<1x10^{-50}$
FIGO stage 3C Residual	316	6.3	31.0	24.4	20.9	17.4	$6.24x10^{-45}$
FIGO stage 3C Residual	846	2.6	21.5	25.3	24.6	26.0	

412 Q1 is the quintile with the best survival and Q5 the worst survival. Samples with missing data are reported in
413 Supplementary Table S11. P-values for BRCA1/2 mutation status were calculated for BRCA1 or BRCA2 mutation
414 vs no mutation. * Treatment: 1 = known to have received first line chemotherapy treatment of ≥ 4 cycles of IV
415 carboplatin AUC 5 or 6 & paclitaxel 135 or 175 mg/m² every 3 weeks. 2 = known to have received first line
416 chemotherapy treatment of ≥ 4 cycles of IV carboplatin & paclitaxel 3-weekly but at doses presumed to be
417 carboplatin AUC 5 or 6 & paclitaxel 135 or 175 mg/m². 3 = all remaining cases with chemo regimens that do
418 not fit criteria 1 or 2 and include unknown or no chemotherapy.
419

420 **Appendix**

421 ***AOCS STUDY GROUP***

422 ***Management Group***: D Bowtell[1,3,4,5,6], G Chenevix-Trench[2], A Green[2], P Webb[2], A DeFazio[7], D Gertig[8]

423

424 ***Project and Data Managers***: N Traficante[1], S Fereday[1], S Moore[2], J Hung[7], K Harrap[2], T Sadkowsky[2], N Pandeya[2]

425

426 ***Research Nurses and Assistants***:

427 M Malt[2], A Mellon[9], R Robertson[9], T Vanden Bergh[10], M Jones[10], P Mackenzie[10], J Maidens[11], K Nattress[12], YE
428 Chiew[7], A Stenlake[7], H Sullivan[7], B Alexander[2], P Ashover[2], S Brown[2], T Corrish[2], L Green[2], L Jackman[2], K Ferguson[2],
429 K Martin[2], A Martyn[2], B Ranieri[2], J White[13], V Jayde[14], P Mamers[15], L Bowes[1], L Galletta[1], D Giles[1], J Hendley[1], K
430 Alsop[1], T Schmidt[16], H Shirley[16], C Ball[17], C Young[17], S Viduka[16], Hoa Tran[16], Sanela Bilic[16], Lydia Glavinas[16], Julia
431 Brooks[18]

432

433 ***Clinical and Scientific Collaborators***:

434 R Stuart-Harris[19], F Kirsten[20], J Rutovitz[21], P Clingan[22], A Glasgow[22], A Proietto[9], S Braye[9], G Otton[9], J Shannon[23],
435 T Bonaventura[24], J Stewart[24], S Begbie[25] M Friedlander[26] D Bell[11], S Baron-Hay[11], A Ferrier[11] (*dec.*), G Gard[11], D
436 Nevell[11], N Pavlakis[11], S Valmadre[11], B Young[11], C Camaris[10], R Crouch[10], L Edwards[10], N Hacker[10], D Marsden[10], G
437 Robertson[10], P Beale[12], J Beith[12], J Carter[12], C Dalrymple[12], R Houghton[12], P Russell[12], M Links[27], J Grygiel[28], J Hill[29],
438 A Brand[30], K Byth[30], R Jaworski[30], P Harnett[30], R Sharma[31], G Wain[30], B Ward[32], D Papadimos[32], A Crandon[33], M
439 Cummings[33], K Horwood[33], A Obermair[33], L Perrin[33], D Wyld[33], J Nicklin[33, 34], M Davy[13], MK Oehler[13], C Hall[13], T
440 Dodd[13], T Healy[35], K Pittman[35,] D Henderson[36], J Miller[37], J Pierdes[37], P Blomfield[14], D Challis[14], R McIntosh[14], A
441 Parker[14], B Brown[38], R Rome[38], D Allen[39], P Grant[39], S Hyde[39], R Laurie[39], M Robbie[39,] D Healy[15], T Jobling[15], T
442 Manolitsas[15], J McNealage[15], P Rogers[15], B Susil[15], E Sumithran[15], I Simpson[15], K Phillips[1], D Rischin[1], S Fox[1], D
443 Johnson[1], S Lade[1], M Loughrey[1], N O'Callaghan[1], W Murray[1], P Waring[3], V Billson[40], J Pyman[40], D Neesham[40], M
444 Quinn[40], C Underhill[41], R Bell[42], LF Ng[43], R Blum[44], V Ganju[45], I Hammond[17], Y Leung[17], A McCartney[17] (*dec.*), M
445 Buck[46,] I Haviv[47], D Purdie[2], D Whiteman[2], N Zeps[16]

[1]Peter MacCallum Cancer Centre, Melbourne, Victoria, 3000, Australia.

[2]QIMR Berghofer Medical Research Institute, Brisbane, Queensland, 4006, Australia.

[3]Department of Pathology, University of Melbourne, Parkville, Victoria, 3052, Australia.

[4]Sir Peter MacCallum Cancer Centre Department of Oncology, University of Melbourne, Parkville, Victoria, 3052, Australia.

[5]Department of Biochemistry and Molecular Biology, University of Melbourne, Parkville, Victoria, 3052, Australia.

[6]Ovarian Cancer Action Research Centre, Department of Surgery and Cancer, Imperial College London, London, England, W12 0HS, UK.

[7]Centre for Cancer Research, The Westmead Institute for Medical Research, The University of Sydney and Department of Gynaecological Oncology, Westmead Hospital, Sydney, New South Wales, 2145, Australia.

[8]Melbourne School of Population and Global Health, University of Melbourne, Parkville, Victoria, 3052, Australia.

[9]John Hunter Hospital, Lookout Road, New Lambton, New South Wales, 2305, Australia

[10]Royal Hospital for Women, Barker Street, Randwick, New South Wales, 2031, Australia

[11]Royal North Shore Hospital, Reserve Road, St Leonards, New South Wales, 2065, Australia

[12]Royal Prince Alfred Hospital, Missenden Road, Camperdown, New South Wales, 2050, Australia

[13]Royal Adelaide Hospital, North Terrace, Adelaide, South Australia, 5000, Australia

[14]Royal Hobart Hospital, 48 Liverpool St, Hobart, Tasmania, 7000, Australia

[15] Monash Medical Centre, 246 Clayton Rd, Clayton, Victoria, 3168, Australia

[16]Western Australian Research Tissue Network (WARTN), St John of God Pathology, 23 Walters Drive, Osborne Park, Western Australia, 6017, Australia

[17]Women and Infant's Research Foundation, King Edward Memorial Hospital, 374 Bagot Road, Subiaco, Western Australia, 6008, Australia

473 [18]St John of God Hospital, 12 Salvado Rd, Subiaco, Western Australia, 6008, Australia

474 [19]Canberra Hospital, Yamba Drive, Garran, Australian Capitol Territory, 2605, Australia

475 [20]Bankstown Cancer Centre, Bankstown Hospital, 70 Eldridge Road, Bankstown, New South
476 Wales, 2200, Australia

477 [21]Northern Haematology & Oncology Group, Integrated Cancer Centre, 185 Fox Valley Road,
478 Wahroonga, New South Wales, 2076, Australia

479 [22]Illawarra Shoalhaven Local Health District, Wollongong Hospital, Level 4 Lawson House,
480 Wollongong, New South Wales, 2500, Australia

481 [23]Nepean Hospital, Derby Street, Kingswood, New South Wales, 2747, Australia

482 [24]Newcastle Mater Misericordiae Hospital, Edith Street, Waratah, New South Wales, 2298, Australia

483 [25]Port Macquarie Base Hospital, Wrights Road, Port Macquarie, New South Wales, 2444, Australia

484 [26]Prince of Wales Clinical School, University of New South Wales, New South Wales, 2031, Australia

485 [27]St George Hospital, Gray Street, Kogarah, New South Wales, 2217, Australia

486 [28]St Vincent's Hospital, 390 Victoria Street, Darlinghurst, New South Wales, 2010, Australia

487 [29]Wagga Wagga Base Hospital, Docker St, Wagga Wagga, New South Wales, 2650, Australia

488 [30]Crown Princess Mary Cancer Centre, Westmead Hospital, Westmead, Sydney, New South Wales,
489 2145, Australia.

490 [31]Department of Pathology, Westmead Clinical School, Westmead Hospital, The University of Sydney,
491 New South Wales, 2006, Australia

492 [32]Mater Misericordiae Hospital, Raymond Terrace, South Brisbane, Queensland, 4101, Australia

493 [33]The Royal Brisbane and Women's Hospital, Butterfield Street, Herston, Queensland, 4006, Australia

494 [34] Wesley Hospital, 451 Coronation Drive, Auchenflower, Queensland, 4066, Australia

495 [35]Burnside Hospital, 120 Kensington Road, Toorak Gardens, South Australia, 5065, Australia

496 [36]Flinders Medical Centre, Flinders Drive, Bedford Park, South Australia, 5042, Australia

497 [37]Queen Elizabeth Hospital, 28 Woodville Road, Woodville South, South Australia, 5011, Australia

498 [38]Freemasons Hospital, 20 Victoria Parade, East Melbourne, Victoria, 3002, Australia

499 [39]Mercy Hospital for Women, 163 Studley Road, Heidelberg, Victoria, 3084, Australia

500 [40]The Royal Women's Hospital, Parkville, Victoria, 3052, Australia

501 [41]Border Medical Oncology, Wodonga, Victoria, 3690, Australia

502 [42]Andrew Love Cancer Centre, 70 Swanston Street, Geelong, Victoria, 3220, Australia

503 [43]Ballarat Base Hospital, Drummond Street North, Ballarat, Victoria, 3350, Australia

504 [44]Bendigo Health Care Group, 62 Lucan Street, Bendigo, Victoria, 3550, Australia

505 [45]Peninsula Health, 2 Hastings Road, Frankston, Victoria, 3199, Australia

506 [46]Mount Hospital, 150 Mounts Bay Road, Perth, Western Australia 6000, Australia

507 [47]Faculty of Medicine, Bar-Ilan University, 8 Henrietta Szold St, Safed, Israel

508

509 The six people underlined are named authors on the manuscript.

510

511 # Supplemental Methods

512 **Gene selection based on transcriptome-wide meta-analysis**

513 A meta-analysis was conducted of six transcriptome-wide microarray studies, including 1,455 participant's
514 tumors, to investigate the role of gene expression in HGSOC OS (Supplemental Table 2)[98-103]. The objective
515 was to identify prognostic genes whose effects were consistent across studies and were not surrogates for
516 molecular subtype. The total number of genes tested was 15,345, however, the number of genes in each study
517 differed due to differing microarray platforms used across the studies as well as different data processing and
518 quality control work flows. Expression data were normalized, batch corrected, and extreme outliers removed
519 on a probe-specific level for each study. Extreme outliers were defined as values greater than 2.5 times the
520 interquartile range from the upper or lower quartile, under the constraint that no more than three percent be
521 classified as such. That is, no more than three percent of observations were removed as outliers. According to
522 the microarray platform design, some genes are represented by more than one expression feature. To yield a
523 single expression feature per gene, principal components analysis (PCA) was applied to probe sets within each
524 study, taking the first PC to represent the gene. This approach is similar to that used for the The Cancer
525 Genome Atlas (TCGA) unified gene expression data[99]. To reduce the dependency of the identified genes on
526 the analytic approach, four types of analyses were conducted to evaluate genes for selection.

527 1) *Consistency of OS association across studies*. Cox proportional hazards regression was conducted separately
528 for each study, adjusting for age, stage and study-provided molecular subtype (C1/mesenchymal, C2/immune,
529 C4/differentiated and C5/proliferative). The median of the study-specific p-values for association between
530 gene expression and prognosis was used as the statistic for an omnibus test across studies. Missing p-values
531 were set to one. Consistency in direction of effect was accounted for by setting the median p-value to one if
532 the signs of the effects differed for any of the studies with p-values equal to or smaller than the median p-
533 value. To account for multiple-testing and possible unknown characteristics of the null distribution of the

534 median p-value, we employed a permutation-based false discovery rate (FDR) approach[104] with 100
535 replicate permutations, considering the median p-value to be our test statistic. Permutation analyses were
536 conducted by permuting sample labels on the expression data within each study, while maintaining the
537 observed relationships between genes and between the outcome and adjustment covariates, including
538 molecular subtypes. Thus, for each gene and each of the 100 permutations, a median p-value was computed
539 under the null by grouping the permutation p-values according to the permutation index. This analysis resulted
540 in 115 genes (Supplemental Table S2) identified at FDR < 0.05 level (Supplemental Figure S1. A), all of which
541 were selected for the follow-up study. Note that this significance threshold corresponds to a median p-value
542 that is slightly greater than 0.05. Nevertheless, this threshold defines statistical significance because it is
543 evaluated against the distribution of median p-values under the null hypothesis of independence between
544 gene expression and outcome.

545 2) *Stratified analysis of marginal effects*. A Cox model was fitted and a likelihood ratio test (LRT) conducted for
546 each gene, adjusted for age, stage, and molecular subtype (differentiated, immunoreactive, mesenchymal, and
547 proliferative)[120], and stratifying on study. The result was an additional 12 genes at an FDR significance level
548 of 0.05.

549 3) *Evidence of interaction with molecular subtype on OS*. For each gene, a stratified Cox model adjusted for
550 covariates was fitted as in (1), but molecular subtype×gene interaction terms were included to identify genes
551 whose effects differed across subtypes. LRTs were conducted for interaction terms, yet no additional genes
552 were identified in this analysis at the 0.05 FDR significance level.

553 4) *Stratified analysis with multi-degree-of-freedom tests of main effects and interactions*. For fitted models in
554 (3), LRTs were conducted to jointly assess gene main effects and interactions. This analysis yielded an
555 additional eight genes.

556 An additional 65 genes were added that were suggestive in one of the meta-analysis described above, but that
557 did not meet the 0.05 significance level, based on evidence in prior literature and public knowledge databases
558 such as MSigDB and REACTOME. In total 200 genes were selected from the meta-analysis (Supplemental Table
559 S3).

560 **Selection of additional candidate genes**

561 An additional 304 candidate genes were selected based on one of the following criteria (a) evidence of
562 association with prognosis or potential drug targets from the literature, (b) residing within a 1 MB region of a
563 potential survival GWAS hit $p<5x10^{-6}$ and showing a survival association in the publicly available TCGA data (c)
564 utility in molecular subtype classification (Talhouk et al, in preparation), (d) other specific hypothesis. Five
565 genes, *RPL19*, *ACTB*, *PGK1*, *SDHA*, and *POLR1B*, were included as house-keeping genes for normalisation[105].
566 An additional six genes, *TBP, GAPDH, KIF3B, GUSB, BMS1,* and *RPL41*, were included to evaluate consistency
567 with previous codeset analysis but were not used in the normalisation[105]. Finally, ten genes were selected as
568 a "tagging" approach to increase representation of the gene expression patterns of other genes that are
569 correlated using the methods in Rudd et al[121]. For this study, we chose a threshold of 99% correlation,
570 observed in all four of the largest publicly-available HGSC ovarian cancer gene expression datasets[98, 99, 103,
571 122]. We determined that the 503 genes already selected included 99% correlated gene expression
572 information for an additional 2,617 genes. Another 10 genes were selected in order to maximize gene
573 expression data in other parts of the transcriptome that were not represented, and these 10 genes represent
574 gene expression for another 49 genes. Seven of the candidate genes overlapped with genes selected from the
575 meta-analysis, therefore 313 additional genes were added to the custom code set (Supplemental Table 4).

576 **Single gene analysis of associations with OS.**

577 Time-to-event analyses were carried out for OS with right-censoring at 10 years and left-truncation of
578 prevalent cases. For most genes, the association between log-transformed normalized gene expression for

579 each gene and survival time was evaluated using Cox proportional hazards models applied to the full data set.
580 However, in the analysis of those genes that were selected due to the meta-analysis results, the 211 cases that
581 were also represented in the meta-analysis data set were excluded. All single-gene models were adjusted for
582 age, race, and stage, and stratified by study. Patients with missing race or stage information were assigned to
583 'unknown' categories. Age was modelled using a B-spline with a knot at the median age, which yielded a
584 better fit than using knots at quartiles or categorical variables. Stage was dichotomized into early
585 (International Federation of Gynecology and Obstetrics [FIGO] stage I/II) and advanced (FIGO stage III/IV).
586 Expression of each gene was scaled to have a standard deviation of one, so hazard ratios correspond to a
587 change of one standard deviation. A Benjamini-Hochberg (BH) false discovery rate (FDR) of less than 0.05 was
588 used to identify notable associations. Advanced stage ovarian cancer usually has disease spread throughout
589 the abdomen, therefore sensitivity analyses were performed to assess effects of the anatomical location of
590 tumor samples included in the study by removing observations corresponding to samples known to be
591 extraovarian (*N* = 437).

592 **Gene-expression profiling**

593 For RNA extractions, all sites performing NanoString reactions followed standard operating procedures
594 outlined in advance. RNA was extracted from FFPE tumor samples using the Qiagen miRNeasy FFPE kit and
595 were processed with the NanoString nCounter technology using a custom codeset. Briefly, each day sites
596 processed a maximum of 24 samples. Our standard operating procedure called for 500ng of total RNA, as
597 measured from NanoDrop, combined with hybridization buffer and a custom NanoString reporter and capture
598 CodeSet allowing hybridization for exactly 16 hours (short-hyb, 12 samples per day) or 20 hours (long-hyb, 12
599 samples per day) at 65°C in a pre-heated thermal cycler. Immediately at the end of the prescribed
600 hybridization period samples were processed on an nCounter prep-station (NanoString) following standard
601 procedures. Loaded cartridges (12 samples) were scanned at maximum resolution on an nCounter Digital
602 Analyser (NanoString). The BC Cancer (Vancouver) site performed scanning on a Gen1 Digital Analyzer, while

603 both USC (Los Angeles) and PMC (Melbourne, sometime denoted as AOC or Australian Ovarian Cancer study)
604 performed scanning on Gen2 Digital Analyzers. Relevant variables including processing date, operator, site,
605 and hybridization time were recorded/embedded into specimen information (CDF) and data files (RCC). In
606 addition to unique HGSOC samples a number of controls and sample replicates were run at all sites to enable
607 evaluation of data quality.

608 *Reference Pools*. To monitor for technical bias across sites and allow for cross-CodeSet comparisons 35, we ran
609 3 distinct control RNA pools. This reference-based normalization strategy is considered best practice for
610 development of NanoString based clinical tests and is similar to the implementation already in use for
611 Prosigna[123] and a number of other in development tests. Pools consisted of high-quality RNA from fresh-
612 frozen ovarian cancer samples believed to be representative of all molecular subtypes and/or various ovarian
613 cancer histotypes. Pools were assembled en-mass and aliquoted (5ul, 100ng total RNA) for single use without
614 multiple freeze thaws at all sites. Control aliquots were stored at -80°C until ready for use and shipped on dry
615 ice to all processing sites. Pool1 was run approximately every month at each site. Pool2 and Pool3 were run
616 alternatingly, every other month, at each site.

617 *Cross-Site Controls*. In addition to control pools, a subset of 48 samples were run once at each of the three
618 processing centres (144 individual run files created, 1 failed QC). The first 36/48 consisted of randomly
619 selected high-grade serous ovarian carcinoma specimens, 12 from each processing centre. In addition, the
620 Vancouver site selected 12 samples from non-High-grade serous histology samples (3 clear cell, 3
621 endometrioid, 3 low-grade serous, 3 mucinous). Aliquots of RNA chosen at each site were sent on dry ice to
622 the other two processing centres. RNA from the 48 of the tumor samples were run on all three instruments to
623 assess concordance and the average r-squared was 0.981 (range 0.758-0.996). RNA from 1-2% of the samples
624 were randomly selected as technical replicates and run a second time to access concordance and to identify
625 any systematic problems with sample labelling. All 98 pairs of samples were concordant and the average r-
626 squared was 0.978 (range 0.753-0.998).

627 **Quality control and normalization of gene-expression data**

628 Raw data were assessed using several quality assurance (QA) metrics to measure imaging quality,
629 oversaturation and overall signal to noise.

630 1. *Imaging quality controls*: Samples were flagged as imaging failures if the percentage of lane images FOV
631 obtained was less than 75% of the requested number of fields.

632 2. *Linearity of the assa*y: Samples were flagged as linearity failures if spiked-in positive control probes at
633 different concentrations had R^2<0.95.

634 3. *Detection of Smallest Positive Control*: Samples were flagged when the 0.5 fM positive control probe
635 smaller than 2 standard deviations from the mean of the negative controls probes.

636 4. *Sample Quality*. Thresholds were set to maximize the number of samples of high quality included in the
637 analysis. **Sample Quality** fails if either the **Limit of Detection** or **Signal to Noise** thresholds are not met.

638 a. *% of Genes above Limit of Detection (LOD) of negative controls*: LOD is an upper bound of the
639 background noise in the system, computed as two standard deviations above the mean of the spiked-
640 in negative control probes. Samples below a 50% threshold were deemed of poor quality and
641 considered failures.

642 b. *Signal to noise ratio (S/N):* calculated as a ratio between the geometric mean of housekeeping
643 genes and lower limit of detection: geometric mean/LOD. Samples with signal to noise ratio below a
644 170 threshold were deemed of poor quality and considered failures.

645 5. Overall QC. This is an overall quality control flag which fails if any of the **Imaging**, **Linearity**, or **Smallest**
646 **Positive Control** conditions fail.

647 **Batch correction using control pools**

648 The reference sample methods described in Talhouk et al was used. Briefly, assuming two batches A and B. To
649 calibrate samples with gene expressing X^B, that were run in batch B to samples with gene expression X^A, that
650 were run in batch A,

651 · Some number of reference samples (R) would be run in both batches A and B, resulting in expression
652 R^A and R^B.

653 · To remove Batch Effect: X^B-R^B *and* X^A-R^A

654 · Or alternatively: $X^B + (R^A - R^B)$ would result in calibrating batch B to batch A.

655 As the same CodeSet was observed at all three sites, little difference was observed across sites; for
656 consistency, everything was calibrated to the Vancouver batch. The count data were $\log_2$ transformed and
657 centered by the arithmetic mean of the selected housekeeping genes, *RPL19*, *ACTB*, *PGK1*, *SDHA*, and *POLR1B*.

658

659

660 **Prognostic signature**

661 To compare and evaluate competing methods for development of the signature, we used an AUC approach
662 implemented in the "survivalROC" R software package, designed for a time-dependent setting with censoring
663 where a lag exists between measurement of the biomarker and the disease outcome[124]. A ten-fold cross-
664 validation approach was used, thus each modelling approach was applied to a randomly selected nine-tenths
665 of the training data, then the AUC was computed on the remaining one-tenth (out-of-bag sample; OOB) based
666 on the continuous biomarker generated by the model. These AUC values guided selection of the best method,
667 which was then applied to the full training dataset to identify the signature. Elastic net was the best

performing approach, hence the approach described below in B was applied to the full training dataset to develop the final model. This signature was then evaluated in the testing data that were set aside.

A. *Stepwise*. Each gene was initially modelled separately, then those significant at a Bonferroni corrected 0.05 level were jointly modelled using a backwards stepwise approach, sequentially removing genes that did not achieve a nominal 0.05 significance level. This process yielded a final model with coefficients that defined a linear combination of adjustment covariates and expression features, forming a prognostic biomarker for OS.

B. *Elastic net*. In this regression method, a mixture of l_1 (lasso) and l_2 (ridge regression) penalties are applied. A version has been developed for the Cox proportional hazards model, available in the glmnet R software package[125]. For each step of the overall cross-validation, the penalty parameter, λ, was selected as that with the minimum mean cross-validated error as measured by the Cox partial likelihood for each randomly selected nine-tenths of the training data, using the R function cv.glmnet from the glmnet R package. Thus, nested cross-validation was conducted. The covariates age and stage were included in the model as mandatory categorical variables by applying no penalization. Genes with non-zero coefficients in the penalized Cox model were selected for the final model.

C. *Boosting*. Component-wise gradient boosting, a machine learning method, employs gradient descent techniques to optimize a combined variable selection and model building strategy. It has been implemented for general linear models and Cox proportional hazards model in the R package mboost[126]. To constrain the model to account for age and stage we included an offset computed from the linear predictors from a Cox proportional hazards model. Boosting is an iterative strategy where at each step, parameters are updated by weak estimators according to a pre-specified loss function. An important tuning parameter that can influence over-fitting is the number of iterations, m_{stop}. We used 10-fold cross-validated estimates of empirical risk to choose m_{stop}. Thus, nested cross-validation was performed for parameter tuning, similar to the elastic net approach.

691 D. *Random Survival Forests*. This machine learning approach designed for survival outcomes[127] is an
692 ensemble method based on multiple decision trees, which are weakly predictive individually but combine to
693 yield what can be a strong predictor. The analysis was conducted using the rfsrc function in the R package,
694 randomForestSRC. To enhance performance in settings with large numbers of predictors, a preliminary feature
695 selection step is sometime performed. We conducted the analysis restricting to genes significant at a 0.05
696 Bonferroni corrected alpha level, similar to the stepwise approach described above. The tuning parameters for
697 these analyses included the number of trees, set to 200, and the maximum number of splits for each tree, set
698 to 20.

699 **Prognostic signature performance in relation to gene set size**

700 To further assess the need for all 101 genes included in the signature, we computed AUC sequentially adding
701 genes according to the magnitudes of the corresponding coefficients. The gene with the most extreme
702 coefficient was added first, followed by the gene with the next most extreme coefficient, and so on. Age and
703 stage were included in the signature, and the AUC was computed using individuals from the test set only.

704 **Genes correlated with genes in the prognostic signature**

705 To find additional genes co-expressed with genes within our NanoString panel a correlation analysis was
706 performed within The Cancer Genome Atlas (TCGA) Ovarian Serous Cystadenocarcinoma RNA-seq data set. A
707 correlation coefficient was estimated for each NanoString panel gene vs all other genes included in the TCGA
708 dataset. Genes with FDR < 0.05 were considered truly coexpressed and were considered strongly co-expressed
709 if both Pearson and Spearman correlation coefficients were greater than 0.75 in absolute value.

710 Biological pathways that are enriched in the prognostic gene signature were investigated by computing
711 overlaps between the 101 genes in the prognostic signature and the Reactome database1[128] using the
712 Broad Institute's Molecular Signature Database (MSigDB)[129]. Results were obtained for the top 20 pathways
713 enriched within the gene set that had a FDR < 0.05. A similar analysis was performed to examine the biological

714 pathways that are enriched in genes that strongly co-expressed with the prognostic signature genes in the
715 TCGA data set.

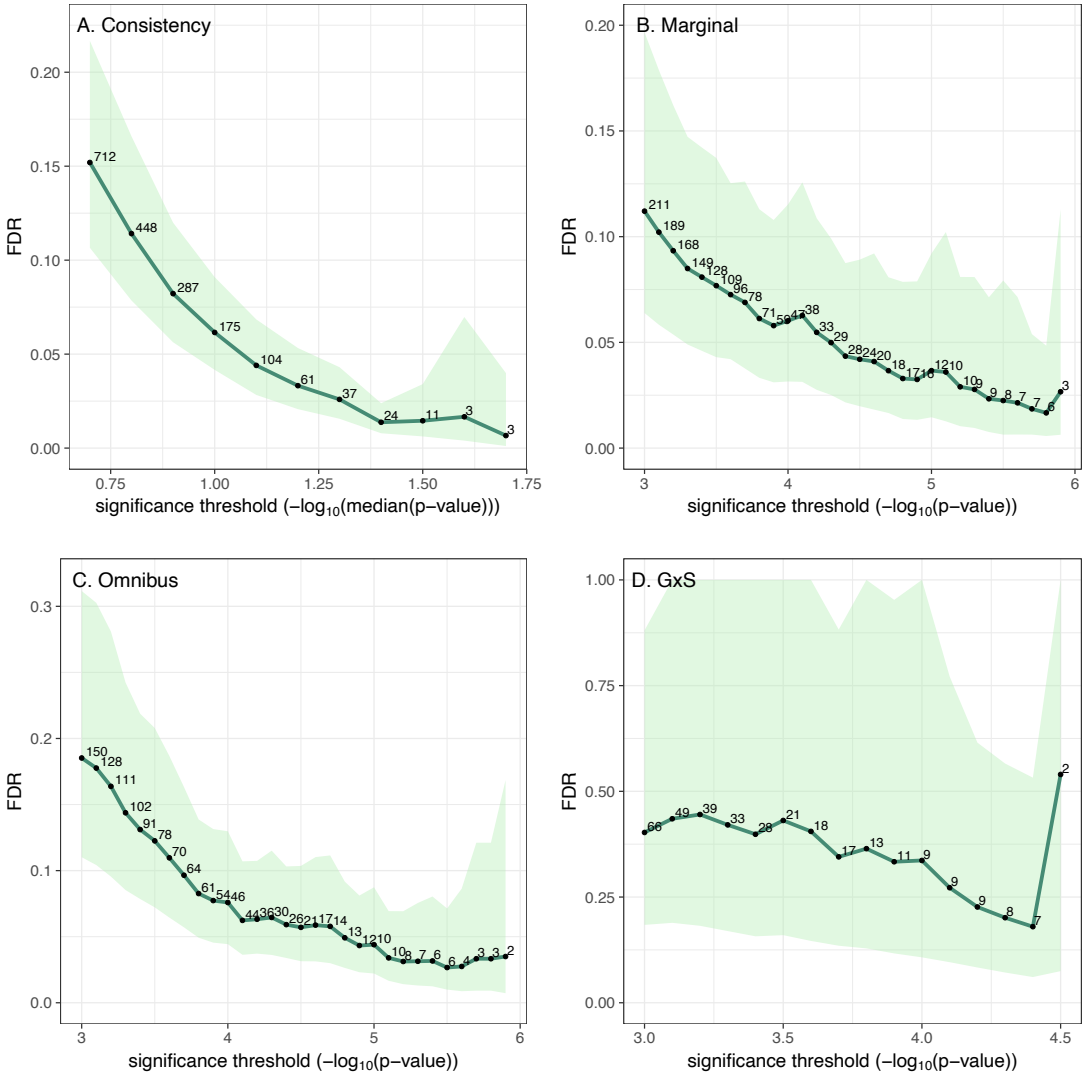
716 **Assigning signature score and hazard ratio to new patients**

717 For future use of this signature with individual patients, it is necessary to compute the signature score based
718 on the coefficients for age, stage, and gene expression that are specified in Supplemental Table S10,
719 GeneSignatureCoefficients. Once a new FFPE sample is obtained, the following steps are taken to generate the
720 score:

721 1. An FFPE tumor sample is processed and profiled using the NanoString platform and codeset developed
722 for the current study (GEO GPL26748) according to methods described above. The quality controls
723 steps and normalization must be adhered to as described, including use of the housekeeping genes
724 included in the codeset.
725 2. Patients are assigned to an age quartile group, and the corresponding variables, age.fq2, age.fq3,
726 age.fq4, are assigned a 1 if the patient fits this category or 0 otherwise. Cut-points were based on our
727 training data, ages 53, 60 and 67. Specifically 53 <= age.fq2 < 60, 60 <= age.fq3 < 67, age.fq4 >= 67.
728 Patients under age 53 are in the reference group, thus the three age.fq variables are all assigned 0 for
729 these patients.
730 3. Patients are assigned to a stage group, and the corresponding variables, stage.f1 and stage.f8 are
731 assigned a 1 if the patients fits this category or 0 otherwise. stage.f1 is defined as FIGO I-II and stage.f8
732 is assigned when FIGO is unknown. Patients with FIGO III-IV are in the reference category, thus stage.f1
733 and stage.f8 are both assigned 0 for these patients.
734 4. Each coefficient in Table S10 is multiplied by its corresponding age, stage or gene expression value.
735 5. The products are then summed to compute the scores. That is, letting *x* denote the vector of values
736 and β denote the vector of coefficients, the score, *s*, is equal to the matrix product, $s = x\beta^T$.

737 The score can be used to estimate the HR relative to the median score observed here, which was -
738 0.1664587 (mean = -0.2965882). The estimated HR for a given patient relative to the median would be
739 computed as exp((observed score – median) * 1.215), where 1.215 was the COX regression coefficient for
740 the signature score in the testing group in the present study. Scores for future patients can be compared
741 to the patient characteristics reported here by the quintile groups shown in Supplemental Table S11 by
742 noting that the cut-points for the scores defining quintiles in the full dataset were -0.7320, -0.3126, -
743 0.0255, and 0.2658.
744

745 **Supplementary figures**



746

747 **Figure S1**. Permutation-based FDR for microarray meta-analysis results. Plot A displays FDR estimates and 95
748 percent confidence regions for a series of increasingly stringent significance thresholds defined by the median
749 p-value across the six studies. Plots B-D are permutation-based FDR results for the joint analyses across the six
750 studies. The Omnibus tests, C, allow the gene effect to vary across molecular subtype, whereas GxS tests, D,
751 assess evidence for differences in effect across molecular subtypes.

752

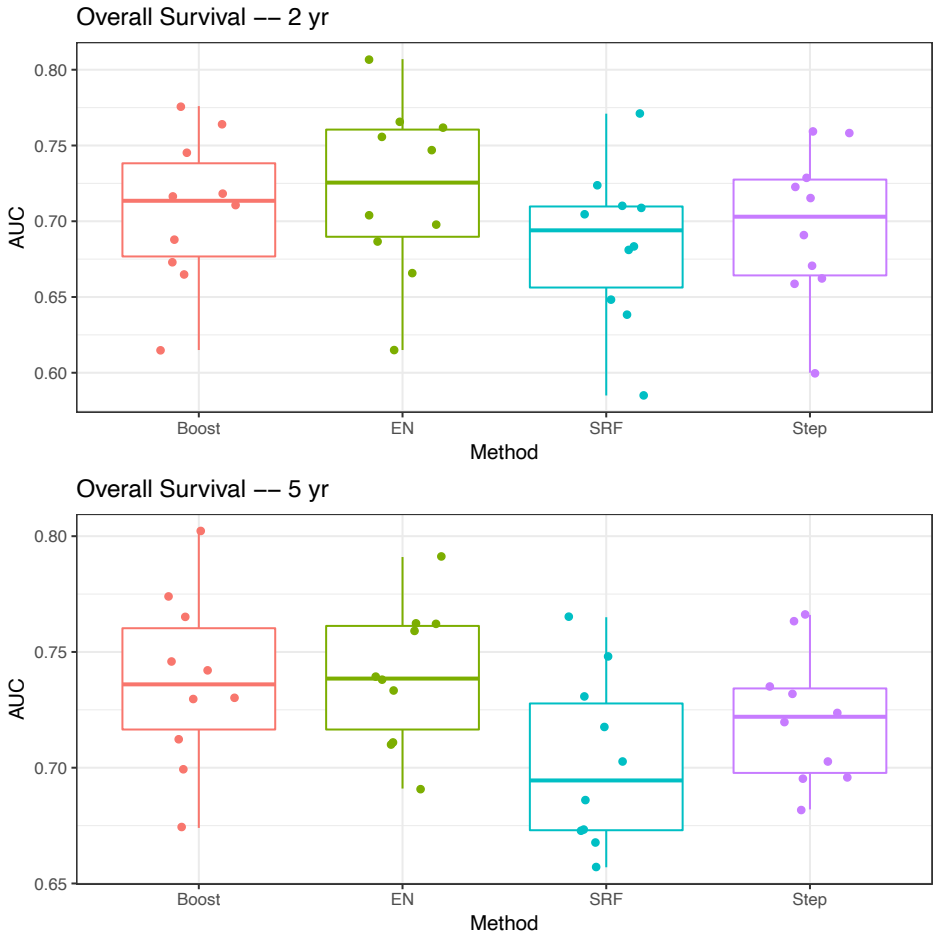


Figure S2. Boxplots of AUC values generated from 10-fold cross validation for competing gene expression prognostic biomarker modelling approaches for two-year and five-year overall survival. Each approach was applied to 10 datasets, each of which was a randomly sampled nine tenths of the training data. AUC values were computed using these models applied to the remaining one tenths of the training data. Methods include boosting (Bst), elastic net (EN), random forests with BH feature selection (RF.BH), random forests with Bonferroni feature selection (RF.Bn), stepwise with BH selection (stp.BH), and stepwise with Bonferroni selection (stp.Bn). The superior performance of the elastic net method as compared to the backward stepwise approach may be explained by the ability of elastic net to account for dependencies among the genes. The fact that it performed better than random survival forests may indicate that Cox proportional hazards models are relatively good at modeling the relation between gene expression and OS.

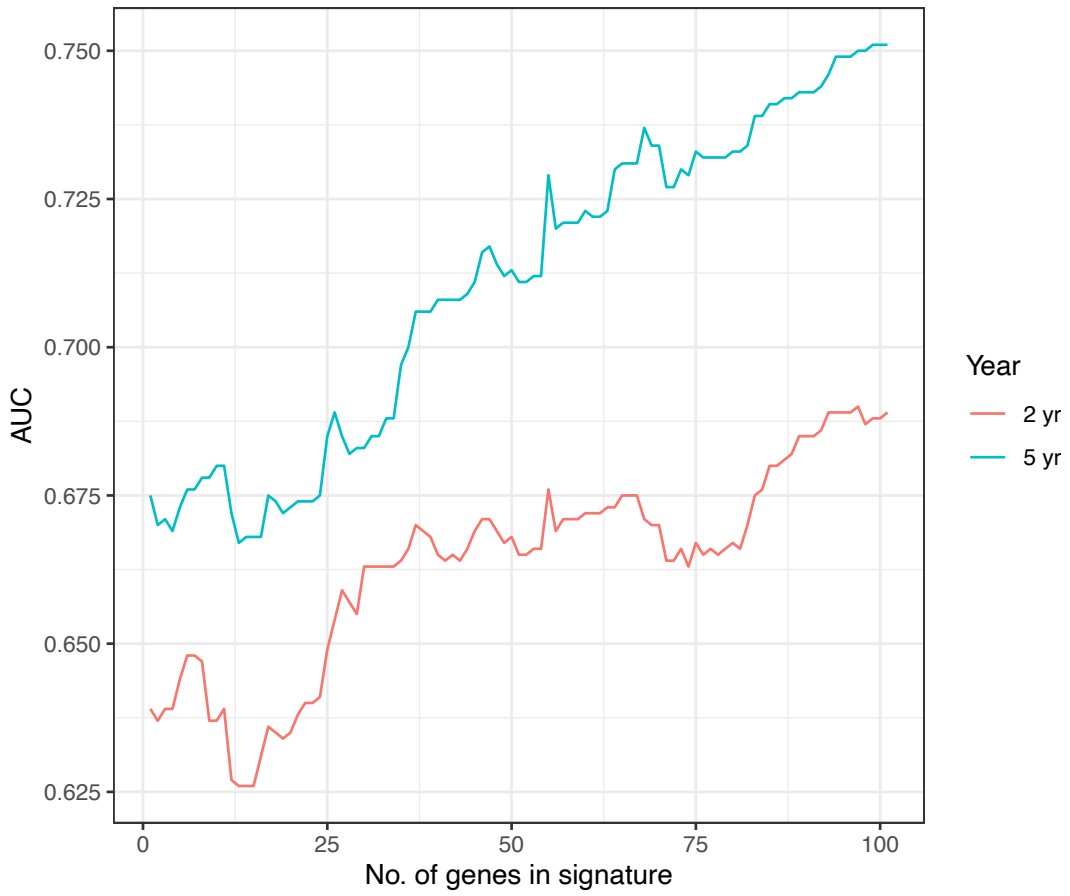


Figure S3. Graph of AUC from sequentially adding genes according to the magnitudes of the corresponding coefficients. Age and stage were included in the signature, and the AUC was computed using individuals from the test set only. The figure shows increasing AUC as the number of genes increases to the final size of 101, with no clear indication of an asymptote. While it may be possible to reduce the number of genes in the signature without impacting signature performance, the above results do not suggest specific genes that could be removed.

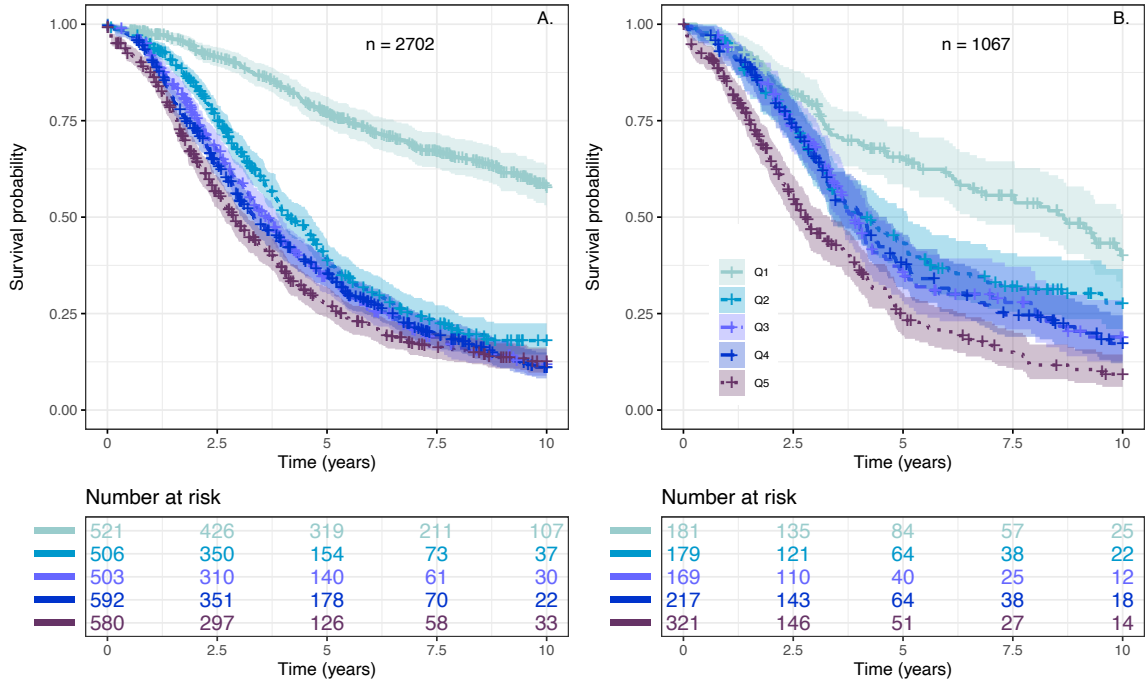


Figure S4. KM curves of overall survival for patients in the training (A) and testing (B) sets. Patients were assigned to quintiles by risk according to age and stage as estimated in Cox models. The smallest quintile has the lightest shade and increasingly darker shades correspond to larger quintiles. Quintiles were calculated independently for the training (A) and testing (B) sets. Shaded regions indicate 95 percent confidence regions.

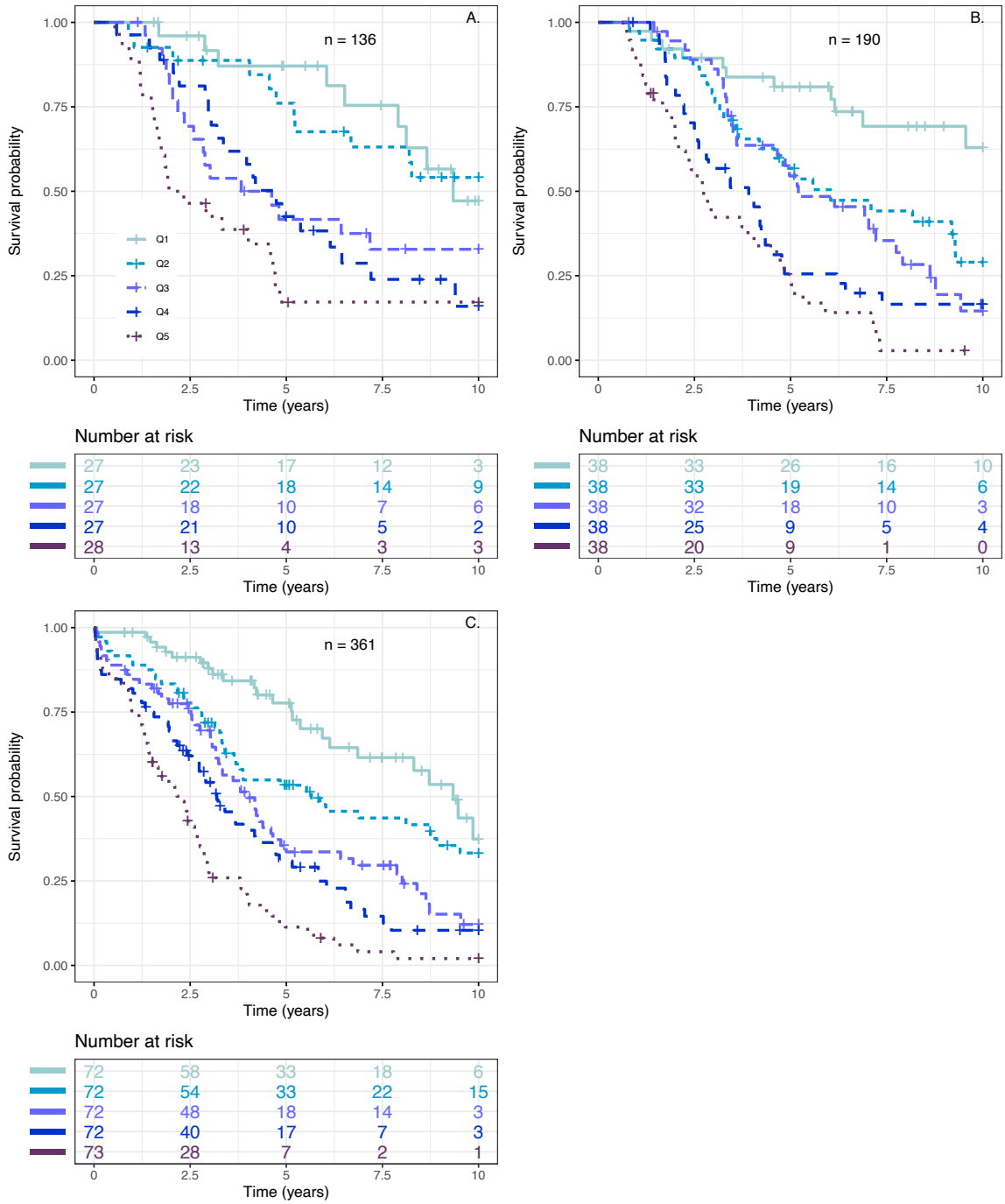


Figure S5. KM curves of overall survival for patients with A) known primary treatment (≥ 4 cycles of IV carboplatin AUC 5 or 6 & paclitaxel 135 or 175 mg/m² every 3 weeks), B) known primary treatment with presumed standard doses (≥ 4 cycles of IV carboplatin & paclitaxel every 3 weeks) and C) all remaining cases with chemotherapy regimens that do not fit criteria in A or B, includes unknown or no chemotherapy. Patients were assigned to quintile groups, Q1 to Q5, based on the signature score, with Q1 having the lightest shade and increasingly darker shades corresponding to quintiles with greater scores. Quintiles were calculated independently for each of the three treatment groups.

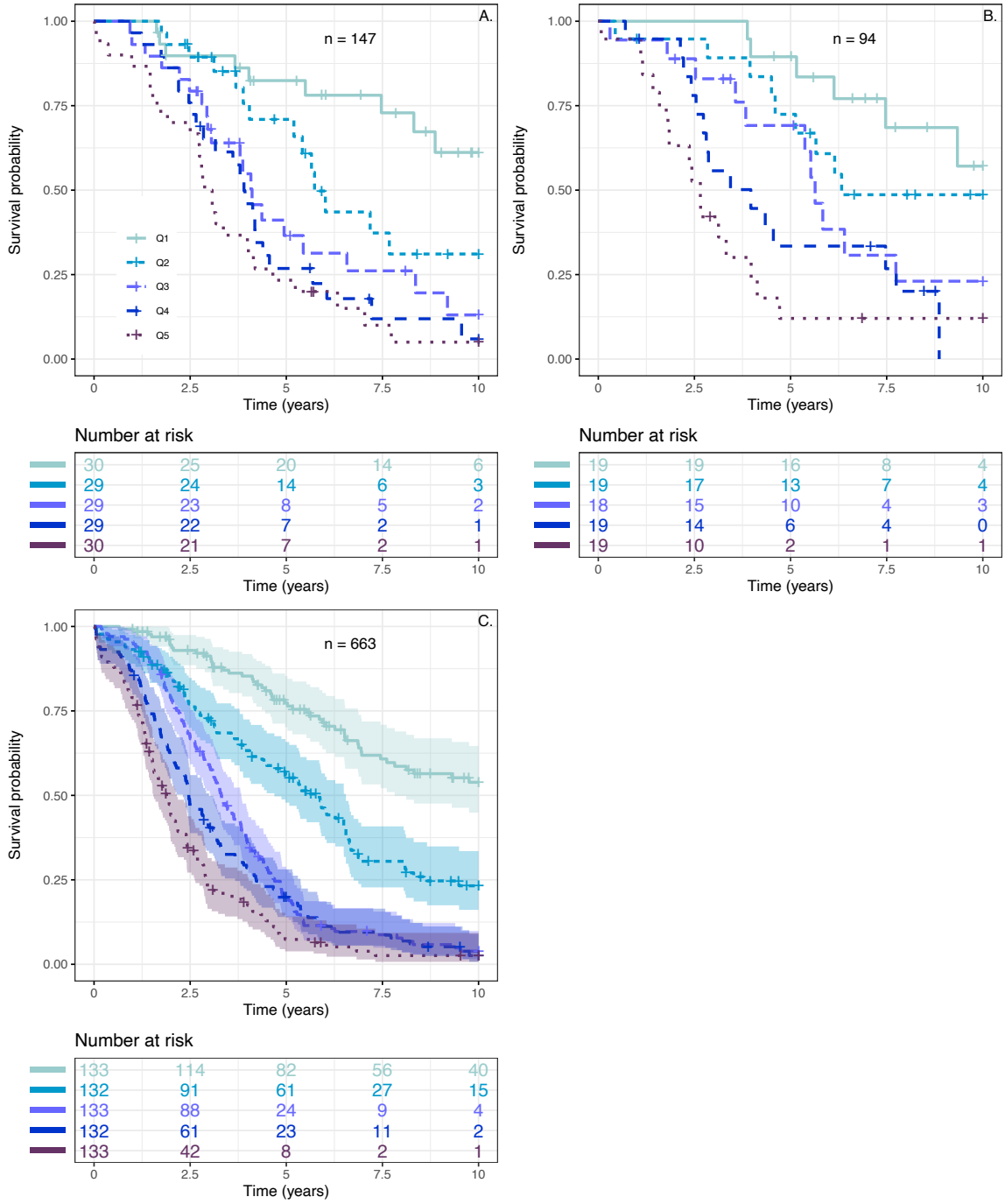


Figure S6. KM curves of overall survival for patients with A) BRCA1, B) BRCA2, and C) no BRCA1 or BRCA2 germline mutations. Patients were assigned to quintile groups, Q1 to Q5, based on the signature score, with Q1 having the lightest shade and increasingly darker shades corresponding to quintiles with greater scores. Quintiles were calculated independently for each of the three BRCA groups.

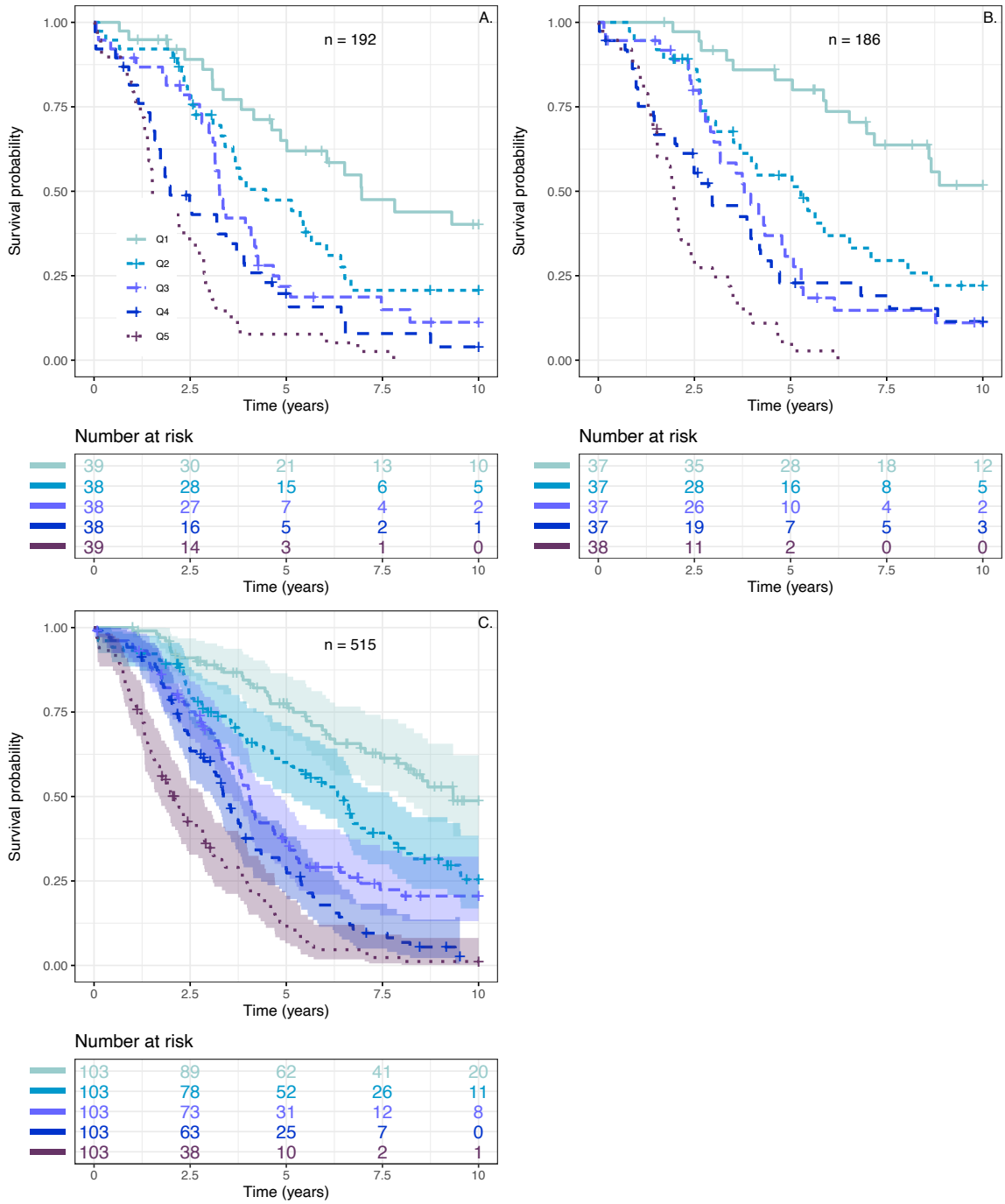


Figure S7. KM curves of overall survival for patients with A) CD8 score equal to 0, B) CD8 score equal to 1 or 2, and C) 2 < CD8 score < 20. Patients were assigned to quintile groups, Q1 to Q5, based on the signature score, with Q1 having the lightest shade and increasingly darker shades corresponding to quintiles with greater scores. Quintiles were calculated independently for each of the three CD8 groups.

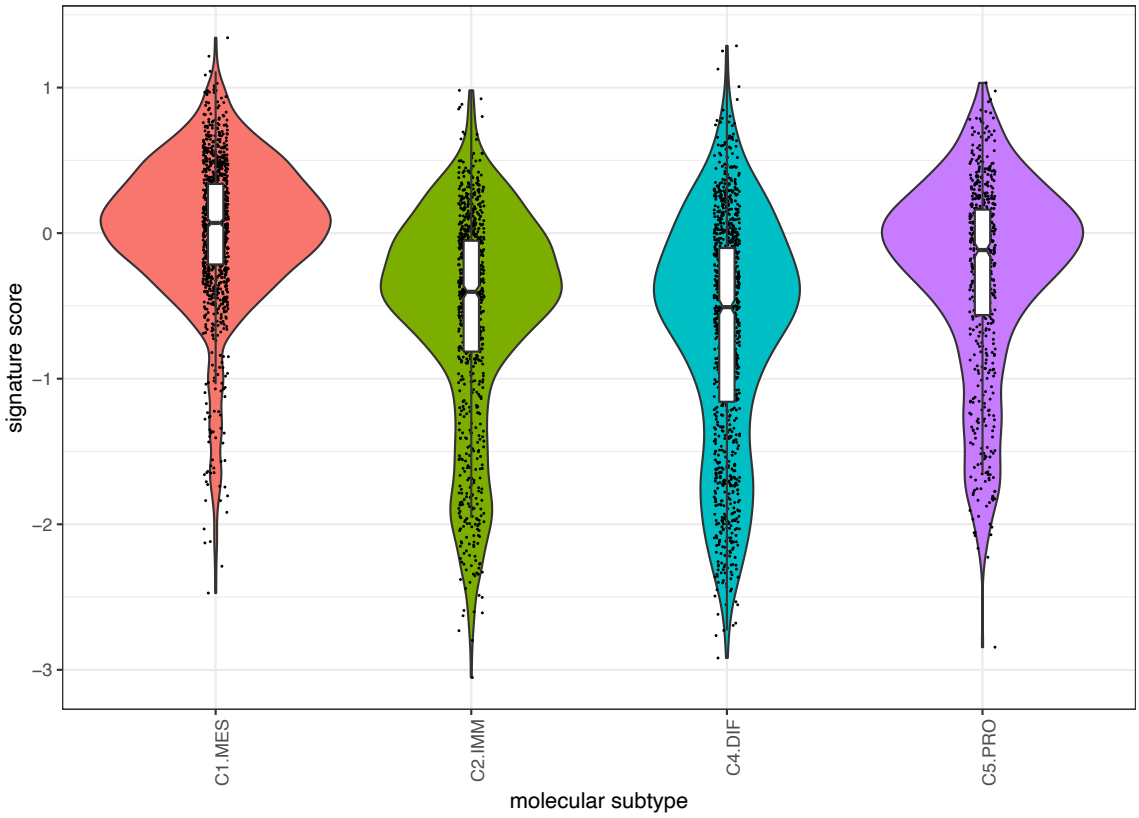
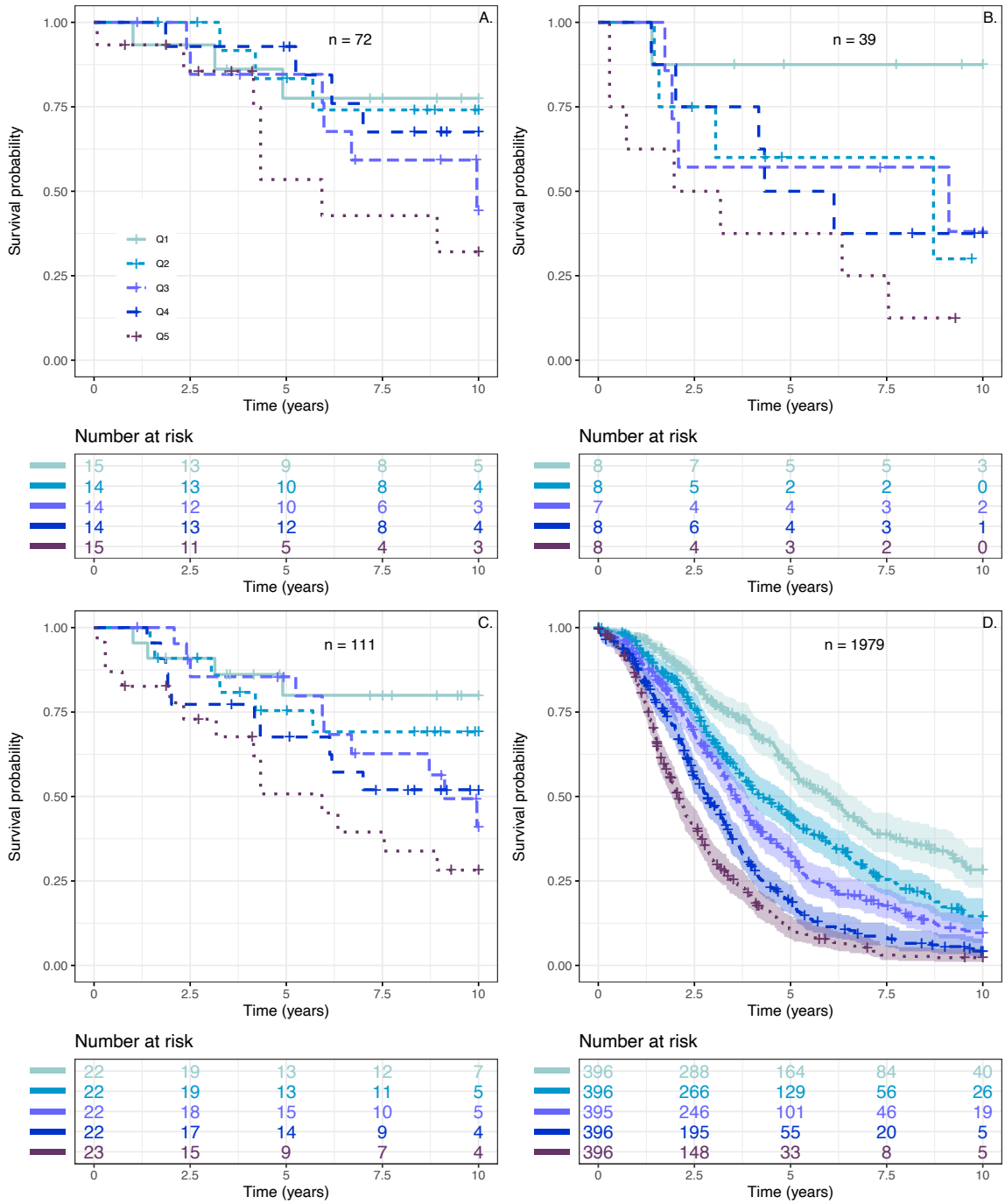


Figure S8. Violin plots of 101-gene signature scores for all 3769 HGSOC patients according to their molecular subtype. The data used to assign molecular subtypes overlapped with the data used to compute the signature score, however, hazard ratios for overall survival across quintiles of the signature were only minimally impacted by molecular subtype adjustment (Table 2), implying that the predictive value of the signature score is largely independent of subtype.


A.
n = 72
B.
n = 39
C.
n = 111
D.
n = 1979
Survival probability
Time (years)
Q1
Q2
Q3
Q4
Q5
Number at risk
15 13 9 8 5
14 13 10 8 4
14 12 10 6 3
14 13 12 8 4
15 11 5 4 3
8 7 5 5 3
8 5 2 2 0
7 4 4 3 2
8 6 4 3 1
8 4 3 2 0
22 19 13 12 7
22 19 13 11 5
22 18 15 10 5
22 17 14 9 4
23 15 9 7 4
396 288 164 84 40
396 266 129 56 26
395 246 101 46 19
396 195 55 20 5
396 148 33 8 5

820

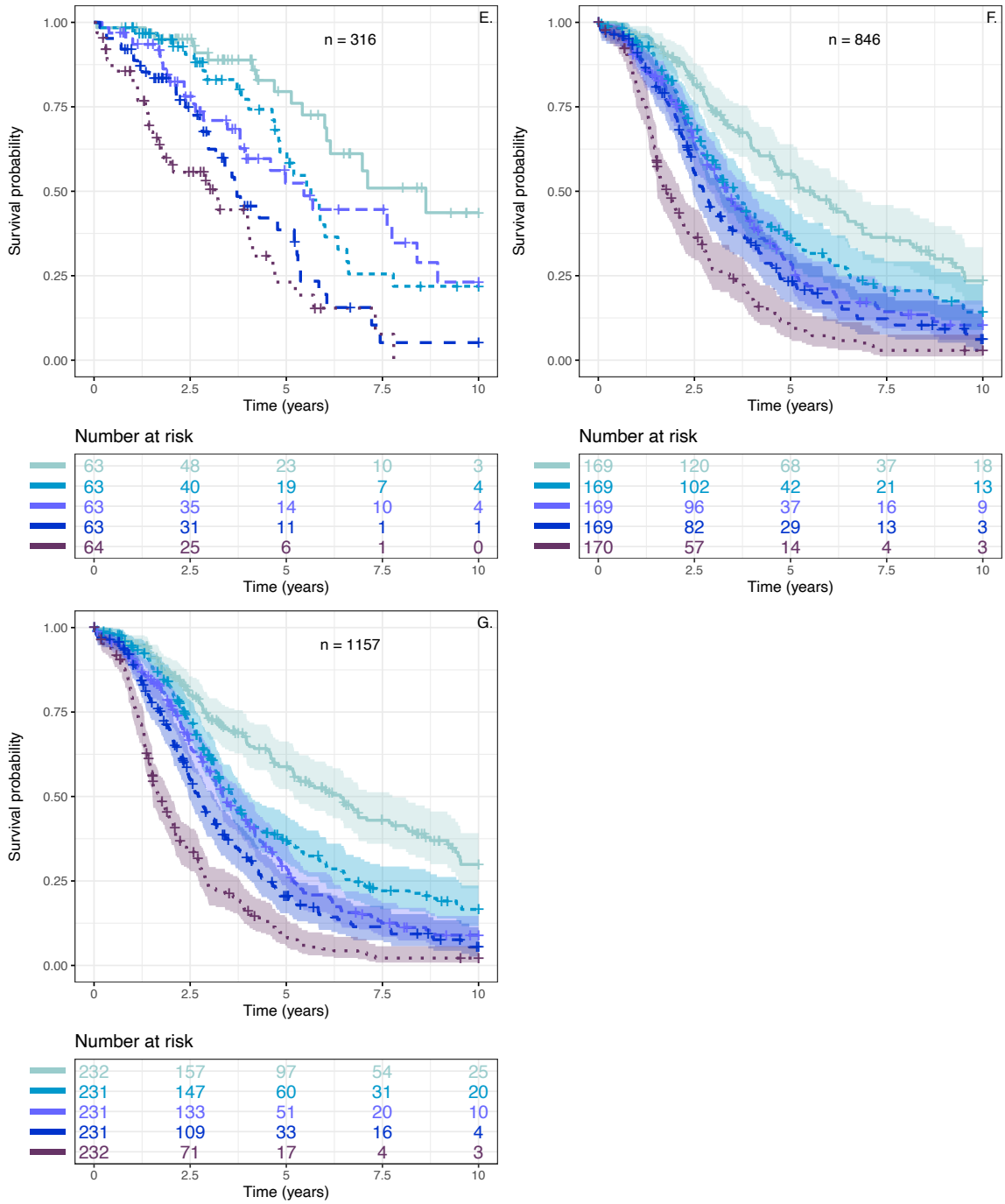


Figure S9. KM curves of overall survival for patients with A) FIGO stage 1A, B) FIGO stage 1B, and C) FIGO stage 1A and 1B, D) FIGO stage 3C, E) FIGO stage 3C patients with no residual disease, F) FIGO stage 3C patients with residual disease, and G) residual disease present, any FIGO stage. Patients were assigned to quintile groups, Q1 to Q5, based on the signature score, with Q1 having the lightest shade and increasingly darker shades corresponding to quintiles with greater scores. Quintiles were calculated independently for each group.

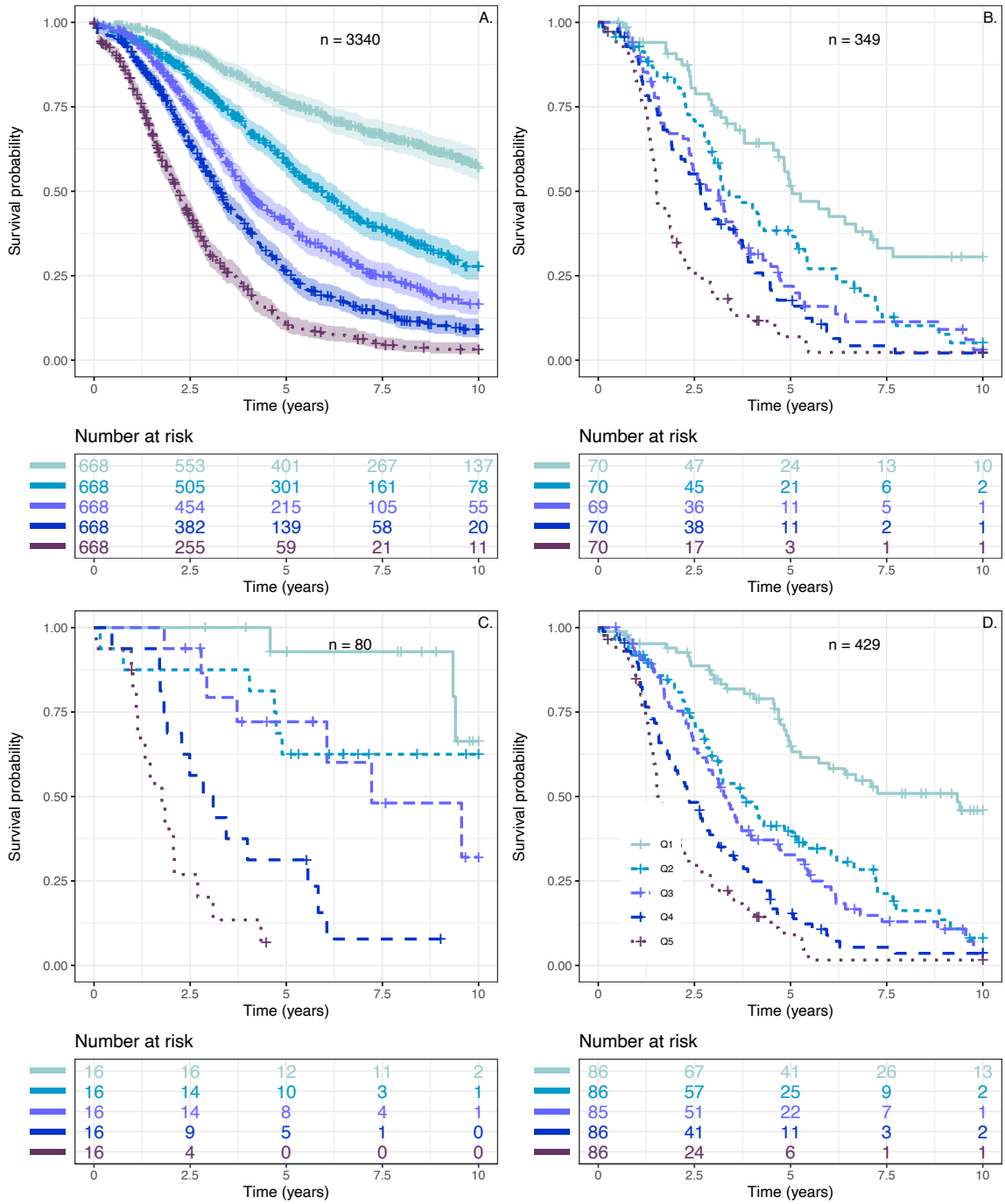


Figure S10. KM curves of overall survival for patients with tissue from A) ovary, B) omentum, C) other extraovarian sites, and D) omentum combined with other extraovarian sites. Patients were assigned to quintile groups, Q1 to Q5, based on the signature score, with Q1 having the lightest shade and increasingly darker shades corresponding to quintiles with greater scores. Quintiles were calculated independently for each of the four tissue groups.

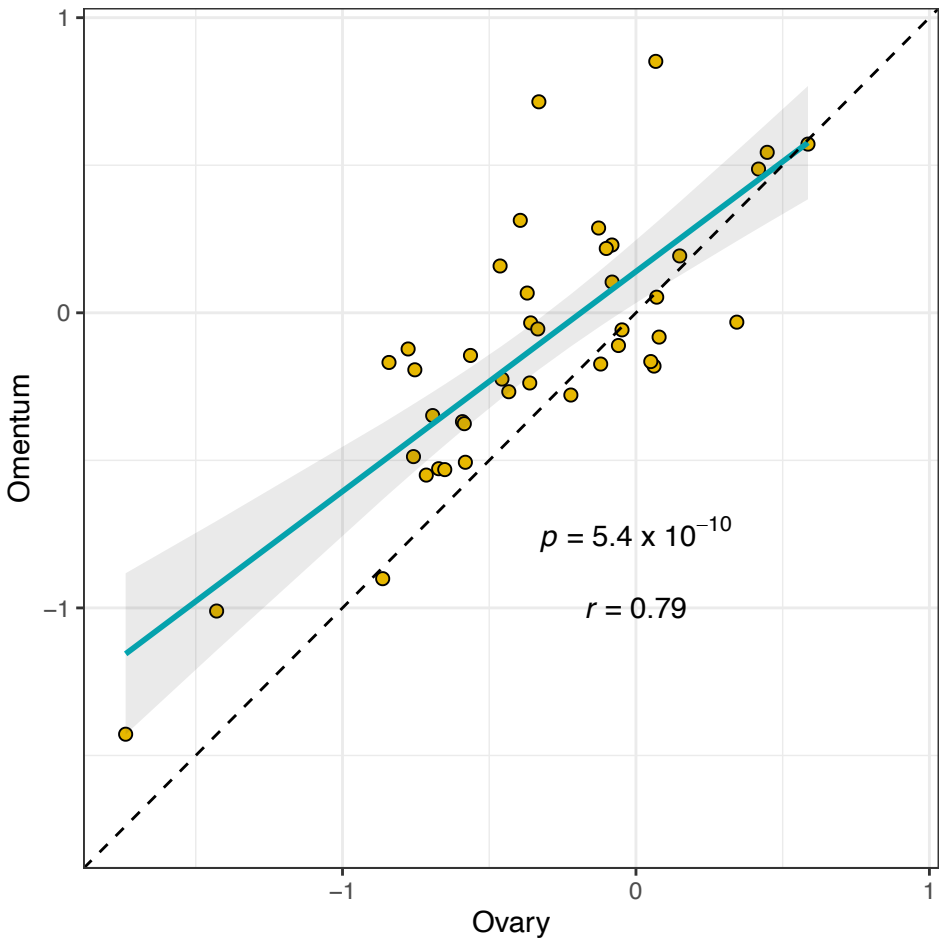


Figure S11. Signature scores in paired omentum vs. ovary tumor tissue samples collected from 42 HGSOC patients. The dashed line represents the line of identity, and the shaded area shows the 95% confidence region around the solid line, which is the least-squares fit to the points. These paired samples were processed and the 101-gene score computed using the Nanostring platform and computations described in Methods. The apparent linear relation observed in the plot demonstrates a strong correspondence between tissue types as suggested by the high Pearson correlation coefficient, r = 0.79 ($p = 5.4 \times 10^{-10}$).

853 **References**

854 1. Division of Biostatistics, Department of Preventive Medicine, Keck School of Medicine. In
855 University of Southern California. Los Angeles: USA.

856 2. School of Women's and Children's Health, Faculty of Medicine. In University of NSW Sydney.
857 Sydney, New South Wales: Australia 2052.

858 3. CRUK Manchester Institute. In The University of Manchester. Manchester: UK.

859 4. Department of Health Science Research, Division of Epidemiology. In Mayo Clinic. Rochester,
860 MN: USA 55905.

861 5. British Columbia's Ovarian Cancer Research (OVCARE) Program. In BC Cancer, Vancouver
862 General Hospital, and University of British Columbia. Vancouver, BC: Canada V5Z 4E6.

863 6. Department of Pathology and Laboratory Medicine. In University of British Columbia.
864 Vancouver, BC: Canada V5Z 4E6.

865 7. Department of Obstetrics and Gynecology. In University of British Columbia. Vancouver, BC:
866 Canada.

867 8. In Peter MacCallum Cancer Center. Melbourne, Victoria: Australia 3000.

868 9. Center for Cancer Prevention and Translational Genomics, Samuel Oschin Comprehensive
869 Cancer Institute. In Cedars-Sinai Medical Center. Los Angeles, CA: USA 90048.

870 10. Department of Pathology, Norris Comprehensive Cancer Center, Keck School of Medicine. In
871 University of Southern California. Los Angeles: USA.

872 11. Anatomical Pathology. In Royal Women's Hospital. Parkville, Victoria: Australia.

873 12. Department of Pathology. In Duke University Hospital. Durham, NC: USA 27710.

874 13. Department of Health Science Research, Division of Biomedical Statistics and Informatics. In
875 Mayo Clinic. Rochester, MN: USA 55905.

876 14. Department of Women's Health. In Tuebingen University Hospital. Tuebingen: Germany
877 72076.

878 15. In The Jackson Laboratory for Genomic Medicine. Farmington, CT: USA CT-06032.

879 16. Department of Genetics and Computational Biology. In QIMR Berghofer Medical Research
880 Institute. Brisbane, Queensland: Australia 4006.

881 17. Centre for Cancer Research, The Westmead Institute for Medical Research. In The University
882 of Sydney. Sydney, New South Wales: Australia 2145.

883 18. Sir Peter MacCallum Department of Oncology. In The University of Melbourne. Parkville,
884 Victoria: Australia 3000.

885 19. Department of Obstetrics and Gynecology, Division of Gynecologic Oncology. In Royal
886 Alexandra Hospital. Edmonton, Alberta: Canada T5H 3V9.

887 20. In Alberta Health Services-Cancer Care. Calgary, AB: Canada.

888 21. Department of Pathology and Laboratory Medicine. In University of Calgary, Foothills Medical
889 Center. Calgary, AB: Canada T2N 2T9.

890 22. Division of Cancer and Ovarian Cancer Action Research Centre, Department Surgery & Cancer.
891 In Imperial College London. London: UK W12 0NN.

892 23. Institute of Cancer Sciences. In University of Glasgow. Glasgow: UK G61 1QH.

893 24. Cancer Research UK Cambridge Institute. In University of Cambridge. Cambridge: UK CB2 0RE.

894 25. Medical Oncology Service. In Hospital Sant Pau. Barcelona: Spain.

895 26. In HM Hospitales Ð Centro Integral Oncol—gico HM Clara Campal. Madrid: Spain.

896 27. Medical Oncology Service. In Hospital Universitario Funcacion Alcorcon. Alcorc—n: Spain.

897 28. Division of Gynecologic Oncology. In NorthShore University HealthSystem, University of
898 Chicago. Evanston, IL: USA 60201.

899 29. Program in Epidemiology, Division of Public Health Sciences. In Fred Hutchinson Cancer
900 Research Center. Seattle, WA: USA 98109.

901 30. Department of Epidemiology. In University of Washington. Seattle, WA: USA 98195.

902 31. Department of Systems Pharmacology and Translational Therapeutics, Perelman School of
903 Medicine. In University of Pennsylvania. Philadelphia: USA PA 19103.

904 32. Department of Genetics. In Geisel School of Medicine at Dartmouth. Hanover, NH: USA.

905 33. Division of Cancer Epidemiology. In German Cancer Research Center (DKFZ). Heidelberg:
906 Germany 69120.

907 34. Department of Pathology, Institute of Pathology. In University Hospital Heidelberg.
908 Heidelberg: Germany 69120.

909 35. National Center for Tumor Diseases. In University Hospital and German Cancer Research
910 Center. Heidelberg: Germany 69120.

911 36. David Geffen School of Medicine, Department of Obstetrics and Gynecology. In University of
912 California at Los Angeles. Los Angeles, CA: USA 90095.

913 37. Women's Cancer Program at the Samuel Oschin Comprehensive Cancer Institute. In Cedars-
914 Sinai Medical Center. Los Angeles, CA: USA 90048.

915 38. Department of Genetics and Pathology. In Pomeranian Medical University. Szczecin: Poland
916 71-252.

917 39. Department of Gynecological Surgery and Gynecological Oncology of Adults and Adolescents.
918 In Pomeranian Medical University. Szczecin: Poland 70-111.

919 40. Department of Epidemiology. In University of Michigan School of Public Health. Ann Arbor,
920 MI: USA 48109.

921 41. Department of Preventive Medicine, Keck School of Medicine. In University of Southern
922 California Norris Comprehensive Cancer Center. Los Angeles, CA: USA 90033.

923 42. Department of Epidemiology and Biostatistics. In Memorial Sloan-Kettering Cancer Center.
924 New York, NY: USA 10065.

925 43. Department of Preventive Medicine, Keck School of Medicine. In University of Southern
926 California. Los Angeles, CA: USA 90033.

927 44. Centre for Cancer Genetic Epidemiology, Department of Oncology. In University of Cambridge.
928 Cambridge: UK CB1 8RN.

929 45. Department of Pathology. In Addenbrooke's Hospital NHS Trust. Cambridge: UK.

930 46. MRC Clinical Trials Unit at UCL, Institute of Clinical Trials & Methodology. In University College
931 London. London: UK WC1V 6LJ.

932 47. Department of Women's Cancer, Institute for Women's Health. In University College London.
933 London: UK W1T 7DN.

934 48. Department of Pathology. In Barts Health National Health Service Trust. London: UK E1 8PR.

935 49. Department of Gynaecological Oncology. In Westmead Hospital. Sydney, New South Wales:
936 Australia 2145.

937 50. Pathology West ICPMR Westmead, Westmead Hospital. In The University of Sydney. Sydney,
938 New South Wales: Australia 2145.

939 51. In University of Western Sydney at Westmead Hospital. Sydney, New South Wales: Australia
940 2145.

941 52. The Crown Princess Mary Cancer Centre Westmead, Sydney-West Cancer Network. In
942 Westmead Hospital. Sydney, New South Wales: Australia 2145.

943 53. Department of Laboratory Medicine and Pathology, Division of Anatomic Pathology. In Mayo
944 Clinic. Rochester, MN: USA 55905.

945 54. Department of Oncology. In Mayo Clinic. Rochester, MN: USA 55905.

946 55. Department of Molecular Pharmacology and Experimental Therapeutics. In Mayo Clinic.
947 Rochester, MN: USA 55905.

948 56. Cancer Epidemiology Program. In University of Hawaii Cancer Center. Honolulu, HI: USA
949 96813.

950 57. Samuel Oschin Comprehensive Cancer Institute, Cancer Prevention and Genetics Program. In
951 Cedars-Sinai Medical Center. Los Angeles, CA: USA 90048.

952 58. John A. Burns School of Medicine, Department of Obstetrics and Gynecology. In University of
953 Hawaii. Honolulu, HI: USA.

954 59. Division of Cancer Epidemiology and Genetics. In National Cancer Institute. Bethesda, MD:
955 USA 20892.

956 60. Department of Cancer Epidemiology and Prevention. In M. Sklodowska-Curie Cancer Center,
957 Oncology Institute. Warsaw: Poland 02-034.

958 61. Department of Health Sciences Research. In Mayo Clinic College of Medicine. Jacksonville, FL:
959 USA 32224.

960 62. Cancer Research UK Clinical Trials Unit, Institute of Cancer Sciences. In University of Glasgow.
961 Glasgow: UK G12 0YN.

962 63. Department of Medical Oncology. In Beatson West of Scotland Cancer Centre and University
963 of Glasgow. Glasgow: UK G12 0YN.

964 64. Gynaecology Unit. In Royal Marsden Hospital. London: UK SW3 6JJ.

965 65. In Centre Leon Berard and University Claude Bernard Lyon 1. Lyon: France 69373.

966 66. Department of Gynecology and Obstetrics. In Ludwig Maximilian University of Munich.
967 Munich: Germany 80336.

968 67. Department of Obstetrics, Gynecology and Women's Health. In University of Minnesota.
969 Minneapolis, MN: USA 55455.

970 68. David Geffen School of Medicine, Department of Medicine Division of Hematology and
971 Oncology. In University of California at Los Angeles. Los Angeles, CA: USA 90095.

972 69. Gynecology Service, Department of Surgery. In Memorial Sloan Kettering Cancer Center. New
973 York, NY: USA 10065.

974 70. Gynecologic Oncology, Laura and Isaac Pearlmutter Cancer Center. In NYU Langone Medical
975 Center. New York, NY: USA 10016.

976 71. Hollings Cancer Center and Department of Public Health Sciences. In Medical University of
977 South Carolina. Charleston, SC: USA 29425.

978 72. In Centro de Investigaci—n en Red de Enfermedades Raras (CIBERER). Madrid: Spain 28029.

979 73. Human Cancer Genetics Programme. In Spanish National Cancer Research Centre (CNIO).
980 Madrid: Spain 28029.

981 74. Cancer Epidemiology Group, University Cancer Center Hamburg (UCCH). In University Medical
982 Center Hamburg-Eppendorf. Hamburg: Germany 20246.

983 75. Department of Public Health Sciences. In University of Virginia. Charlottesville, VA: USA 22908.

984 76. Division of Cancer and Ovarian Cancer Action Research Centre, Department of Surgery and
985 Cancer. In Imperial College London. London: UK W12 0NN.

986 77. Department of Gynecologic Oncology. In Duke University Hospital. Durham, NC: USA 27710.

987 78. Center for Bioinformatics and Functional Genomics and the Cedars Sinai Genomics Core. In
988 Cedars-Sinai Medical Center. Los Angeles, CA: USA 90048.

989 79. In Biomedical Network on Rare Diseases (CIBERER). Madrid: Spain 28029.

990 80. Children's Cancer Institute, Lowy Cancer Research Centre. In University of NSW Sydney.
991 Sydney, New South Wales: Australia 2052.

992 81. Division of Epidemiology, Department of Medicine, Vanderbilt Epidemiology Center,
993 Vanderbilt-Ingram Cancer Center. In Vanderbilt University School of Medicine. Nashville, TN: USA
994 37232.

995 82. Department of Gynecology and Obstetrics, Comprehensive Cancer Center ER-EMN. In
996 University Hospital Erlangen, Friedrich-Alexander-University Erlangen-Nuremberg. Erlangen: Germany
997 91054.

998 83. Department of Molecular Oncology. In BC Cancer Research Centre. Vancouver, BC: Canada
999 V5Z 4E6.

1000 84. Huntsman Cancer Institute, Department of Population Health Sciences. In University of Utah.
1001 Salt Lake City, UT: USA 84112.

1002 85. Centre for Cancer Genetic Epidemiology, Department of Public Health and Primary Care. In
1003 University of Cambridge. Cambridge: UK CB1 8RN.

1004 86. In The Kinghorn Cancer Centre, Garvan Institute of Medical Research. Sydney, New South
1005 Wales: Australia 2010.

1006 87. Adult Cancer Program, Lowy Cancer Research Centre. In University of NSW Sydney. Sydney,
1007 New South Wales: Australia 2052.

1008 88. Vaughan S, Coward JI, Bast RC, Jr. et al. Rethinking ovarian cancer: recommendations for
1009 improving outcomes. Nat Rev Cancer 2011; 11: 719-725.

1010 89. Torre LA, Trabert B, DeSantis CE et al. Ovarian cancer statistics, 2018. CA Cancer J Clin 2018;
1011 68: 284-296.

1012 90. Bowtell DD, Bohm S, Ahmed AA et al. Rethinking ovarian cancer II: reducing mortality from
1013 high-grade serous ovarian cancer. Nat Rev Cancer 2015; 15: 668-679.

1014 91. du Bois A, Reuss A, Pujade-Lauraine E et al. Role of surgical outcome as prognostic factor in
1015 advanced epithelial ovarian cancer: a combined exploratory analysis of 3 prospectively randomized
1016 phase 3 multicenter trials: by the Arbeitsgemeinschaft Gynaekologische Onkologie Studiengruppe
1017 Ovarialkarzinom (AGO-OVAR) and the Groupe d'Investigateurs Nationaux Pour les Etudes des Cancers
1018 de l'Ovaire (GINECO). Cancer 2009; 115: 1234-1244.

1019 92. Bolton KL, Chenevix-Trench G, Goh C et al. Association between BRCA1 and BRCA2 mutations
1020 and survival in women with invasive epithelial ovarian cancer. JAMA 2012; 307: 382-390.

1021 93. Candido-dos-Reis FJ, Song H, Goode EL et al. Germline mutation in BRCA1 or BRCA2 and ten-
1022 year survival for women diagnosed with epithelial ovarian cancer. Clin Cancer Res 2015; 21: 652-657.

1023 94. Goode EL, Block MS, Kalli KR et al. Dose-Response Association of CD8+ Tumor-Infiltrating
1024 Lymphocytes and Survival Time in High-Grade Serous Ovarian Cancer. JAMA Oncol 2017; 3: e173290.

1025 95. Zhang L, Conejo-Garcia JR, Katsaros D et al. Intratumoral T cells, recurrence, and survival in
1026 epithelial ovarian cancer. N Engl J Med 2003; 348: 203-213.

1027 96. Pujade-Lauraine E, Ledermann JA, Selle F et al. Olaparib tablets as maintenance therapy in
1028 patients with platinum-sensitive, relapsed ovarian cancer and a BRCA1/2 mutation (SOLO2/ENGOT-
1029 Ov21): a double-blind, randomised, placebo-controlled, phase 3 trial. Lancet Oncol 2017; 18: 1274-
1030 1284.

1031 97. Moore K, Colombo N, Scambia G et al. Maintenance Olaparib in Patients with Newly
1032 Diagnosed Advanced Ovarian Cancer. N Engl J Med 2018; 379: 2495-2505.

1033 98. Tothill RW, Tinker AV, George J et al. Novel molecular subtypes of serous and endometrioid
1034 ovarian cancer linked to clinical outcome. Clin Cancer Res 2008; 14: 5198-5208.

1035 99. Cancer Genome Atlas Research N. Integrated genomic analyses of ovarian carcinoma. Nature
1036 2011; 474: 609-615.

1037 100. Bonome T, Levine DA, Shih J et al. A gene signature predicting for survival in suboptimally
1038 debulked patients with ovarian cancer. Cancer Res 2008; 68: 5478-5486.

1039 101. Karlan BY, Dering J, Walsh C et al. POSTN/TGFBI-associated stromal signature predicts poor
1040 prognosis in serous epithelial ovarian cancer. Gynecol Oncol 2014; 132: 334-342.

1041 102. Konecny GE, Haluska P, Janicke F et al. A phase II, multicenter, randomized, double-blind,
1042 placebo-controlled trial of ganitumab or placebo in combination with carboplatin/paclitaxel as front-
1043 line therapy for optimally debulked primary ovarian cancer: The TRIO14 trial. Journal of Clinical
1044 Oncology 2014; 32: 5529.

1045 103. Konecny GE, Wang C, Hamidi H et al. Prognostic and therapeutic relevance of molecular
1046 subtypes in high-grade serous ovarian cancer. J Natl Cancer Inst 2014; 106.

1047 104. Millstein J, Volfson D. Computationally efficient permutation-based confidence interval
1048 estimation for tail-area FDR. Front Genet 2013; 4: 179.

1049 105. Talhouk A, Kommoss S, Mackenzie R et al. Single-Patient Molecular Testing with NanoString
1050 nCounter Data Using a Reference-Based Strategy for Batch Effect Correction. PLoS One 2016; 11:
1051 e0153844.

1052 106. Edgar R, Domrachev M, Lash AE. Gene Expression Omnibus: NCBI gene expression and
1053 hybridization array data repository. Nucleic Acids Res 2002; 30: 207-210.

1054 107. Jin C, Xue Y, Li Y et al. A 2-Protein Signature Predicting Clinical Outcome in High-Grade Serous
1055 Ovarian Cancer. Int J Gynecol Cancer 2018; 28: 51-58.

1056 108. Mankoo PK, Shen R, Schultz N et al. Time to recurrence and survival in serous ovarian tumors
1057 predicted from integrated genomic profiles. PLoS One 2011; 6: e24709.

1058 109. Nymoen DA, Hetland Falkenthal TE, Holth A et al. Expression and clinical role of
1059 chemoresponse-associated genes in ovarian serous carcinoma. Gynecol Oncol 2015; 139: 30-39.

1060 110. Jimenez-Sanchez A, Memon D, Pourpe S et al. Heterogeneous Tumor-Immune
1061 Microenvironments among Differentially Growing Metastases in an Ovarian Cancer Patient. Cell 2017;
1062 170: 927-938.e920.

1063 111. Wang C, Cicek MS, Charbonneau B et al. Tumor hypomethylation at 6p21.3 associates with
1064 longer time to recurrence of high-grade serous epithelial ovarian cancer. Cancer Res 2014; 74: 3084-
1065 3091.

1066 112. Gorbachev AV, Kobayashi H, Kudo D et al. CXC chemokine ligand 9/monokine induced by IFN-
1067 gamma production by tumor cells is critical for T cell-mediated suppression of cutaneous tumors. J
1068 Immunol 2007; 178: 2278-2286.

1069 113. Bronger H, Singer J, Windmuller C et al. CXCL9 and CXCL10 predict survival and are regulated
1070 by cyclooxygenase inhibition in advanced serous ovarian cancer. Br J Cancer 2016; 115: 553-563.

1071 114. Dose-Response Association of CD8+ Tumor-Infiltrating Lymphocytes and Survival Time in High-
1072 Grade Serous Ovarian Cancer. JAMA Oncol 2017; 3: e173290.

1073 115. Tokunaga R, Zhang W, Naseem M et al. CXCL9, CXCL10, CXCL11/CXCR3 axis for immune
1074 activation - A target for novel cancer therapy. Cancer Treat Rev 2018; 63: 40-47.

1075 116. Xiao P, Guo Y, Zhang H et al. Myeloid-restricted ablation of Shp2 restrains melanoma growth
1076 by amplifying the reciprocal promotion of CXCL9 and IFN-gamma production in tumor
1077 microenvironment. Oncogene 2018.

1078 117. Zhang R, Tian L, Chen LJ et al. Combination of MIG (CXCL9) chemokine gene therapy with low-
1079 dose cisplatin improves therapeutic efficacy against murine carcinoma. Gene Ther 2006; 13: 1263-
1080 1271.

1081 118. Zhang W, Ota T, Shridhar V et al. Network-based survival analysis reveals subnetwork
1082 signatures for predicting outcomes of ovarian cancer treatment. PLoS Comput Biol 2013; 9:
1083 e1002975.

1084 119. Reinartz S, Finkernagel F, Adhikary T et al. A transcriptome-based global map of signaling
1085 pathways in the ovarian cancer microenvironment associated with clinical outcome. Genome Biol
1086 2016; 17: 108.

1087 120. Verhaak RG, Tamayo P, Yang JY et al. Prognostically relevant gene signatures of high-grade
1088 serous ovarian carcinoma. J Clin Invest 2013; 123: 517-525.

1089 121. Rudd J, Zelaya RA, Demidenko E et al. Leveraging global gene expression patterns to predict
1090 expression of unmeasured genes. BMC Genomics 2015; 16: 1065.

1091 122. Yoshihara K, Tsunoda T, Shigemizu D et al. High-risk ovarian cancer based on 126-gene
1092 expression signature is uniquely characterized by downregulation of antigen presentation pathway.
1093 Clin Cancer Res 2012; 18: 1374-1385.

1094 123. Wallden B, Storhoff J, Nielsen T et al. Development and verification of the PAM50-based
1095 Prosigna breast cancer gene signature assay. BMC Med Genomics 2015; 8: 54.

1096 124. Heagerty PJ, Lumley T, Pepe MS. Time-dependent ROC curves for censored survival data and a
1097 diagnostic marker. Biometrics 2000; 56: 337-344.

1098 125. Simon N, Friedman J, Hastie T, Tibshirani R. Regularization Paths for Cox's Proportional
1099 Hazards Model via Coordinate Descent. J Stat Softw 2011; 39: 1-13.

1100 126. Hofner B, Mayr A, Robinzonov N, Schmid M. Model-based boosting in R: a hands-on tutorial
1101 using the R package mboost. Computational Statistics 2012; 29: 3-35.

1102 127. Ishwaran H, Kogalur UB, Blackstone UH, Lauer MS. Random survival forests. Ann. Appl. Stat
1103 2008; 2: 841-860.

1104 128. Fabregat A, Jupe S, Matthews L et al. The Reactome Pathway Knowledgebase. Nucleic Acids
1105 Res 2018; 46: D649-D655.

1106 129. Subramanian A, Tamayo P, Mootha VK et al. Gene set enrichment analysis: a knowledge-
1107 based approach for interpreting genome-wide expression profiles. Proc Natl Acad Sci U S A 2005; 102:
1108 15545-15550.

1109