Gene expression profiling of mucinous ovarian tumors and comparison with upper and lower gastrointestinal tumors identifies markers associated with adverse outcomes.

Gene expression profiling of mucinous ovarian tumors and comparison with upper and lower gastrointestinal tumors identifies markers associated with adverse outcomes.

Nicola S. Meagher1, 2, Kylie L. Gorringe3,4, Matthew Wakefield5-7, Adelyn Bolithon1, 2, Chi Nam Ignatius Pang8, 9, Derek S. Chiu10, Michael S. Anglesio10, 11, Kylie-Ann Mallitt1, 12, Jennifer A. Doherty13, Holly R. Harris14, 15, Joellen M. Schildkraut16, Andrew Berchuck17, Kara L. Cushing-Haugen14, Ksenia Chezar18, Angela Chou19, 20, 21, Adeline Tan22, 23, Jennifer Alsop24, Ellen Barlow25, Matthias W. Beckmann26, Jessica Boros21,27, 28; David D.L. Bowtell3, 4; for the AOCS Group, Alison H. Brand21, 28, James D. Brenton29, Ian Campbell3, 4, Dane Cheasley3, 4, Joshua Cohen30, Cezary Cybulski31, Esther Elishaev32, Ramona Erber33, Rhonda Farrell21, 34, Anna Fischer35, Zhuxuan Fu36, Blake Gilks37; Anthony J. Gill19, 20, 21; for the Australian Pancreatic Genome Initiative, Charlie Gourley38, Marcel Grube39, Paul R. Harnett21, 40, Arndt Hartmann35, Anusha Hettiaratchi41, Claus K. Høgdall42, Tomasz Huzarski31, 43 Anna Jakubowska31, 44, Mercedes Jimenez-Linan45, Catherine J. Kennedy21, 27, 28, Byoung-Gie Kim46, Jae-Weon Kim47, Jae-Hoon Kim48, Kayla Klett49, Jennifer M. Koziak50, Tiffany Lai30, Angela Laslavic51, Jenny Lester30, Yee Leung22, 52, 53, Na Li3, 54, Winston Liauw1, 55, Belle W.X. Lim3, Anna Linder56, Jan Lubiński31, Sakshi Mahale3, Constantina Mateoiu57, Simone McInerny3, 54, Janusz Menkiszak58, Parham Minoo18, Suzana Mittelstadt39, David Morris59, Sandra Orsulic30, Sang-Yoon Park60, Celeste Leigh Pearce61, 62, John V Pearson63, Malcolm C. Pike62, 64, Carmel M. Quinn41, Ganendra Raj Mohan52, 65,66 Jianyu Rao67, Marjorie J. Riggan17, Matthias Ruebner26, Stuart Salfinger65, Clare L Scott4-7, Mitul Shah24, Helen Steed68, 69, Colin J.R. Stewart70, Deepak Subramanian3, Soseul Sung71-73, Katrina Tang74, Paul Timpson19, Robyn L Ward21, Rebekka Wiedenhoefer35; Heather Thorne3; for the kConFab Investigators, Paul A. Cohen22, 75, Philip Crowe1, 76, Peter A. Fasching26, Jacek Gronwald31, Nicholas J. Hawkins1, Estrid Høgdall77, David G. Huntsman11, 78, Paul A. James3, 54 Beth Y. Karlan30, Linda E. Kelemen79, Stefan Kommoss39, Gottfried E. Konecny30, Francesmary Modugno36, 49, 51, Sue K. Park72, 73, 80, Annette Staebler35, Karin Sundfeldt56, Anna H. Wu62, Aline Talhouk10, 11, 37, Paul D.P. Pharoah24, 81, Lyndal Anderson21, 82, Anna DeFazio21, 27, 28, 83, Martin Köbel18, Michael L. Friedlander1, 25, 84, Susan J. Ramus1, 2

1 School of Clinical Medicine, Faculty of Medicine and Health, University of NSW Sydney, Sydney, New South Wales, Australia.
2 Adult Cancer Program, Lowy Cancer Research Centre, University of NSW Sydney, Sydney, New South Wales, Australia.
3 Peter MacCallum Cancer Centre, Melbourne, Victoria, Australia.

4 Sir Peter MacCallum Department of Medical Oncology, The University of Melbourne, Parkville, Victoria, Australia.
5 The Walter and Eliza Hall Institute of Medical Research, Parkville, Victoria, Australia.
6 Department of Medical Biology, The University of Melbourne, Melbourne, Victoria, Australia.
7 Department of Obstetrics and Gynaecology, The University of Melbourne, Melbourne, Victoria, Australia.
8 School of Biotechnology and Biomolecular Sciences, The University of New South Wales, Sydney, New South Wales, Australia.
9 Bioinformatics Unit, Children's Medical Research Institute, Westmead, Sydney, Australia.
10 British Columbia's Gynecological Cancer Research Team (OVCARE), University of British Columbia, BC Cancer, and Vancouver General Hospital, Vancouver, BC, Canada.
11 Department of Obstetrics and Gynecology, University of British Columbia, Vancouver, BC, Canada.
12 Centre for Big Data Research in Health, University of New South Wales Sydney, Sydney, New South Wales, Australia.
13 Huntsman Cancer Institute, Department of Population Health Sciences, University of Utah, Salt Lake City, UT, USA.
14 Program in Epidemiology, Division of Public Health Sciences, Fred Hutchinson Cancer Center, Seattle, WA, USA.
15 Department of Epidemiology, University of Washington, Seattle, WA, USA.
16 Department of Epidemiology, Rollins School of Public Health, Emory University, Atlanta, GA, USA.
17 Department of Obstetrics and Gynecology, Division of Gynecologic Oncology, Duke University Medical Center, Durham, NC, USA.
18 Department of Pathology and Laboratory Medicine, University of Calgary, Foothills Medical Center, Calgary, AB, Canada.
19 The Kinghorn Cancer Centre, Garvan Institute of Medical Research, Sydney, New South Wales, Australia.
20 Department of Anatomical Pathology, Royal North Shore Hospital, Sydney, New South Wales, Australia.
21 The University of Sydney, Sydney, New South Wales, Australia.
22 Division of Obstetrics and Gynaecology, Medical School, University of Western Australia, Crawley, Western Australia, Australia.
23 Western Women's Pathology, Western Diagnostic Pathology, Wembley, Australia.
24 Centre for Cancer Genetic Epidemiology, Department of Oncology, University of Cambridge, Cambridge, UK.
25 Gynaecological Cancer Centre, Royal Hospital for Women, Sydney, New South Wales, Australia.
26 Department of Gynecology and Obstetrics, Comprehensive Cancer Center Erlangen-EMN, Friedrich-Alexander University Erlangen-Nuremberg, University Hospital Erlangen, Erlangen, Germany.
27 Centre for Cancer Research, The Westmead Institute for Medical Research, Sydney, New South Wales, Australia.
28 Department of Gynaecological Oncology, Westmead Hospital, Sydney, New South Wales, Australia.
29 Cancer Research UK Cambridge Institute, University of Cambridge, Cambridge, UK.
30 David Geffen School of Medicine, Department of Obstetrics and Gynecology, University of California at Los Angeles, Los Angeles, CA, USA.
31 Department of Genetics and Pathology, International Hereditary Cancer Center, Pomeranian Medical University, Szczecin, Poland.
32 Department of Pathology, University of Pittsburgh School of Medicine, Pittsburgh, PA, USA.
33 Institute of Pathology, Comprehensive Cancer Center Erlangen-EMN, Friedrich-Alexander Universität Erlangen-Nürnberg, University Hospital Erlangen, Erlangen, Germany.

34 Prince of Wales Private Hospital, Randwick, New South Wales, Australia.
35 Institute of Pathology and Neuropathology, Tübingen University Hospital, Tübingen, Germany.
36 Department of Epidemiology, University of Pittsburgh School of Public Health, Pittsburgh, PA, USA.
37 Department of Pathology and Laboratory Medicine, University of British Columbia, Vancouver, BC, Canada.
38 Nicola Murray Centre for Ovarian Cancer Research, Cancer Research UK Scotland Centre, University of Edinburgh, Edinburgh, UK.
39 Department of Women's Health, Tübingen University Hospital, Tübingen, Germany.
40 Crown Princess Mary Cancer Centre, Westmead Hospital, Sydney, New South Wales, Australia.
41 The Health Precincts Biobank (formerly the Health Science Alliance Biobank), UNSW Biospecimen Services, Mark Wainwright Analytical Centre, University of New South Wales Sydney, Sydney, New South Wales, Australia.
42 Department of Gynaecology, Rigshospitalet, University of Copenhagen, Copenhagen, Denmark.
43 Department of Genetics and Pathology, University of Zielona Góra, Zielona Góra, Poland.
44 Independent Laboratory of Molecular Biology and Genetic Diagnostics, Pomeranian Medical University, Szczecin, Poland.
45 Department of Histopathology, Addenbrooke's Hospital, Cambridge, UK.
46 Department of Obstetrics and Gynecology, Samsung Medical Center, Sungkyunkwan University School of Medicine, Seoul, Korea.
47 Department of Obstetrics and Gynecology, Seoul National University College of Medicine, Seoul, Korea.
48 Department of Obstetrics and Gynecology, Gangnam Severance Hospital, Yonsei University College of Medicine, Seoul, Republic of Korea.
49 Women's Cancer Research Center, Magee-Womens Research Institute and Hillman Cancer Center, Pittsburgh, PA, USA.
50 Alberta Health Services-Cancer Care, Calgary, AB, Canada.
51 Division of Gynecologic Oncology, Department of Obstetrics, Gynecology and Reproductive Sciences, University of Pittsburgh School of Medicine, Pittsburgh, PA, USA.
52 Department of Gynaecological Oncology, King Edward Memorial Hospital, Subiaco, Western Australia, Australia.
53 Australia New Zealand Gynaecological Oncology Group, Camperdown, Australia.
54 Parkville Familial Cancer Centre, The Royal Melbourne Hospital and Peter MacCallum Cancer Centre, Melbourne, Victoria, Australia
55 Cancer Care Centre, St George Hospital, Sydney, New South Wales, Australia.
56 Department of Obstetrics and Gynecology, Inst of Clinical Science, Sahlgrenska Center for Cancer Research, University of Gothenburg, Gothenburg, Sweden.
57 Department of Pathology, Sahlgrenska University Hospital, Gothenburg, Sweden.
58 Department of Gynecological Surgery and Gynecological Oncology of Adults and Adolescents, Pomeranian Medical University, Szczecin, Poland.
59 St George and Sutherland Clinical School, University of New South Wales Sydney, Sydney, New South Wales, Australia.
60 Center for Gynecologic Cancer, National Cancer Center Institute for Cancer Control, Goyang, Republic of Korea.
61 Department of Epidemiology, University of Michigan School of Public Health, Ann Arbor, MI, USA.
62 Department of Preventive Medicine, Keck School of Medicine, University of Southern California Norris Comprehensive Cancer Center, Los Angeles, CA, USA.
63 QIMR Berghofer Medical Research Institute, Brisbane, Queensland, Australia.
64 Department of Epidemiology and Biostatistics, Memorial Sloan-Kettering Cancer Center, New York, NY, USA.

65 Department of Gynaecological Oncology, St John of God Subiaco Hospital, Subiaco, Western Australia, Australia.
66 School of Medicine, University of Notre Dame, Fremantle, Western Australia, Australia.
67 Department of Pathology and Laboratory Medicine, David Geffen School of Medicine, University of California and Los Angeles, Los Angeles, CA, USA.
68 Division of Gynecologic Oncology, Department of Obstetrics and Gynecology, University of Alberta, Edmonton, Alberta, Canada.
69 Section of Gynecologic Oncology Surgery, North Zone, Alberta Health Services, Edmonton, Alberta, Canada.
70 School for Women's and Infants' Health, University of Western Australia, Perth, Australia.
71 Department of Biomedical Sciences, Seoul National University Graduate School, Seoul, Korea.
72 Cancer Research Institute, Seoul National University, Seoul, Korea.
73 Department of Preventive Medicine, Seoul National University College of Medicine, Seoul, Korea.
74 Department of Anatomical Pathology, Prince of Wales Hospital, Sydney, New South Wales, Australia.
75 Department of Gynaecological Oncology, St John of God Subiaco Hospital, Subiaco, Western Australia, Australia.
76 Department of Surgery, Prince of Wales Private Hospital, Randwick, New South Wales, Australia.
77 Department of Pathology, Herlev Hospital, University of Copenhagen, Copenhagen, Denmark.
78 Department of Molecular Oncology, BC Cancer Research Centre, Vancouver, BC, Canada.
79 Hollings Cancer Center, Medical University of South Carolina, Charleston, SC, USA.
80 Integrated Major in Innovative Medical Science, Seoul National University College of Medicine, Seoul, South Korea.
81 Centre for Cancer Genetic Epidemiology, Department of Public Health and Primary Care, University of Cambridge, Cambridge, UK.
82 Department of Tissue Pathology and Diagnostic Oncology, Royal Prince Alfred Hospital and NSW Health Pathology, Sydney, New South Wales, Australia.
83 The Daffodil Centre, a joint venture with Cancer Council NSW, The University of Sydney, Sydney, New South Wales, Australia.
84 Nelune Comprehensive Cancer Centre, Prince of Wales Hospital, Sydney, New South Wales, Australia.

Corresponding authors:

Nicola S Meagher nicola.meagher@sydney.edu.au

Level 2, Lowy Cancer Research Centre, UNSW Sydney NSW 2052 Australia

Professor Susan Ramus s.ramus@unsw.edu.au

+61 9385 1720 Level 2, Lowy Cancer Research Centre, UNSW Sydney NSW 2052 Australia

Running title (60 characters): Prognostic Features of Mucinous Ovarian Carcinoma.

Conflicts of interest statement

NSM has received a travel grant from NanoString technologies, unrelated to this work, and the company were not involved in the design or interpretation of this work. CG has received

grants and personal fees outside of the reported work from AstraZeneca, MSD, GSK, Clovis, and Nucana; grants outside of the reported work from Aprea, Novartis, Bergenbio, and Medannexin; personal fees outside of the reported work from Chugai, Cor2Ed, and Takeda. AH had an advisory role and received honoraria from BMS, MSD, Roche, Cepheid, Qiagen, Agilent, Diaceutics, Lilly, AstraZeneca, Boehringer Ingelheim, Abbvie, Jansen-Cilag, Pfizer, and Ipsen. RE has received honoraria from Roche, Eisai, Pfizer, BioNTech, Veracyte (PROCURE), Diaceutics, and Novartis. The institution of AH and RE conducts research for AstraZeneca, Roche, Janssen-Cilag, NanoString Technologies, Biocartis, ZytoVision, Novartis, Cepheid, and BioNTech. MLF has participated in Advisory Boards/Consulting from Astra Zeneca, Novartis, GSK, Takeda, Lilly, MSD Eisei and received honoraria/speakers' fees from AstraZeneca, GSK, ACT-Genomics; Research funding to institution: from AstraZeneca, Novartis, Beigene. ADeF has received funding from AstraZeneca, unrelated this work. The other authors declare no potential conflicts of interest.

ABSTRACT

Purpose

Advanced stage MOC have poor chemotherapy response and prognosis and lack biomarkers to aid Stage I adjuvant treatment. Differentiating primary mucinous ovarian carcinoma (MOC) from gastrointestinal (GI) metastases to the ovary is also challenging due to phenotypic similarities. Clinicopathological and gene expression data were analysed to identify prognostic and diagnostic features.

Experimental Design

Discovery analyses selected 19 genes with prognostic/diagnostic potential. Validation was performed through the Ovarian Tumor Tissue Analysis consortium and GI cancer biobanks comprising 604 patients with MOC (n=333), mucinous borderline ovarian tumors (MBOT, n=151), upper GI (n=65), and lower GI tumors (n=55).

Results

Infiltrative pattern of invasion was associated with decreased overall survival (OS) within 2-years from diagnosis, compared with expansile pattern in Stage I MOC (hazard ratio HR 2.77 (1.04-7.41, p=0.042). Increased expression of *THBS2* and *TAGLN* were associated with shorter OS in MOC patients, (HR 1.25 (95% CI 1.04-1.51, p=0.016)) and (1.21 (1.01-1.45, p=0.043)) respectively. *ERBB2* (HER2)-amplification or high mRNA expression was evident in 64/243 (26%) of MOCs, but only 8/243 (3%) were also infiltrative (4/39, 10%) or Stage III/IV (4/31, 13%).

Conclusion

An infiltrative growth pattern infers poor prognosis within 2-years from diagnosis and may help select Stage I patients for adjuvant therapy. High expression of *THBS2* and *TAGLN* in

MOC confer an adverse prognosis and is upregulated in the infiltrative subtype which warrants further investigation. Anti-HER2 therapy should be investigated in a subset of patients. MOC samples clustered with upper GI, yet markers to differentiate these entities remain elusive, suggesting similar underlying biology and shared treatment strategies.

Statement of translational relevance

Mucinous ovarian cancer (MOC) is a rare histological subtype of epithelial ovarian cancer, lacking prognostic markers in Stage I tumors, with poor prognosis and low response to chemotherapy at advanced stage. Phenotypic similarities between MOC and lower and upper gastrointestinal (GI) tumors create diagnostic challenges when they spread to the ovary. In the largest series to-date of Stage I MOC characterised pathologically by pattern of invasion, we confirm that an infiltrative pattern is a poor prognostic factor, supporting consideration of adjuvant chemotherapy. We identified two prognostic markers *THBS2* and *TAGLN* in MOC worthy of further investigation. Despite higher frequency of HER2+ in low stage and expansile pattern MOC, just 3% of patients with HER2+ MOC are of poor prognosis (advanced stage or infiltrative) and should be considered for anti-HER2 therapy. Comparisons with GI cancers at the mRNA expression level concludes that distinction from pancreatic and gastric cancers remains a challenge.

INTRODUCTION

Mucinous ovarian carcinoma (MOC) is a rare histological type that is less well characterized compared to more common ovarian cancer histotypes. A clinical problem frequently encountered in patients diagnosed with advanced stage MOC is the uncertainty as to whether the primary cancer is ovarian or metastatic from other sites. Metastases typically originate from the gastrointestinal (GI) tract, and the primary tumor may not be evident at surgery or on imaging.(1-3) Earlier literature has focused on differentiating MOC from lower GI tumors, due to the relatively high frequency of reclassification from “primary MOC” to primary colorectal or appendiceal neoplasms metastatic to the ovary following expert pathological review. (1) Gene and protein expression studies have led to improved diagnostic algorithms

for lower GI tumors (4) but robust markers to differentiate primary MOC from metastases of upper GI origin are lacking.(5)

Patients with MOC diagnosed at advanced stage (International Federation of Gynecology and Obstetrics (FIGO), Stage III/IV) have very poor survival (5-year survival 15%). (6) Treatment guidelines for FIGO Stage IC – IV MOC are primary cytoreductive surgery and adjuvant chemotherapy with carboplatin and paclitaxel (+/- bevacizumab), similar to the treatment of patients with more common ovarian cancer histotypes.(7) However, given the poor outcomes of patients with advanced stage MOC there is a great need for more effective treatment strategies. This has proven to be difficult due to the rarity of MOC, and difficulties in making a definitive diagnosis based on routine histopathology. The only randomized trial designed to compare carboplatin and paclitaxel (+/- bevacizumab) with a gastrointestinal chemotherapy regimen, capecitabine and oxaliplatin (+/- bevacizumab) for MOC was closed prematurely.(8) The major obstacles were a limited number of sites participating due to the cost of opening trials with low accrual, and the high frequency of misclassified GI metastases on central pathology review.(8) The United States National Comprehensive Cancer Network (US NCCN) guidelines now recommend either ovarian or GI regimens for patients with MOC based on expert opinion and small retrospective series but the evidence base is low.(9-11) A better understanding of the molecular differences and similarities between MOC and mucinous carcinomas arising in the GI tract is needed. This could guide treatment recommendations and inform the design of future basket clinical trials that include advanced stage mucinous cancers irrespective of site of origin.

For most patients diagnosed with Stage I MOC (~70-80% of all MOC) prognosis is good, however the clinical challenge is identifying the subset of patients with a higher mortality risk. Notwithstanding the limited evidence for efficacy, the US NCCN guidelines recommend adjuvant chemotherapy for MOC FIGO Stage IC or higher(9), while the European guidelines include consideration of adjuvant chemotherapy for patients with FIGO Stage IA or IB MOC with an infiltrative growth pattern.(12) This pathological feature exhibits destructive invasion of haphazardly arranged and angulated tumor cell nests into a desmoplastic stroma,(13) and has been suggested to confer an increased risk of relapse and mortality. This contrasts with expansile invasion characterized by complex tumor nodules with confluent epithelial growth.(14) Published series to-date in Stage I MOC have reported inconsistent results and are limited by small sample sizes (n=21 to 64).(15-20) Determining the role of pattern of invasion in a large Stage I MOC cohort is needed to help inform treatment recommendations if higher risk of recurrence is confirmed.

We analysed clinical, pathological and gene expression data, in tumor samples from a large cohort of patients with MOC. We aimed to identify new prognostic biomarkers, as well as validate the prognostic association between pattern of invasion and survival in a well-powered, adjusted analysis. We also aimed to differentiate MOC from primary and metastatic GI cancers based on mRNA expression of key genes, or to identify shared markers that may help select targeted therapeutic options independent of site of origin.

MATERIALS AND METHODS

Patient cohort

Samples and data were submitted from 848 patients diagnosed with ovarian or GI tumors. These were from 24 sites from the Ovarian Tumor Tissue Analysis (OTTA) consortium, the Australian Pancreatic Genome Initiative, the Molecular and Cellular Oncology colorectal biobank (UNSW) and the Department of Pathology, University of Calgary. Clinical data including patient age at diagnosis, tumor stage, histopathological grade and overall survival (OS) were provided by the respective studies. The study was approved by the UNSW Human Research Ethics Committee (approval HC17182), all contributing sites obtained written informed patient consent or had relevant ethical/institutional review board approval for waiver of consent, and all studies were conducted in accordance with recognized ethical guidelines (Supplementary Table S1).

Hematoxylin and eosin (H&E) stained slides were reviewed to confirm diagnosis, identify the anatomical site of the tissue sample used in this study, mark the region for RNA extraction, estimate the percentage of tumor cells within the extraction area, and to classify the pattern of invasion. Centralized pathology review was performed by expert gynecological or gastrointestinal pathologists (MK, LA, AT, NH, AC). An infiltrative pattern of invasion in MOC was classified with linear extent of stromal invasion >5 mm. (21) Samples from 178 patients were excluded (Supplementary Figure 1) due to; low (<20%) tumor cellularity (n=52), ineligible diagnosis following pathology review, including ‘seromucinous’ tumors (n=55), unknown or unclassifiable discordant diagnosis (n=54, Supplementary Table S2) or no tumor in the block (n=17). For 77 cases with 2 or more slides suitable for inclusion, the slide with the most representative and/or highest tumor cellularity was selected. Following RNA extraction, another 36 samples with yield less than 32 ng/uL were excluded.

RNA samples from a total of 634 patients were eligible for the NanoString plexset assay, extracted from either formalin-fixed, paraffin embedded (FFPE) whole sections (n=403), FFPE cores (n=191), or fresh-frozen sections (n=40). Samples from the prognostic gene discovery analysis were excluded from validation analyses to preclude overfitting of the data (n=54). A second sample was analysed in a subset of 33 patients: either multiple blocks from the same tumor or multiple tumor tissue sites.

Gene selection

We analysed two datasets to select 19 genes of potential prognostic or diagnostic value in MOC. Candidate prognostic genes were identified based on analysis of 513 genes run on a NanoString platform (Appendix 1, Supplementary methods). The dataset included 60 MOCs among a study of predominantly high-grade serous ovarian cancers which have been published elsewhere. (22, 23) We identified four genes (THBS2, TAGLN, DCN, PLA2R1) that were differentially expressed between low (I/II, n=49)) and high (III/IV, n=11) stage MOC (Supplementary methods Table A), and increased expression of three of these (THBS2, TAGLN, DCN) were associated with poorer OS on univariate analysis (Supplementary Methods, Table B).

Candidate diagnostic classification genes (*MUC16* (encoding CA125), *GKN1*, *PGC*, *MEP1A*, *KRT20* (encoding CK20), *MUC5AC*, *CLDN18*, *VSIG1* and *ANXA10*) were from an analysis by the Genomic Analysis of Mucinous Tumours (GAMuT) study (24), whereby an exploratory RNASeq cluster analysis was performed to differentiate between benign mucinous ovarian tumors, mucinous borderline ovarian tumors (MBOT), MOC and upper and lower GI metastases to the ovary (Supplementary

Methods). The goal was to identify differential markers between entities with biological plausibility and available antibodies for future potential validation by IHC. We selected six additional genes (*ERBB2, TYMS, SATB2, MUC2, PD-1, PD-L1*) for diagnostic or therapeutic interest from the literature.(4, 25-27) Housekeeping genes (*DNAH6*, *LDHA*, *MTG1*, *POLR1B*, *TBP*) were selected based on consistent expression across different cancer types using publicly available TCGA RNASeq data for colorectal adenocarcinomas (COAD), ovarian (OV), pancreas (PAAD), stomach (STAD) and in the GAMuT RNASeq dataset for mucinous histology (Supplementary Methods).

NanoString PlexSet™ assay

Extraction of RNA and sample preparation for the NanoString assay was as described previously.(22, 23) A Plexset-24 assay of 24 customized probes (Supplementary Table S3) was used and due to the multiplex design, one patient sample with adequate quantity was selected as an internal calibrator. The assay was run by the Ramaciotti Centre for Genomics (UNSW Sydney, Australia).

Data Quality Assurance and Normalisation

We performed single sample data normalization as previously described (28), with adjustments to account for the Plexset assay. Raw counts were normalized to the housekeeping genes and then to the calibrator sample. Expression of the housekeeping gene *DNAH6* was at the limit of detection and the data were therefore excluded. We transformed the normalized gene expression data by taking the logarithm with base 2. Quality control (QC) measures were assessed by sample, by codeset and by cartridge to examine relevant levels of variability. Measures included

the signal to noise ratio (SNR <150), percentage of genes detected (above background plus two standard deviations) and expected expression of housekeeping genes.

Immunohistochemistry (IHC) and silver in-situ hybridisation (SISH)

We performed ERBB2/HER2 IHC using anti-HER2/neu (4B5), Roche Diagnostics, (6 ug/mL) and SISH using HER2/Ch17 Dual ISH DNA Probe Cocktail, Roche Diagnostics, concentration (14.24 ug/mL). Staining was performed on the Ventana Benchmark ULTRA Platform on 4 micron tissue microarray sections for a subset of cases from one study (WMH). For ERBB2/HER2 IHC, we used serous endometrial scoring guidelines(29) and a score of 3+ was given where >30% tumor cells showed intense complete membrane or basolateral membrane staining. Positive amplification was defined as either clusters (signal in >20 cells) or HER2/CEP17 Ratio ≥2 or >6 copies/nucleus and IHC 2+.

Statistical analysis

Overall survival (OS) was estimated using Cox proportional hazards, with right censoring at 10 years, and left truncation of prevalent cases. Validation of the association between gene expression and survival for the 4 candidate prognostic genes (*THBS2*, *TAGLN*, *DCN*, *PLA2R1*) was limited to new cases, removing the 54 overlapping samples from the discovery dataset. All multivariable analyses were adjusted for age and tumor stage and stratified by study site. Survival analyses of gene expression data used continuous normalized mRNA expression, examining one gene per model. The proportional hazards assumption was tested using the cox.zph function in the survival package in R. Survival curves were produced using

the Kaplan-Meier method. For visualization, survival curves of expression by tertile for significant genes were plotted. A time-dependent analysis was performed to assess pattern of invasion in MOC (all stages and Stage I alone) using the survSplit function in R(30), with stratification applied at 0-2 years versus >2 years based on inspection of the survival curves. This was run with and without left truncation to ensure consistent results for time from diagnosis as well as from study entry. Comparisons of gene expression between groups were performed using either the Student's t-test for 2 group comparisons or one-way ANOVA with Tukey's post hoc test for multiple comparisons. Correlation between mRNA expression and immunohistochemistry scores for ERBB2/HER2 were calculated with Spearman's correlation coefficients. Correlation between expression of all 19 candidate genes in different tumor blocks from the same patient was calculated using the Pearson's correlation coefficient. All statistical analyses were performed using R v4.1.2.
We performed all analyses of gene expression data on samples where the original diagnosis was concordant with the pathology review of the tissue being run on the assay to avoid misclassification.

Bioinformatics analysis

We used unsupervised hierarchical clustering and clustered samples based on gene expression profiles. We used the 'complete' agglomeration method and measured the Euclidean distance between samples. The heat maps were drawn using the iheatmapr package (v0.5.1) in R.(31) Diagnosis groups in the clustering were MBOT, low stage (I/II) MOC, advanced stage (III/IV) MOC, pancreas, gastric and lower GI (colorectal and appendiceal combined). We used random forest analysis and stratified bootstrapping (32) to assess the ability of the gene expression profiles to

predict the disease class (diagnosis group) of each sample. The cohort was divided into independent training and testing sets using stratified random sub-sampling, maintaining a balanced proportion of samples of each disease class. The training dataset was used to train a random forest classifier (the randomForest package in R, version 4.6-14) using default parameters and the classifier was benchmarked against the test set to obtain an error rate (Supplementary Methods). We repeated the above analyses 100 times to obtain a distribution of error rates, the mean overall error rate, and the mean and standard deviation of each element of the confusion matrix, to tabulate the number of samples associated with the actual and predicted class.

Data Availability

The data generated in this study are publicly available in the Gene Expression Omnibus (GEO), (accession number GSE203611).

RESULTS

Patient cohort

We generated RNA expression for 19 candidate genes from 634 patients, on a NanoString plexset assay, of which one patient sample was used as a calibrator and excluded from further analysis. Technical replicates (n=13) showed high correlation (intraclass correlation coefficient range 0.94-0.99). Following data processing, 29 samples failed QC and were excluded. 54 samples and 7 genes overlapped the discovery NanoString dataset and the Plexset and the observed adjusted intraclass coefficient was 0.69 (median R=0.90, range 0.34 *PD-L1* to 0.98 *ERBB2*). The final analytic cohort of 604 patients was divided into four diagnostic groups, MOC

(n=333), MBOT (n=151), upper GI (n=65), and lower GI (n=55) (Table 1). Of the 333 MOCs, 226 were low stage (I/II) (86% of cases with known stage). Upper GI included primary and metastatic pancreatic ductal adenocarcinoma (PDAC), intraductal papillary mucinous neoplasms (IPMN) with invasion, pancreatic mucinous cystadenocarcinomas and gastric adenocarcinomas. Lower GI included primary and metastatic mucinous and non-mucinous colorectal and appendiceal tumors.

Pathology review concordance and data analysis

Pathology review found 107 of the 604 cases were discordant between the original diagnosis and the review diagnosis of the sample run on the Plexset (Supplementary Table S4). Given the known intra-tumoral pathological heterogeneity of large mucinous ovarian tumors, and the focal nature of some MOC, we considered that these may be cases where the tissue submitted was not representative of the overall patient diagnosis (e.g. a block from a MOC case that contains only mucinous borderline tumor tissue). These patients were included in survival analysis that were unrelated to specific tissue features, based on their highest pathological diagnosis. For analyses involving features of the tissue itself (pattern of invasion and gene expression), we only included the concordant cases (n=497) to avoid misclassification. For each analysis, samples with missing clinical data were also removed where relevant, while attempting to maximise the sample size in this rare histotype (Figure 1).

Prognosis

Overall survival (OS) by tumor group

Survival analysis included all patients with a concordant diagnosis and those where a non-representative block was submitted (604 patients), of which 582 had complete clinical data. The 5-year unadjusted OS was highest in MBOT (88%), intermediate for MOC (71%), and considerably lower for lower GI (56%) and upper GI (29%, log-rank p<0.0001) (Figure 2A). We also examined OS in MOC by FIGO stage (n=184) and observed decreasing OS with increasing FIGO stage (p<0.0001). (Figure 2B)

Pattern of invasion in Mucinous Ovarian Carcinoma

Pattern of invasion was available for 208 MOC cases, with 167 (80%) classified as expansile and 41 (20%) as infiltrative. The proportion of cases with an infiltrative pattern increased with more advanced stage, 18% of Stage I MOC had an infiltrative pattern, as did 27% of Stage II, 29% of Stage III and 80% of Stage IV MOC (Supplementary Figure S2). Of the cases with FIGO stage data, 178 had survival data. Univariate survival analysis demonstrated that an infiltrative growth pattern was associated with poorer OS (HR 2.20 (1.33-3.64), p<0.01, Table 2, Figure 2C), however multivariable modelling adjusting for age, stage, and stratified by study site violated the proportional hazards assumption, suggestive of a time-dependent association. A time split analysis was performed for the periods 0-2 years and >2years after diagnosis based on inspection of the survival curves. This showed a significant time-dependent association between infiltrative growth pattern and poorer OS at 0-2 years after diagnosis (adj-HR 3.06 (1.49-6.29), p=0.002), but was not significant during the period >2 years (p=0.297). Similarly, within Stage I MOC (n=134), the Kaplan-Meir curves showed that most deaths in the infiltrative type occurred within the first 2 years of diagnosis (Figure 2D). A significant association between infiltrative subtype and poorer OS in Stage I MOC was observed within the first 2 years following diagnosis (adj-HR 2.77 (1.04-7.41), p=0.042), Table 2).

Of the 19 genes analysed, 12 had a statistically significant difference in mean expression between expansile and infiltrative subtypes (n=208) (Supplementary Figure S3). Eight genes were significantly higher in infiltrative (*THBS2*, *TAGLN*, *DCN*, *SATB2*, *GKN1*, *MUC16*, *PLA2R1, MUC2)* and the expansile subtype had significantly higher expression of *ERBB2*, *PGC, ANXA10* and *CLDN18* (Supplementary Figure S3). In FIGO Stage I cases (n=134), 6 of these genes were higher in the infiltrative subtype (*THBS2*, *TAGLN*, *DCN*, *PLA2R1, SATB2*, *GKN1*), and 1 higher in expansile (*ERBB2*, Supplementary Figure S3).

Gene expression and overall survival (OS)

We assessed the association between gene expression and survival in 233 MOC patients. Univariate analysis found five genes associated with OS - *THBS2*, *TAGLN*, *DCN*, *PLA2R1, ERBB2* (Table 3). After adjusting for age and stage and stratifying by study site, increased expression of two genes were associated with poorer OS: *THBS2*, HR 1.25 (95% CI 1.04-1.51), p=0.016 and *TAGLN* 1.21 (1.01-1.45), p=0.043. We plotted tertiles of expression for each gene for visualization (Supplementary Figure S4). These two genes were also upregulated in tumors with an infiltrative pattern of invasion (Supplementary Figure S5).

Survival was also assessed in upper and lower GI patients. Increased expression of *MUC2* was associated with better OS in lower GI tumors adjusted for age, stage, tumor type (colon/appendix) and stratified by study site (HR 0.72 (0.55-0.95), p=0.020), Supplementary Table S5). There were no prognostic associations between gene expression and OS in upper GI tumors in multivariable analyses.

Diagnosis

Clustering and diagnostic predictions

We identified nine genes in the RNAseq analysis (Figure A, Supplementary methods) with the goal of differentiating between MBOT, MOC and upper and lower GI cancers. A random forest model was trained and tested after stratified bootstrapping to produce balanced proportions in each diagnostic group class (Supplementary Table S6). We then used unsupervised hierarchical clustering to visualize clusters. To replicate the discovery analysis, we only included tissue samples from the ovary (MBOT, MOC and upper and lower GI metastases to the ovary, n=397). The mean testing error rate was 0.38 (Supplementary Figure S6, Table S7), and this poor validation is also reflected in the heatmap (Supplementary Figure S7). Following this, we ran a model with all 19 candidate genes and all pathology-concordant samples (with stage data for MOC, n=479). The mean testing error rate of the model was 0.33 (equivalent to an overall accuracy of 67%, Supplementary Figure S6, Table S7). Lower GI samples were most accurately predicted (9/12, 75%) and upper GI were no greater than chance (50%). A heatmap of these samples shows the lower GI samples clustering out in one main group, the pancreatic samples mainly in cluster 2, and the five gastric samples across clusters 1 and 4 along with MOC samples (Figure 3).

Comparison of gene expression across tumor groups

To examine similarities and differences between MOC, upper GI and lower GI cancers we compared gene expression between all pathology-concordant, invasive cases (n=363). While the random forest models and clustering showed that this gene set had limited ability to distinguish tumor groups overall, the mean expression of

several individual genes differed significantly between tumor groups (Supplementary Figure S8). Sixteen genes significantly differed between MOC and lower GI tumors. *ANXA10, CLDN18, ERBB2/HER2, MUC16, MUC5AC, PGC, VSIG1* all showed significantly higher expression in MOC and *MEP1A*, *PD1*, *DCN*, *TAGLN*, *THBS2*, *GKN1*, *CK20*, *MUC2*, *SATB2* were significantly lower in MOC. Twelve genes differed significantly between MOC and upper GI. Two genes contrasted with opposing directions – *MEP1A* higher in MOC compared to upper GI, *MUC16* higher in MOC compared with lower GI. Expression of the immune marker *PD-1* was lower in MOC compared with both upper and lower GI (Supplementary Figure S9) and *PD-L1* was relatively similar across all groups, but slightly lower in MOC compared with upper GI ($p=0.03$).

ERBB2/HER2 expression and implications for therapy

We analysed mRNA expression of ERBB2/HER2 by NanoString for MBOT (n=134), MOC (n=243), lower GI (n=55) and upper GI (n=65). Expression of *ERBB2*/*HER2* was higher in MOC compared with both lower GI and upper GI and higher in Stage I and II MOC respectively compared with MBOT ($p<0.001$ and $p<0.001$, Supplementary Figure S10). A subset of cases (n=37) was examined for *ERBB2*/*HER2* amplification using SISH, showing clear delineation between *ERBB2*/*HER2* amplified (n=7) and non-amplified (n=28, plus two equivocal) with respect to their mRNA expression levels (Supplementary Figure S11A). We used these data to estimate that a threshold of normalized mRNA expression ≥2.5, represented potential amplification. When we applied this threshold from the SISH subset to all 243 MOC cases, 26% were considered high expressing, i.e., estimated to be amplified. The threshold was supported by the comparison of mRNA with

protein expression from IHC, whereby 15 of 17 ERBB2/HER2 3+ staining tumors had >2.5 mRNA expression (Supplementary Figure S11B). The proportion of high expression/amplified was higher in low stages of MOC compared with advanced stage (Stage I 36/139, 26%; Stage II 6/16, 38%; versus Stage III/IV 4/31, 13%, Supplementary Table S8). We did not observe differences by grade, however *ERBB2*/*HER2* high/amplified were more common in MOC with an expansile pattern of invasion (chi-square p=0.008), 50/163 (31%) compared with 4/39 (10%) of infiltrative cases. In summary of known poor prognosis features only, just 8/243 (3%) of MOC cases were either infiltrative (4/39, 10%) or Stage III/IV (4/31, 13%) *and* *ERBB2*/*HER2* high/amplified. There was no association between high expression/amplified *ERBB2*/*HER2* and OS (log-rank p=0.2) (Supplementary Figure S12).

Prediction modelling of non-representative tissue

We trained and tested models using 246 concordant MOC and 139 concordant MBOT samples, to predict the diagnosis of 90 discordant samples that were submitted as carcinoma (MOC), but pathology review deemed MBOT. The random forest model had a relatively low mean testing error rate of 0.18 (Supplementary Figure S6), and out of the 90 discordant cases, 53 were predicted to be MBOT, i.e. 59% of predictions concorded with the pathologist review, and the rest were predicted to be MOC. (Supplementary Table S9).

Paired sample analysis

There were 33 pairs of samples from the same patient and the same diagnostic episode, consisting of seven cases with MBOT and MOC, one case benign and

MBOT blocks, one case left and right ovary blocks, 16 cases with two MBOT blocks, two cases with primary appendix and metastases to the ovary, and six cases with different metastatic tissue sites (Supplementary Table S10). We examined the correlation in gene expression between samples, and from 16 sets of MBOT tissue from different blocks for the same patient, correlation was variable: seven sets R>0.9, four sets R =0.7-0.9, five sets <0.7. Two sets of primary low grade appendiceal mucinous neoplasm (LAMN) and metastases to the ovary had very high correlation, R=0.94 and R=0.95. Three out of seven sets of MBOT and MOC from the same patient had strong correlation (R>0.9), three moderate R=0.7-0.9 and one with poor correlation R=0.40. Differences in correlation across sets of tumor samples was not related to differences in cellularity between samples, with 82% (9/11) of pairs with R<0.8 both having tumor cellularity of <60%, as did 86% (18/21) of pairs with R>0.8 (Chi-squared test, p=0.8).

DISCUSSION

We found that increased expression of two markers, thrombospondin 2 (THBS2) and transgelin (*TAGLN*), were associated with poorer OS in MOC after adjustment for age and tumor stage. Thrombospondin 2 (THBS2) is a glycoprotein with a role in tumor growth, angiogenesis and metastases, with high expression found to be associated with poorer survival colorectal cancer at the mRNA and protein level.(33) In ovarian cancer, *THBS2* mRNA expression has been shown to be upregulated in more aggressive tumors (malignant compared to borderline), advanced stage and high grade.(34) There may be role variations in different tumor types as *THBS2* has been reported to be an inhibitor of angiogenesis in cervical cancer.(35) The role of *THBS2* in prognosis may be driven by an interaction with the extra-cellular matrix,

enabling tumor progression and metastases. Transgelin (*TAGLN*) is an actin-binding protein, expressed in smooth muscle cells. Multiple studies in colorectal, gastric, pancreas, non-small cell lung cancer have shown increased *TAGLN* expression is associated with migration, invasion and poor survival,(36-38), however others have suggested it is a tumor suppressor in colorectal cancer.(39) Both prognostic genes appear to be expressed in the stroma, with upregulation of *TAGLN* in gastric stromal carcinoma-associated fibroblasts, (40) and increased expression of *THBS2* implicated in tumor progression and poor prognosis in pancreatic cancer, excreted by stromal fibroblasts.(41) This apparent stromal localisation could also explain the higher expression levels observed in the infiltrative MOC compared with expansile and subsequent prognostic association. Indeed, both *THBS2* and *TAGLN* expression were higher in the samples with low tumor cellularity, inferring at least some expression may be due to the higher stromal content of the samples (Supplementary Figure S13). Expression of *TAGLN* has been implicated with *KRAS* signalling in promoting proliferation in pancreatic cancer,(42) *KRAS* mutations being the most frequent aberration in MOC.(24) When both genes were combined in the same survival model, the associations were no longer significant, and correlation was high (R=0.8, Supplementary Table S11) suggesting a possible contributory effect of the two genes. Examination of the role of both *THBS2* and *TAGLN* in large clinical cohorts is critical, and validation of the current finding is needed to confirm the prognostic potential of these markers and to further explore their role in the biology of MOC.

We observed a time-dependent association between pattern of invasion and OS with an infiltrative pattern associated with poorer OS within two years from diagnosis, but not significant after two years. This finding was consistent when assessing FIGO

Stage I cases alone. This subset is arguably the most clinically relevant for the prognostic value of pattern of invasion: a poor outcome marker will influence decision making when considering adjuvant chemotherapy or more vigilant monitoring for recurrence. Prior studies have reported varying outcomes with regard to progression-free survival (PFS) and OS in the infiltrative subtype(15-20), however most have not adjusted survival models for age and stage, and no single study has observed a prognostic association in Stage I MOC alone. The largest series to-date(16) included 67 patients and no multivariable analyses were performed. A similarly sized study of Stage I only MOC (n=64)(17) found no statistically significant difference in PFS (p=0.49) or OS (p=0.18). Hada et. al reported that an infiltrative pattern was associated with poorer PFS HR 9.01 (2.28–61.41) p<0.01 and OS HR 17.56 (2.58-393.24), p<0.01 but this study was underpowered to analyse Stage I alone. Combining Stage I and II in univariate analysis (n=38) they found a significant impact on PFS (p=0.03), but OS was not evaluable.(20) Time-dependent associations of prognostic factors have been described in other cancers such as triple negative breast cancer,(43) similar to our observation in patients with Stage I MOC with an infiltrative pattern with early recurrence and death, and a low risk beyond the 2-year mark. The proportion of infiltrative cases here is lower than many series, and it is possible that others used a different threshold, stringency in excluding metastases or heterogeneity between blocks has led to this difference. We classified an infiltrative pattern at >5 mm and cases with only a small focus of destructive invasion were grouped with the expansile. Tabrizi et al. report a similar low frequency at 13% (4/31) in a population-based series and suggests that other institutional studies with higher rates of infiltrative cases may reflect more complex, selected populations.(19) Of note in the current study, four of the Stage I cases were

reported to have an anaplastic component: two infiltrative and one expansile, all of whom died within two years and one infiltrative case alive after seven years. Although anaplastic carcinoma arising in mural nodules is considered to infer more aggressive disease, some report that their presence in Stage I disease does not influence outcome.(18) It cannot be ruled out that the small number in this study influenced our findings. Given that infiltrative invasion is a feature of metastatic neoplasms to the ovary and was observed more frequently in higher stage MOC, we also cannot rule out that some of these cases represent undiagnosed metastases from a different primary site or inadequate staging of disease. Expression of the two prognostic markers *THBS2* and *TAGLN* was significantly higher in the infiltrative subtype compared with expansile. In contrast, *ERRB2* encoding HER2 expression was lower in infiltrative MOC compared with expansile MOC, which is consistent with other reports of HER2 positive MOC on IHC associated with the expansile subtype and better prognosis.(44)

This study has replicated the survival patterns seen in the literature for MOC (6) and GI tumors, showing that advanced stage MOC and upper GI cancers have significantly poorer survival than MBOT, Stage I MOC and lower GI cancers. Notably, the difference observed in 5-year survival between Stage I (79%) and II (69%) indicates that studies in MOC should not combine these ‘low’ stages together in analyses, which is the practice for ovarian endometrioid carcinomas (45).

The discovery RNASeq analysis identified a 9-gene classifier to help differentiate between MBOT, MOC and metastases to the ovary, however we did not validate this in the larger cohort. This could be due to cohort differences, such as inclusion of benign and ‘seromucinous’ cases and few GI tumors in the RNASeq dataset. We were limited by the 19-gene panel in this large follow-up study using formalin-fixed

paraffin-embedded tissue, and more work is needed to identify other possible diagnostic classifiers that may have been missed by this study. Despite this, clustering of the whole gene set found that most lower GI tumors separated out prominently in one main cluster, but MOCs and upper GI grouped together. This, along with differences in expression between groups revealed more differences between MOC and lower GI compared with MOC and upper GI. Recent improvements in diagnostic classification now includes use of CK7 and SATB2 for lower GI metastases(4), but in contrast differential markers for upper GI tumors remain elusive. One prior study showed potential for *MEP1A* with lower membranous staining in MOC compared with pancreatic cancers(46), however in our cohort the mean mRNA expression was higher in MOC compared with upper GI tumors (Figure S8), including in comparison with pancreatic tumors alone (p=0.006, Supplementary Figure S8). It is possible that the challenges of differentiating MOC from mucinous pancreatic and gastric cancers could shift the therapeutic strategies for MOC. Considering the similarities between MOC and pancreatic tumors, we see high rates of co-existing *CDK2NA* inactivation (76%), and a similar frequency of *TP53* mutations (~60%).(24) Likewise, the gastroesophageal junction tumors share the features of *ERBB2* amplification and *TP53* mutations.(47) There is an argument to shift focus from trying to seek differences between groups and look at opportunities for basket-style clinical trials of either systemic or targeted therapies by including advanced stage MOC together with GI cancers based on shared molecular alterations.(48) For example, FOLFIRINOX is standard of care in metastatic pancreatic cancers but has not been investigated in advanced stage MOC.(49) In addition, 20-30% of MOCs have been reported to harbor *ERBB2* amplification(26, 27, 50), consistent with our finding (26%). Our findings on ERBB2/HER2

amplification/overexpression confirm the results of previous studies (26, 27, 44, 50), including the observation of a lower frequency of *ERBB2/HER2* high/amplified cases in advanced stage MOC (4/31, 13%). Similarly, 4/39 (10%) of infiltrative subtype MOC were *ERBB2/HER2* high/amplified compared with 31% with an expansile pattern, consistent with the study by Kim et al who reported 0/9 infiltrative and 14/37 expansile were HER2 positive.(44) If the suitable population for anti-HER therapy were limited to high stage or infiltrative MOC, our data suggest that about ~3% of patients may be considered. Despite this, in addition to high/amplified cases, there have been promising developments in treatment of HER2-low (IHC 1+) in breast cancer(51) which may broaden eligibility to these therapies for patients with advanced stage MOC and HER2 1+ or 2+ on IHC. Additional important developments in anti-HER2 directed therapy in gastric cancer now include antibody-drug conjugates such as Trastuzumab Deruxtecan in the advanced setting (52), and a potential role for XELOX-T (oxaliplatin, capecitabine and trastuzumab) in locally advanced, resectable gastric cancer.(53) The latter therapy regimen is based on a small phase II study, however future large randomized studies could arguably adopt a basket design to include *ERBB2/HER2* amplified MOCs as well as potentially all tumors with high ERBB2/HER2 expression on IHC. Indeed, the current Bouquet-ENGOT-gyn2 rare ovarian cancers basket trial (ClinicalTrials.gov identifier: NCT04931342) includes an arm for *ERBB2/HER2* amplified/mutated cases for treatment with Trastuzumab emtansine.

The current study did not provide a simple mRNA profile that can be used diagnostically to distinguish MBOT from MOC and it highlighted the heterogeneity through varying concordance of expression between borderline and invasive carcinoma and between multiple borderline tumor blocks from the same patient.

Whether the 25% of MOC cases considered borderline on pathology review reflect a genuine discrepancy between pathologists, or the submission of a non-representative block from a heterogeneous tumor remains unclear and should be the subject of further studies. Interestingly, a recently reported French cohort (n=79) with access to all blocks or a minimum of 5 blocks also reclassified 18% of MOC as MBOT.(16)

In the context of exploring better therapeutic options for MOC, we observed lower levels of expression of *PD-1* in comparison to GI tumors, and similar levels of *PD-L1*. As immunotherapy is now being investigated in multiple cancer types, further studies with appropriate immunohistochemical scoring for *PD-1* and *PD-L1* should be carried out to understand whether a subset of MOC may benefit from immunotherapy.

There are several limitations to the current study that has combined samples and data on a large scale over many years. Since tumor heterogeneity is well recognized in MOC, it is possible that the blocks sectioned for the study were not representative. This was highlighted by the discordant diagnoses which may be due to sampling or individual pathologist's interpretation. Although 30% of mucinous ovarian tumors (n=104) had IHC performed for CK7, CK20, CDX2, SATB2 and PAX8 in a prior study(4), we were unable to perform this diagnostic panel on all cases and could not confirm whether they were done as part of routine pathological assessment. Notwithstanding this limitation, the majority of misclassified samples related to discordance between MOC and MBOT. The IHC panel would be of limited assistance to differentiate these entities as the diagnosis is based on H&E. This panel would have limited utility in differentiating *upper GI* from MOC due to their similarities in the IHC phenotype. In addition, the lack of guidelines for HER2 scoring in MOC could have misclassified some cases on IHC, however the follow-up SISH to

determine HER2 amplification would mitigate against this. Survival analyses lacked residual disease, progression and cause of death data. A major strength is that this is the largest series to-date of gene expression profiles of MOC and includes comparisons with upper and lower gastrointestinal tumors on the same profiling platform. Future work could also assess mutation profiles to identify mRNA expression differences in *KRAS*/*TP53* mutant and wildtype subsets.

This analysis of a large series of mucinous ovarian carcinomas has identified two potential prognostic biomarkers in *THBS2* and *TAGLN* which could have clinical utility and deserve further investigation. In addition, we confirmed the importance of an infiltrative pattern of invasion as a risk indicator for early recurrence and mortality. Given their rarity, there is a strong argument supporting inclusion of MOC in basket trials with similar and much more common gastrointestinal mucinous cancers.

Acknowledgements

We thank all the patients who participated in the studies combined for this research, and all the scientists, clinicians, technical and administrative staff who have contributed. Special thanks to the staff of the Ramaciotti Centre for Genomics (UNSW Sydney, Australia). The AOCS gratefully acknowledges additional support from S. Boldeman, the Agar family, Ovarian Cancer Action (UK), Ovarian Cancer Australia and the Peter MacCallum Foundation. The AOCS also acknowledges the cooperation of the participating institutions in Australia and acknowledges the contribution of the study nurses, research assistants and all clinical and scientific collaborators to the study. The complete AOCS Study Group can be found at http://www.aocstudy.org. We thank Mie Konno, Michelle Darago, Faye Chambers, SEARCH team, Craig Luccarini, Caroline Baynes, Don Conroy, I. Jacobs, S. Gayther, E. Wozniak, J. Ford and N. Balogun. We thank the Genetic Pathology Evaluation Centre (GPEC) for technical support in construction of tissue micro-arrays; and the OVCARE clinical informatics core, Gynecological tissue bank (GTB), and Cheryl Brown Gynecologic Cancer Outcomes Unit for facilitating recruitment of patients, banking biospecimens, and maintaining clinical data. We thank the Gynaecological Oncology Biobank at Westmead, a member of the Australasian Biospecimen Network-Oncology group. Biospecimens and clinical data

were provided by the Australian Pancreatic Cancer Genome Initiative (APGI, www.pancreaticcancer.net.au) which is supported by an Avner Pancreatic Cancer Foundation Grant, www.avnersfoundation.org.au. We wish to thank Heather Thorne, Eveline Niedermayr, Sharon Guo, all the kConFab research nurses and staff, the heads and staff of the Family Cancer Clinics, and the Clinical Follow Up Study (which has received funding from the NHMRC, the National Breast Cancer Foundation, Cancer Australia, and the National Institute of Health (USA)) for their contributions to this resource, and the many families who contribute to kConFab. The ViP study thanks the participants and acknowledges the contributions of the site principal investigators Geoffrey Lindeman, Marion Harris, Lucinda Salmon, Ingrid Winship and Yoland Antill, and the staff at the Peter MacCallum Cancer Centre, Royal Melbourne Hospital, Monash Health, Cabrini Health and Barwon Health Familial Cancer Centres and the Austin and Tasmanian Clinical Genetics Services.

The design of this study was informed in part by data generated by the TCGA Research Network: https://www.cancer.gov/tcga.

Funding information

We are grateful for financial support for this study from UNSW Faculty of Medicine and Health. Funding for this study came from Michael L Friedlander NHMRC program grant APP1092856, and the accelerator grant of NSW Health PhD Scholarship Program awarded to Nicola S Meagher.

Funding for individual studies

AOCS: The Australian Ovarian Cancer Study was supported by the U.S. Army Medical Research and Materiel Command under DAMD17-01-1-0729, The Cancer Council Victoria, Queensland Cancer Fund, The Cancer Council New South Wales, The Cancer Council South Australia, The Cancer Foundation of Western Australia (Multi-State Applications 191, 211, and 182), The Cancer Council Tasmania and the National Health and Medical Research Council of Australia (NHMRC; ID199600; ID400413 and ID400281); AOV: Canadian Institutes of Health Research (MOP-86727).; BAV: ELAN Funds of the University of Erlangen-Nuremberg; DOV: NCI/NIH R01CA168758, NCI/NIH R01CA112523 and NCI/NIH R01CA087538; HSA: The Health Science Alliance (HSA) Biobank acknowledges financial support from the Translational Cancer Research Network, funded by the Cancer Institute NSW. KRA: This study was partially supported by a grant from the Korea Health Technology R&D Project through the Korea Health Industry Development Institute (KHIDI), and the National R&D Program for Cancer Control, Ministry of Health & Welfare, Republic of

Korea (HI16C1127; 0920010); POC: POC study was supported by the program of the Minister of Science and Higher Education under the name "Regional Initiative of Excellence" in 2019-2022 project number 002 / RID / 2018/19 amount of financing 12 000 000 PLN; PRM: The University of Pittsburgh School of Medicine Dean's Faculty Advancement Award (Modugno) provided support for this research. Funding was also provided by the National Science Foundation (DGE-2217399-Modugno). The views expressed are those of the authors and do not necessarily reflect the views of the National Science Foundation. This project used the UPMC Hillman Cancer Center and Tissue and Research Pathology/Pitt Biospecimen Core shared resource which is supported in part by National Cancer Institute award P30CA047904; SEA: Cancer Research UK C490/A16561, the UK National Institute for Health Research Biomedical Research Centres at the University of Cambridge, Cambridge Cancer Centre. The University of Cambridge has received salary support for PDPP from the NHS in the East of England through the Clinical Academic Reserve. VAN: BC's Gynecological Cancer Research Team (OVCARE) receives core funding from The BC Cancer Foundation and the VGH and UBC Hospital Foundation; WMH: National Health and Medical Research Council of Australia, Enabling Grants ID 310670 and ID 628903. Cancer Institute NSW Grants 12/RIG/1-17 and 15/RIG/1-16. The Westmead GynBiobank acknowledges financial support from the Sydney West Translational Cancer Research Centre, funded by the Cancer Institute NSW. VIP: The Variants in Practice study was supported by the National Health and Medical Research Council of Australia (NHMRC; APP1023698; 1092856).

NSM is supported by the NSW Ministry of Health and UNSW Sydney under the NSW Health PhD Scholarship Program, and the Translational Cancer Research Network. KLG supported by US Dept of Defense (OC170121), Australian National Health and Medical Research Council and the Peter MacCallum Cancer Foundation. SJR is supported by National Health and Medical Research Council of Australia (NHMRC) grant APP2009840.
DDLB is supported by the National Health and Medical Research Council of Australia (NHMRC Fellowship (APP1117044) and Program Grant APP1092856) and the United States Department of Defence (DoD) Ovarian Cancer Transitional Leverage Award (W81XWH-12-1-0104); MK received support through the Calgary Laboratory Services research support fund (RS19-612). AB received support through the Ovarian Cancer Research Fund (OCRF). JDB acknowledges funding and support from Cancer Research UK (grant numbers A22905, A15601 and A17197). MA is funded through a Michael Smith Foundation for Health Research Scholar Award and the Janet D. Cottrelle Foundation Scholars programme managed by the BC Cancer Foundation. DGH receives support from the Dr. Chew Wei Memorial Professorship in Gynecologic Oncology and the Canada

Research Chairs programme (Research Chair in Molecular and Genomic Pathology). ADeF. was funded by Cancer Institute NSW grant 15/TRC/1-01. BK is funded through an American Cancer Society Early Detection Professorship (SIOP-06-258-01-COUN) and the National Center for Advancing Translational Sciences (NCATS), Grant UL1TR000124. KS received support from the Swedish Cancer Foundation (CAN 18-384) and the Swedish state under the agreement between the Swedish government and the county council (*ALFGBG-965552).* The contents of the published material are solely the responsibility of the authors and do not reflect the views of NHMRC.

References

1. Kelemen LE, Kobel M. Mucinous carcinomas of the ovary and colorectum: different organ, same dilemma. The Lancet Oncology. 2011;12(11):1071-80.

2. Zaino RJ, Brady MF, Lele SM, Michael H, Greer B, Bookman MA. Advanced stage mucinous adenocarcinoma of the ovary is both rare and highly lethal: a Gynecologic Oncology Group study. Cancer. 2011;117(3):554-62.

3. Seidman JD, Kurman RJ, Ronnett BM. Primary and metastatic mucinous adenocarcinomas in the ovaries: incidence in routine practice with a new approach to improve intraoperative diagnosis. Am J Surg Pathol. 2003;27(7):985-93.

4. Meagher NS, Wang L, Rambau PF, Intermaggio MP, Huntsman DG, Wilkens LR, et al. A combination of the immunohistochemical markers CK7 and SATB2 is highly sensitive and specific for distinguishing primary ovarian mucinous tumors from colorectal and appendiceal metastases. Modern Pathology. 2019;32(12):1834-46.

5. Ackroyd SA, Goetsch L, Brown J, Houck K, Wang C, Hernandez E. Pancreaticobiliary metastasis presenting as primary mucinous ovarian neoplasm: A systematic literature review. Gynecol Oncol Rep. 2019;28:109-15.

6. Peres LC, Cushing-Haugen KL, Kobel M, Harris HR, Berchuck A, Rossing MA, et al. Invasive Epithelial Ovarian Cancer Survival by Histotype and Disease Stage. J Natl Cancer Inst. 2019;111(1):60-68

7. Morice P, Gouy S, Leary A. Mucinous Ovarian Carcinoma. New England Journal of Medicine. 2019;380(13):1256-66.

8. Gore M, Hackshaw A, Brady WE, Penson RT, Zaino R, McCluggage WG, et al. An international, phase III randomized trial in patients with mucinous epithelial ovarian cancer (mEOC/GOG 0241) with long-term follow-up: and experience of conducting a clinical trial in a rare gynecological tumor. Gynecol Oncol. 2019;153(3):541-548.

9. Armstrong DK, Alvarez RD, Bakkum-Gamez JN, Barroilhet L, Behbakht K, Berchuck A, et al. Ovarian Cancer, Version 2.2020, NCCN Clinical Practice Guidelines in Oncology. Journal of the National Comprehensive Cancer Network : JNCCN. 2021;19(2):191-226.

10. Kurnit KC, Sinno AK, Fellman BM, Varghese A, Stone R, Sood AK, et al. Effects of Gastrointestinal-Type Chemotherapy in Women With Ovarian Mucinous Carcinoma. Obstetrics & Gynecology. 2019;134(6):1253-9.

11. Schlappe BA, Zhou QC, O'Cearbhaill R, Iasonos A, Soslow RA, Abu-Rustum NR, et al. A descriptive report of outcomes of primary mucinous ovarian cancer patients receiving either an adjuvant gynecologic or gastrointestinal chemotherapy regimen. International Journal of Gynecologic Cancer. 2019:ijgc-2018-000150.

12. Colombo N, Sessa C, Bois Ad, Ledermann J, McCluggage W, McNeish I, et al. ESMO–ESGO consensus conference recommendations on ovarian cancer: pathology and molecular biology, early and advanced stages, borderline tumours and recurrent disease. International Journal of Gynecologic Cancer. 2019;29(4):728-60.

13. Lee KR, Scully RE. Mucinous tumors of the ovary: a clinicopathologic study of 196 borderline tumors (of intestinal type) and carcinomas, including an evaluation of 11 cases with 'pseudomyxoma peritonei'. Am J Surg Pathol. 2000;24(11):1447-64.

14. Hart WR. Mucinous tumors of the ovary: a review. Int J Gynecol Pathol. 2005;24(1):4-25.

15. Busca A, Nofech-Mozes S, Olkhov-Mitsel E, Gien LT, Bassiouny D, Mirkovic J, et al. Histological grading of ovarian mucinous carcinoma – an outcome-based analysis of traditional and novel systems. 2020;77(1):26-34.

16. Genestie C, Auguste A, Al Battal M, Scoazec J-Y, Gouy S, Lacroix L, et al. Histological classification of mucinous ovarian tumors: inter-observer reproducibility, clinical relevance, and role of genetic biomarkers. Virchows Archiv. 2021;478(5):885-91.

17. Gouy S, Saidani M, Maulard A, Bach-Hamba S, Bentivegna E, Leary A, et al. Characteristics and Prognosis of Stage I Ovarian Mucinous Tumors According to Expansile or Infiltrative Type. International Journal of Gynecologic Cancer. 2018;28(3):493.

18. Rodríguez IM, Prat J. Mucinous Tumors of the Ovary: A Clinicopathologic Analysis of 75 Borderline Tumors (of Intestinal Type) and Carcinomas. 2002;26(2):139-52.

19. Tabrizi AD, Kalloger SE, Kobel M, Cipollone J, Roskelley CD, Mehl E. Primary ovarian mucinous carcinoma of intestinal type: significance of pattern of invasion and immunohistochemical expression profile in a series of 31 cases. Int J Gynecol Pathol. 2010;29(2):99-107

20. Hada T, Miyamoto M, Ishibashi H, Matsuura H, Sakamoto T, Kakimoto S, et al. Survival and biomarker analysis for ovarian mucinous carcinoma according to

invasive patterns: retrospective analysis and review literature. Journal of Ovarian Research. 2021;14(1):33.

21. WHO classification of tumours of female reproductive organs. 5th edition, Vol 4. International Agency for Research on Cancer, Lyon, France. 2020.

22. Millstein J, Budden T, Goode EL, Anglesio MS, Talhouk A, Intermaggio MP, et al. Prognostic gene expression signature for high-grade serous ovarian cancer. Annals of Oncology. 2020;31(9):1240-50.

23. Talhouk A, George J, Wang C, Budden T, Tan TZ, Chiu DS, et al. Development and Validation of the Gene Expression Predictor of High-grade Serous Ovarian Carcinoma Molecular SubTYPE (PrOTYPE). Clin Cancer Res. 2020;26(20):5411-23.

24. Cheasley D, Wakefield MJ, Ryland GL, Allan PE, Alsop K, Amarasinghe KC, et al. The molecular origin and taxonomy of mucinous ovarian carcinoma. Nature Communications. 2019;10(1):3935.

25. Meagher NS, Schuster K, Voss A, Budden T, Pang CNI, deFazio A, et al. Does the primary site really matter? Profiling mucinous ovarian cancers of uncertain primary origin (MO-CUP) to personalise treatment and inform the design of clinical trials. Gynecol Oncol. 2018;150(3):527-33.

26. Gorringe KL, Cheasley D, Wakefield MJ, Ryland GL, Allan PE, Alsop K, et al. Therapeutic options for mucinous ovarian carcinoma. Gynecologic Oncology. 2020;156(3):552-60.

27. Anglesio MS, Kommoss S, Tolcher MC, Clarke B, Galletta L, Porter H. Molecular characterization of mucinous ovarian tumours supports a stratified treatment approach with HER2 targeting in 19 % of carcinomas. J Pathol. 2013;229(1):111-20.

28. Talhouk A, Kommoss S, Mackenzie R, Cheung M, Leung S, Chiu DS, et al. Single-Patient Molecular Testing with NanoString nCounter Data Using a Reference-Based Strategy for Batch Effect Correction. PloS one. 2016;11(4):e0153844-e.

29. Buza N. HER2 Testing in Endometrial Serous Carcinoma: Time for Standardized Pathology Practice to Meet the Clinical Demand. Archives of pathology & laboratory medicine. 2020;145(6):687-91.

30. Therneau T, Crowson C, Atkinson E. Using Time Dependent Covariates and Time Dependent Coefficients in the Cox Model. 2022.

31. Schep AN, Kummerfeld SK. iheatmapr: Interactive complex heatmaps in R. Journal of Open Source Software. 2017;2(16):359.

32. Breiman L. Random Forests. Machine Learning. 2001;45(1):5-32.

33. Wang X, Zhang L, Li H, Sun W, Zhang H, Lai M. THBS2 is a Potential Prognostic Biomarker in Colorectal Cancer. Scientific Reports. 2016;6(1):33366.

34. Kodama J, Hashimoto I, Seki N, Hongo A, Yoshinouchi M, Okuda H, et al. Thrombospondin-1 and -2 messenger RNA expression in epithelial ovarian tumor. Anticancer research. 2001;21(4b):2983-7.

35. Wu XG, Zhou CF, Zhang YM, Yan RM, Wei WF, Chen XJ, et al. Cancer-derived exosomal miR-221-3p promotes angiogenesis by targeting THBS2 in cervical squamous cell carcinoma. Angiogenesis. 2019;22(3):397-410.

36. Shi YY, Wang HC, Yin YH, Sun WS, Li Y, Zhang CQ, et al. Identification and analysis of tumour-associated antigens in hepatocellular carcinoma. British journal of cancer. 2005;92(5):929-34.

37. Zhou H-m, Fang Y-y, Weinberger PM, Ding L-l, Cowell JK, Hudson FZ, et al. Transgelin increases metastatic potential of colorectal cancer cells in vivo and alters expression of genes involved in cell motility. BMC Cancer. 2016;16(1):55.

38. Zhou L, Zhang R, Zhang L, Sun Y, Yao W, Zhao A, et al. Upregulation of transgelin is an independent factor predictive of poor prognosis in patients with advanced pancreatic cancer. 2013;104(4):423-30.

39. Zhao L, Wang H, Deng Y-J, Wang S, Liu C, Jin H, et al. Transgelin as a suppressor is associated with poor prognosis in colorectal carcinoma patients. Modern Pathology. 2009;22(6):786-96.

40. Yu B, Chen X, Li J, Qu Y, Su L, Peng Y, et al. Stromal fibroblasts in the microenvironment of gastric carcinomas promote tumor metastasis via upregulating TAGLN expression. BMC Cell Biology. 2013;14(1):17.

41. Nan P, Dong X, Bai X, Lu H, Liu F, Sun Y, et al. Tumor-stroma TGF-β1-THBS2 feedback circuit drives pancreatic ductal adenocarcinoma progression via integrin αvβ3/CD36-mediated activation of the MAPK pathway. Cancer Letters. 2022;528:59-75.

42. Sun Y, Peng W, He W, Luo M, Chang G, Shen J, et al. Transgelin-2 is a novel target of KRAS-ERK signaling involved in the development of pancreatic cancer. Journal of Experimental & Clinical Cancer Research. 2018;37(1):166.

43. Foulkes WD, Smith IE, Reis-Filho JS. Triple-negative breast cancer. The New England journal of medicine. 2010;363(20):1938-48.

44. Kim SK, Cho NH. HER2-positive mucinous adenocarcinomas of the ovary have an expansile invasive pattern associated with a favorable prognosis. Int J Clin Exp Pathol. 2014;7(7):4222-30.

45. Krämer P, Talhouk A, Brett MA, Chiu DS, Cairns ES, Scheunhage DA, et al. Endometrial Cancer Molecular Risk Stratification is Equally Prognostic for Endometrioid Ovarian Carcinoma. Clin Cancer Res. 2020;26(20):5400-10.

46. Heinzelmann-Schwarz VA, Scolyer RA, Scurry JP, Smith AN, Gardiner-Garden M, Biankin AV, et al. Low meprin α expression differentiates primary ovarian mucinous carcinoma from gastrointestinal cancers that commonly metastasise to the ovaries. 2007;60(6):622-6.

47. Cassier PA, Peyramaure C, Attignon V, Eberst L, Pacaud C, Boyault S, et al. Precision medicine for patients with gastro-oesophageal cancer: A subset analysis of the ProfiLER program. Translational oncology. 2022;15(1):101266.

48. Craig O, Salazar C, Gorringe KL. Options for the Treatment of Mucinous Ovarian Carcinoma. Current treatment options in oncology. 2021;22(12):114.

49. Conroy T, Desseigne F, Ychou M, Bouché O, Guimbaud R, Bécouarn Y, et al. FOLFIRINOX versus Gemcitabine for Metastatic Pancreatic Cancer. 2011;364(19):1817-25.

50. Chay W-Y, Chew S-H, Ong W-S, Busmanis I, Li X, Thung S, et al. HER2 Amplification and Clinicopathological Characteristics in a Large Asian Cohort of Rare Mucinous Ovarian Cancer. PLOS ONE. 2013;8(4):e61565.

51. Modi S, Jacot W, Yamashita T, Sohn J, Vidal M, Tokunaga E, et al. Trastuzumab Deruxtecan in Previously Treated HER2-Low Advanced Breast Cancer. N Engl J Med. 2022;387(1):9-20.

52. Shitara K, Bang Y-J, Iwasa S, Sugimoto N, Ryu M-H, Sakai D, et al. Trastuzumab Deruxtecan in Previously Treated HER2-Positive Gastric Cancer. 2020;382(25):2419-30.

53. Rivera F, Izquierdo-Manuel M, García-Alfonso P, Martínez de Castro E, Gallego J, Limón ML, et al. Perioperative trastuzumab, capecitabine and oxaliplatin in patients with HER2-positive resectable gastric or gastro-oesophageal junction

adenocarcinoma: NEOHX phase II trial. European Journal of Cancer. 2021;145:158-67.

Table 1: Patient characteristics and analytical cohorts

	MOC		MBOT		Upper GI[a]		Lower GI[b]	
Clinical and gene expression data n=604	**333**		**151**		**65**		**55**	
Age at diagnosis (years)								
Median	53		47		67		68	
Range	18-95		18-91		31-85		39-89	
	n	% of known	n	% of known	n	% of known	n	% of known
Stage								
I	206	78%	98	95%	5	8%	8	18%
II	20	8%	2	2%	26	41%	11	25%
III	31	12%	3	3%	26	41%	15	34%
IV	7	3%	0	0%	6	10%	10	23%
unknown	69		48		2		11	
Sex								
Female	333	100%	151	100%	33	51%	47	85%
Male	0	0%	0	0%	32	49%	8	15%
Grade								
1	136	46%		n/a	3	5%	12	23%
2	113	38%		n/a	44	72%	32	62%
3	46	16%		n/a	14	23%	8	15%
unknown	38			n/a	4		3	
Residual disease								
Nil macroscopic	145	86%	62	90%	18	82%	25	84%
Yes	24	14%	7	10%	4	18%	7	25%
Unknown	164		82		43		23	

MOC Mucinous Ovarian Carcinoma; MBOT Mucinous Borderline Ovarian Tumor; GI Gastrointestinal

[a]Pancreas cancer (n=57), gastric cancer (n=5), upper GI metastases, unknown primary (n=3)

[b]Colorectal cancer (n=36), appendiceal cancer (n=15); lower GI metastases, unknown primary (n=4)

Table 2: Overall survival in mucinous ovarian carcinoma by pattern of invasion

				univariate		multivariable*	
ALL STAGES		n	deaths	HR (95% CI)	p	HR (95% CI)	p
	Expansile	138	45	Ref		Ref	
	Infiltrative	**40**	**23**	**2.20 (1.33-3.64)**	**0.002**	**1.86 (1.02-3.42)**	**0.044**
Stratification by time							
0-2 years	Expansile	138	25	Ref		Ref	
	Infiltrative	**40**	**19**	**3.12 (1.72-5.67)**	**1.88E-04**	**3.06 (1.49-6.29)**	**0.002**
>2 years	Expansile	101	20	Ref		Ref	
	Infiltrative	20	4	0.71 (0.21-2.39)	0.580	0.45 (0.10-2.03)	0.297
STAGE I	Expansile	109	28	Ref		Ref	
	Infiltrative	25	10	1.52 (0.71-3.21)	0.278	1.40 (0.59-3.33)	0.447
Stratification by time							
0-2 years	Expansile	108	14	Ref		Ref	
	Infiltrative	24	**8**	**2.67 (1.12-6.37)**	**0.027**	**2.77 (1.04-7.41)**	**0.042**
>2 years	Expansile	87	14	Ref		Ref	
	Infiltrative	16	2	0.34 (0.04-2.60)	0.299	0.34 (0.04-2.69)	0.309

*All stage multivariable analysis adjusted for age, stage, stratified by site; Stage I multivariable analysis adjusted for age and site

Table 3: Associations between gene expression and stage group and overall survival in mucinous ovarian carcinoma

Gene	Mean expression by stage group			Univariate			Multivariable		
	I/II (n=189)	III/IV (n=36)	p	n	HR (95% CI)	p	n	HR (95% CI)	p
THBS2	2.49	3.29	**0.009**	179	1.42 (1.21-1.68)	**2.91E-05**	179	1.25 (1.04-1.51)	**0.016**
TAGLN	1.93	2.59	**0.034**	179	1.36 (1.16-1.58)	**9.87E-05**	179	1.21 (1.01-1.45)	**0.043**
DCN	1.7	2.16	0.168	179	1.24 (1.07-1.44)	**0.005**	179	1.05 (0.88-1.26)	0.584
PLA2R1	-1.96	-1.69	0.208	179	1.33 (1.06-1.67)	**0.015**	179	1.28 (0.97-1.69)	0.082
ERBB2	1.18	-0.06	**0.003**	233	0.87 (0.78-0.98)	**0.019**	224	0.99 (0.87-1.13)	0.921
ANXA10	-1.28	-2.38	0.034	233	0.95 (0.89-1.01)	0.107			
SATB2	-0.28	0.22	**0.007**	233	1.12 (0.89-1.40)	0.34			
PD-L1 (CD274)	-1.23	-1.06	0.627	233	1.07 (0.93-1.24)	0.341			
CK20 (KRT20)	-0.45	-1.31	0.019	233	1.03 (0.97-1.10)	0.354			
VSIG1	-0.13	-0.84	0.329	233	0.97 (0.91-1.04)	0.455			
MUC2	1.3	1.17	0.829	233	1.03 (0.95-0.46)	0.458			
PGC	0.13	0.36	0.782	233	0.98 (0.93-1.03)	0.5			
MUC16	-2.78	-1.77	0.130	233	1.02 (0.96-1.08)	0.567			
CLDN18	-1.79	-2.24	0.403	233	0.98 (0.92-1.05)	0.649			
MEP1A	0.07	-0.05	0.851	233	0.98 (0.91-1.06)	0.69			
GKN1	-0.06	0.06	0.294	233	1.07 (0.73-1.57)	0.721			
TYMS	0.1	0.02	0.658	233	1.04 (0.81-1.34)	0.746			
MUC5AC	0.44	-0.07	0.461	233	1.00 (0.94-1.07)	0.941			
PD-1 (PDCD1)	-1.25	-0.82	0.138	233	1.01 (0.86-1.17)	0.943			

Difference in expression between stage groups Student's t-test; HR hazard ratio, 95% confidence interval; Multivariable analysis adjusted for age and stage, stratified by study site

Figure legends

Figure 1: Schema of study numbers for each analysis to describe different cohort numbers due to pathology review and missing data. Mucinous Ovarian Carcinoma (MOC); Mucinous Borderline Ovarian Tumor (MBOT); Lower Gastrointestinal (LGI); Upper Gastrointestinal (UGI); Silver *in situ* hybridisation (SISH); Immunohistochemistry (IHC).

Figure 2: Kaplan-Meir curves of overall survival in A) main tumor groups (n=582) Mucinous Borderline Ovarian Tumor (MBOT), Mucinous Ovarian Carcinoma (MOC), Lower Gastrointestinal (LGI), and Upper Gastrointestinal (UGI); B) patients with MOC by FIGO Stage (n=184); C) patients with MOC by pattern of invasion in all stages (n=178); D) patients with Stage I MOC (n=134) by pattern of invasion.

Figure 3: Heatmap of unsupervised clustering analysis. Contains all samples with a concordant pathology diagnosis (n=497), with MOC grouped by FIGO Stage. Labels show main clusters and diagnoses. Gene expression values are normalized and logarithm base 2 transformed. Diagnosis (dx), Mucinous Ovarian Carcinoma (MOC), Mucinous Borderline Ovarian Tumor (MBOT), Lower Gastrointestinal (LGI).

Figure 1

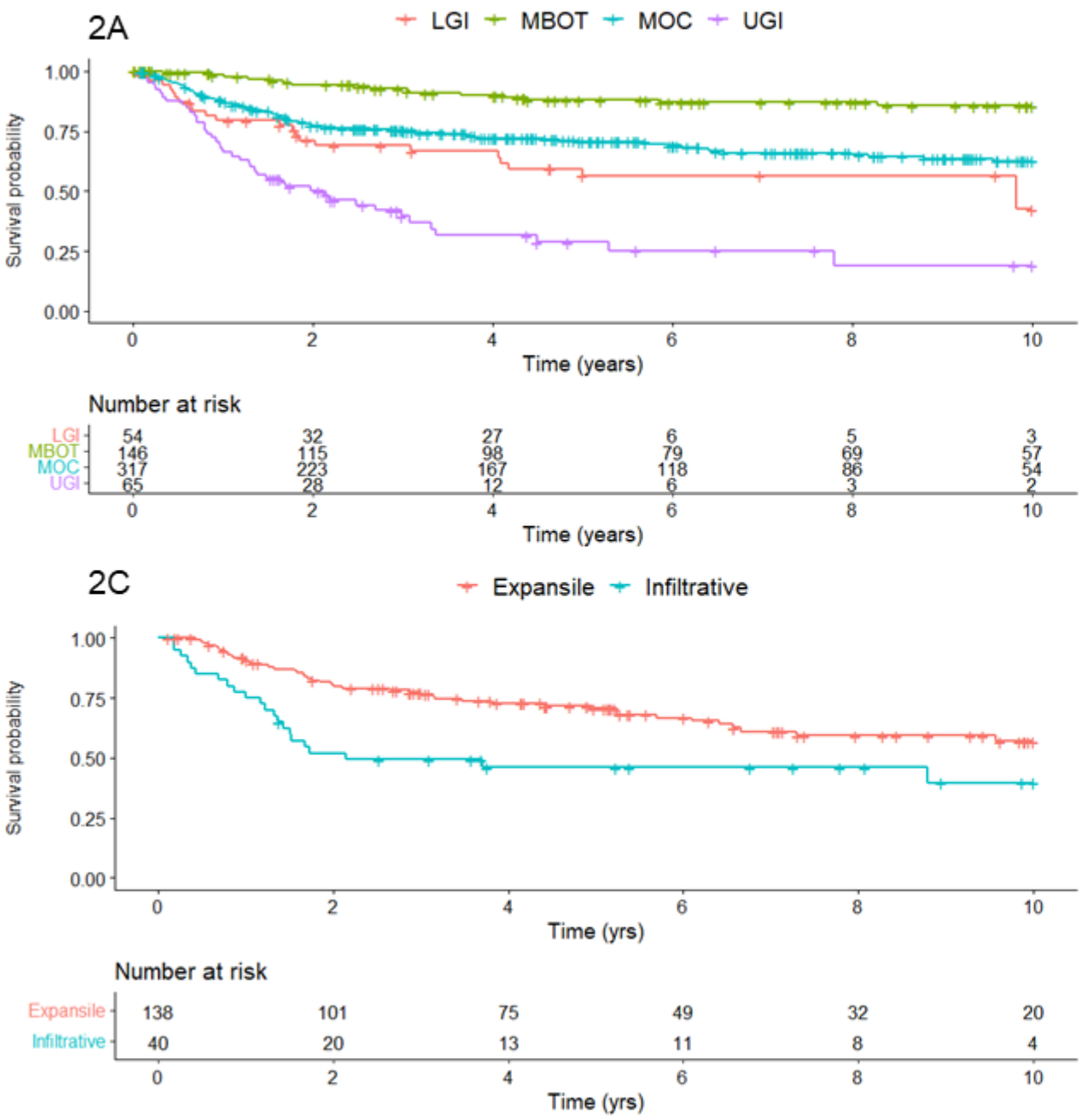
	Analysis	Tumor group	n[a]	Exclusions
PROGNOSTIC	Survival by tumor group[b] By MOC stage[c]	All MOC	582 184	Missing survival data Missing MOC FIGO stage
	Pattern of invasion MOC[c] MOC stage I only[c]	MOC	208 178	Missing survival data Missing MOC FIGO stage
	Gene expression and survival	MOC UGI LGI	233 65 55	Non-invasive (MBOT) Missing survival data MOC missing stage group (I/II vs. III/IV) MOC from discovery dataset (n=54)
DIAGNOSTIC	Clustering & predictions[c]	All	479	MOC missing stage group (I/II vs. III/IV)
	Expression by tumor groups[c]	MOC/UGI/LGI	363	Non-invasive
	ERBB2/HER2 invasive[c] MOC and MBOT SISH data mRNA vs IHC	MOC/UGI/LGI MOC/MBOT MOC MOC	363 377 37 94	Non-invasive Missing MOC FIGO stage
	Discordant diagnosis predictions[b]	MOC/MBOT Discordant	377 90	Discordant MBOT with benign tissue
	Compare paired samples[b]	All	33	Single samples

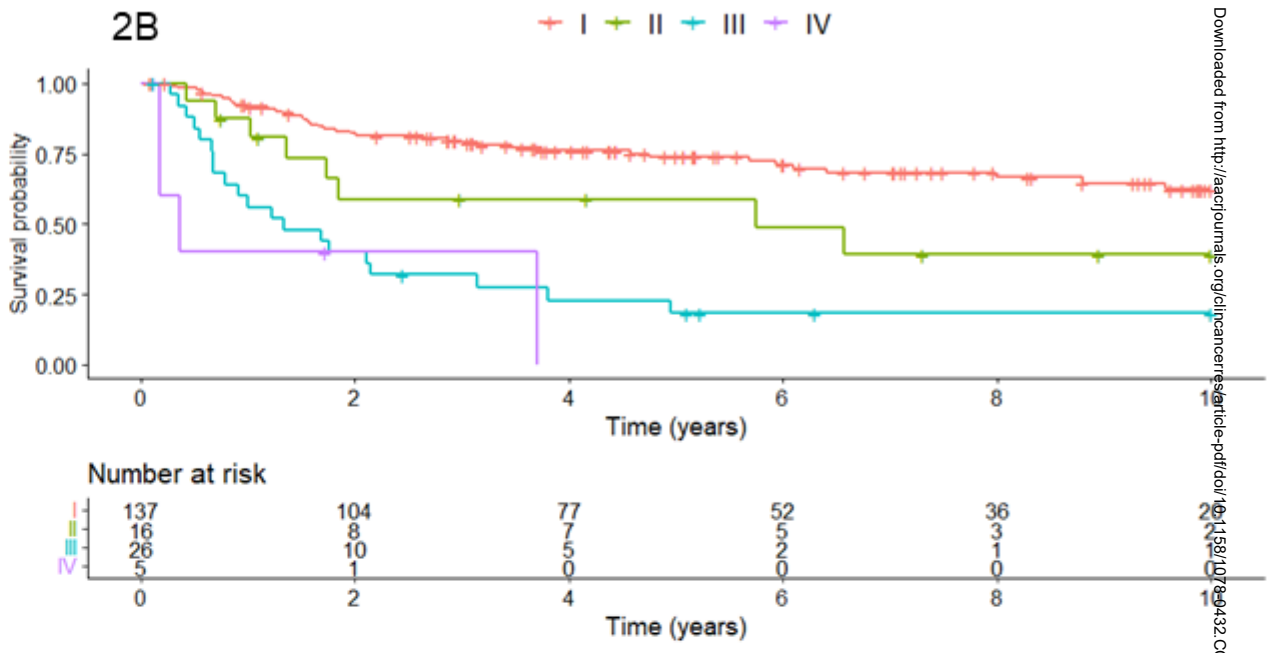
[a] Final number of cases included in each analysis

[b] Cases with concordant diagnosis on review, including those where the submitted block was not representative of the overall diagnosis.

[c] Only cases with both a concordant diagnosis on review, and a submitted block containing tissue representative of that diagnosis.

Figure 2





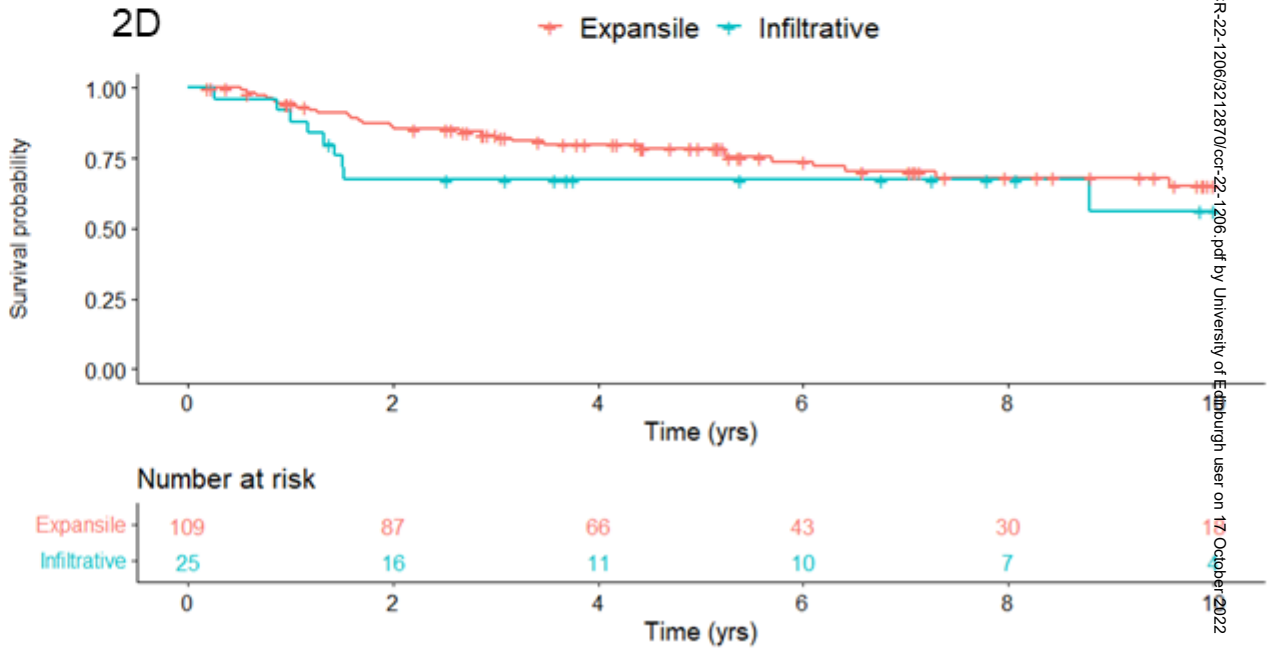


Figure 3