Development and validation of a microRNA-based signature (MiROvaR) to predict early relapse or progression of epithelial ovarian cancer: a cohort study

Marina Bagnoli, Silvana Canevari*, Daniela Califano, Simona Losito, Massimo Di Maio, Francesco Raspagliesi, Maria Luisa Carcangiu, Giuseppe Toffoli, Erika Cecchin, Roberto Sorio, Vincenzo Canzonieri, Daniela Russo, Giosué Scognamiglio, Gennaro Chiappetta, Gustavo Baldassarre, Domenica Lorusso, Giovanni Scambia, Gian Franco Zannoni, Antonella Savarese, Mariantonia Carosi, Paolo Scollo, Enrico Breda, Viviana Murgia, Francesco Perrone, Sandro Pignata†, Loris De Cecco†, Delia Mezzanzanica†, for the Multicentre Italian Trials in Ovarian cancer (MITO) translational group*

Summary

Background **Risk of relapse or progression remains high in the treatment of most patients with epithelial ovarian cancer, and development of a molecular predictor could be a valuable tool for stratification of patients by risk. We aimed to develop a microRNA (miRNA)-based molecular classifier that can predict risk of progression or relapse in patients with epithelial ovarian cancer.**

Methods **We analysed miRNA expression profiles in three cohorts of samples collected at diagnosis. We used 179 samples from a Multicenter Italian Trial in Ovarian cancer trial (cohort OC179) to develop the model and 263 samples from two cancer centres (cohort OC263) and 452 samples from The Cancer Genome Atlas epithelial ovarian cancer series (cohort OC452) to validate the model. The primary clinical endpoint was progression-free survival, and we adapted a semi-supervised prediction method to the miRNA expression profile of OC179 to identify miRNAs that predict risk of progression. We assessed the independent prognostic role of the model using multivariable analysis with a Cox regression model.**

Findings **We identified 35 miRNAs that predicted risk of progression or relapse and used them to create a prognostic model, the 35-miRNA-based predictor of Risk of Ovarian Cancer Relapse or progression (MiROvaR). MiROvaR was able to classify patients in OC179 into a high-risk group (89 patients; median progression-free survival 18 months [95% CI 15–22]) and a low-risk group (90 patients; median progression-free survival 38 months [24–not estimable]; hazard ratio [HR] 1·85 [1·29–2·64], p=0·00082). MiROvaR was a significant predictor of progression in the two validation sets (OC263 HR 3·16, 95% CI 2·33–4·29, p<0·0001; OC452 HR 1·39, 95% CI 1·11–1·74, p=0·0047) and maintained its independent prognostic effect when adjusted for relevant clinical covariates using multivariable analyses (OC179: adjusted HR 1·48, 95% CI 1·03–2·13, p=0·036; OC263: adjusted HR 3·09 [2·24–4·28], p<0·0001; and OC452: HR 1·41 [1·11–1·79], p=0·0047).**

Interpretation **MiROvaR is a potential predictor of epithelial ovarian cancer progression and has prognostic value independent of relevant clinical covariates. MiROvaR warrants further investigation for the development of a clinical-grade prognostic assay.**

Funding **AIRC and CARIPLO Foundation.**

Introduction

Epithelial ovarian cancer is a life-threatening disease characterised by late-stage presentation and high pathological and molecular heterogeneity.[1] The standard treatment for epithelial ovarian cancer is aggressive primary surgery followed by platinum-based chemotherapy. However, the risk of relapse is high, even in patients who achieve a pathological complete response, and most patients develop platinum-resistant progressive disease which restricts available therapeutic options. Despite impressive advances in surgical approaches and drugs for epithelial ovarian cancer, overall survival has improved little during the past 30 years,[2] and 5-year survival for patients with advanced disease remains about 30%.[3] Substantial efforts have been made to develop gene expression-based molecular signatures to predict the prognosis of epithelial ovarian cancer, but few molecular prognostic classifiers have been developed,[4–9] even fewer have been externally validated, and none are clinically available. One reason for this shortage of prognostic tools is the fact that epithelial ovarian cancer is a genetically plastic disease, which evolves during progression and is highly heterogeneous at the time of initial diagnosis. To find a way to predict risk of progression for patients with epithelial ovarian cancer, we decided to focus our attention on microRNAs (miRNAs) as miRNAs act as a master layer of regulation for gene expression but their number is at least one order of magnitude lower than that of genes. Despite studies of miRNA profiles in epithelial ovarian cancer having been

See **Comment** page 1032

*Contributed equally as first authors

†Contributed equally as senior authors

Molecular Therapies Unit, Department of Experimental Oncology and Molecular Medicine (M Bagnoli PhD, D Mezzanzanica PhD), **Functional Genomics, Department of Experimental Oncology and Molecular Medicine** (S Canevari PhD, L De Cecco PhD), **Unit of Gynaecological Oncology** (F Raspagliesi MD, D Lorusso MD), **and Anatomic Pathology 1 Unit** (M L Carcangiu MD), **Fondazione IRCCS Istituto Nazionale dei Tumori, Milan, Italy; Functional Genomic Unit** (D Califano PhD, G Chiappetta PhD, D Russo PhD), **Surgical Pathology Unit** (S Losito MD, G Scognamiglio PhD), **Clinical Trials Unit** (Prof M Di Maio MD, F Perrone MD), **and Department of Urogynaecological Oncology** (S Pignata MD), **Istituto Nazionale per lo Studio e la Cura dei Tumori "Fondazione G Pascale", IRCCS, Naples, Italy; Experimental and Clinical Pharmacology Unit** (G Toffoli MD, E Cecchin PhD), **Medical Oncology C** (R Sorio MD), **Unit of Pathology** (V Canzonieri MD), **and Division of Experimental Oncology 2** (G Baldassarre MD), **Centro di Riferimento Oncologico, Istituto Ricovero e Cura Carattere Scientifico (CRO-IRCCS), Aviano, Italy; Department of Obstetrics and Gynecology, Gynecologic**

Oncology Unit (Prof G Scambia MD) **and Department of Human Pathology, Division of Gynecologic Pathology** (Prof G F Zannoni MD)**, Catholic University of the Sacred Heart, Rome, Italy; Division of Medical Oncology 1** (A Savarese, MD) **and Division of Pathology** (M Carosi MD)**, Regina Elena Cancer Institute, Rome, Italy; Department of Obstetrics and Gynecology, Azienda Ospedaliera Cannizzaro, Catania, Italy** (P Scollo MD)**; Medical Oncology Unit Ospedale S Giovanni Calibita Fatebenefratelli, Rome, Italy** (E Breda MD)**; and Medical Oncology Unit Ospedale S Chiara, Trento, Italy** (V Murgia MD)

Correspondence to: Dr Delia Mezzanzanica, Unit of Molecular Therapies, Department of Experimental Oncology and Molecular Medicine, Fondazione IRCCS Istituto Nazionale dei Tumori, Milan 20133, Italy **delia.mezzanzanica@istitutotumori.mi.it**

Research in context

Evidence before this study

We systematically searched PubMed up to Jan 11, 2016, for research articles containing the terms "miRNA AND predictive OR predictor AND ovarian cancer AND progression OR relapse", without a date or language restrictions. Our search did not identify any previous high-throughput studies that had investigated the potential predictive role of microRNA (miRNA) expression profiles in epithelial ovarian cancer tissues.

Added value of this study

We did a multicentre retrospective study to test the ability of miRNA expression in primary epithelial ovarian cancer to predict the risk of disease relapse or progression at diagnosis. 894 epithelial ovarian cancer cases were analysed for miRNA expression, which we believe to be the largest collection studied so far. We developed MiROvaR, a molecular predictor for disease relapse or progression that is based on the expression of 35 miRNAs. MiROvaR was able to classify patients with epithelial ovarian cancer as being at high or low risk of relapse or progression, irrespective of the tumour characteristics at presentation. We validated the performance of MiROvaR in two independent datasets and MiROvaR maintained its independent prognostic effect after adjustment for the two strongest prognostic clinical variables so far available for epithelial ovarian cancer: disease stage and residual disease after primary surgery. To our knowledge, this is the first study to investigate the effectiveness of a miRNA-based predictor for ovarian cancer relapse, and our findings suggest that MiROvaR is a promising and valuable classifier for the identification of patients at high risk of relapse or progression.

Implications of all the available evidence

Our findings show the potential use of miRNAs, a master layer of gene expression regulation, to more precisely classify patients according to their risk of progression. The effectiveness of MiROvaR warrants further investigation for the development of a useful clinical-grade assay.

published,[10–12] no data are available on their use in the classification of risk of relapse or progression in this setting. We aimed to analyse miRNA expression profiles from epithelial ovarian cancer samples to develop and validate a miRNA-based predictor of risk of relapse or progression.

Methods

Study design and participants

In total, we used three treatment-naive cohorts of samples from women with epithelial ovarian cancer for this study: OC179 (training set), OC263 (validation set), and OC452 (validation set). The training set, OC179, was derived from the MITO-2 clinical trial (NCT00326456).[13] Blocks of formalin-fixed, paraffin-embedded tissue from the primary tumours of patients with epithelial ovarian cancer enrolled in the MITO-2 trial[13] were provided by 17 (40%) of 43 centres participating in the trial. Of 549 samples, 305 (56%) were excluded because they had insufficient amounts of tumour tissue for RNA extraction, 24 because of the poor quality of the extracted RNA, ten because the tumours derived from peritoneal lesions, 30 because they were not chemotherapy naive (samples were collected at interval debulking surgery after three cycles of chemotherapy), and one because of miRNA hybridisation failure. After clinical–pathological review of the available paraffin blocks, RNA extraction, quality control, and profiling on human miRNA arrays, 179 samples were eligible for data analysis (appendix pp 2, 10).

We used two independent series, OC263 (collected at the Istituto Nazionale dei Tumori [Milan, Italy] and the Centro di Riferimento Oncologico [Aviano, Italy]) and OC452 (data collected from the epithelial ovarian cancer dataset of The Cancer Genome Atlas [TCGA]), as validation sets. The OC263 study population consisted of all data for 130 previously profiled patients with epithelial ovarian cancer[14] (the OC130 cohort) and 133 tissue samples collected at Centro di Riferimento Oncologico Aviano (OC133 cohort); both cohorts were profiled with the Illumina microchip platform (Illumina, San Diego, CA, USA) at Istituto Nazionale dei Tumori. All experimental and clinical data for OC452[5] and OC130[14] have been reported previously and are publicly accessible through the Gene Expression Omnibus for OC130 (superseries GSE25204, which includes GSE25202,[14] GSE25203,[14] and GSE67819[14]) and TCGA website for OC452. For OC133, freshly frozen tumour samples were collected at the Centro di Riferimento Oncologico from patients with primary epithelial ovarian cancer who underwent surgical resection before receiving any chemotherapy. All clinical data and complete follow-up information for patients included in OC133 were available from Centro di Riferimento Oncologico. Tumours were staged in accordance with the FIGO criteria and graded with the WHO criteria. A pathologist (VC) with expertise in gynaecological pathology reviewed all pathological data from patients in OC133, confirming the pathological diagnosis and the presence of the required representative percentage of tumour cells in each sample. None of the tumour samples included in our analysis were macrodissected. Tumour cellularity (ie, the percentage of tumour cells present in the sample) was similar between the three cohorts OC179, OC263, and OC452. OC452 included samples with at least 70% tumour cellularity and less than 20% necrosis and the same sample criteria were used for the OC179 training cohort and OC263 validation cohort.

Signed informed consent was obtained from all patients included in the study. For both OC179 and OC263, the investigation was approved by the

For **The Cancer Genome Atlas website** see https://tcga-data.nci.nih.gov/tcga/

institutional review boards of participating institutions. In the case of OC179, which was derived from the MITO-2 trial,[13] samples were collected for translational research following the approval of a study amendment in 2011. For OC263, the study was approved by the Independent Ethics Committee of the Istituto Nazionale dei Tumori, where the miRNA profiling was done. Ethics approval for OC452 has been reported previously.[15]

	Reference	Type of material	Number of samples	miRNA platform
OC179 (training cohort)	Present study	FFPE	179	Agilent miR-Base 17
OC263 (validation cohort)				
OC130	Bagnoli et al[14]	Frozen/FFPE	130	Illumina miR-Base 12
OC133	Present study	Frozen	133	Illumina miR-Base 12
OC452 (validation cohort)	Kang et al[5]	Frozen	452	Agilent miR-Base 10

miRNA=microRNA. FFPE=formalin-fixed paraffin-embedded samples.

Table 1: **Summary of cohorts, miRNA platforms, and chip arrays used for model development**

Procedures

The RNA extraction and quality control procedures for OC179 and the OC133 subset of the OC263 validation set and the miRNA expression profiling procedures for all the cohorts analysed are described in detail in the appendix (p 2) and summarised in table 1. Samples in OC263 were collected in our institutions and included both frozen and formalin-fixed paraffin-embedded samples, OC452 contained only frozen samples, and OC179 only formalin-fixed paraffin-embedded samples. We already shown that data obtained from frozen samples could be reproduced in formalin-fixed paraffin-embedded samples and vice versa.[14] The inter-platform reproducibility of miRNA microarray profiles was shown in our previous study.[16] We have deposited the Minimum Information About a Microarray Experiment (MIAME)-compliant data reported in this study into the NCBI Gene Expression Omnibus[17] and they are accessible with the superseries accession number GSE73583, which includes GSE73581 (OC179) and GSE73582 (OC133).

Data were analysed with R (version 3.1.0) and the R Bioconductor packages.[18]

We integrated the four datasets in the OC263 dataset (OC133 [GSE73582], GSE25202,[14] GSE25203,[14] and GSE67819[14]) with the virtualArray R BioConductor package version 1.2.1.[19] The datasets were produced with the same version of the microarray chip (Human v2 miRNA panel; Illumina, San Diego, CA, USA), which identified 1146 miRNAs annotated on miRBase version 12.0. We applied the ComBat algorithm[20] to the normalised and log2-transformed data matrices to reduce the likelihood of batch effects from non-biological technical biases.

OC452 contains the ovarian cancer miRNA microarray profile from the TCGA consortium. We downloaded the level 1 raw data and the clinical annotations from the TCGA website on Nov 8, 2014. miRNA expression was profiled with 8x15K Human miRNA Microarray Kits (Agilent Technologies, Santa Clara, CA, USA), which identified 799 miRNAs annotated on miRBase version 10. Data were normalised with the robust multiarray average algorithm, log2-transformed, and filtered with the AgiMicroRna R package version 1.4.0 as described for the training set in the appendix (p 2).

To create a prognostic model from the miRNA expression profiles, we adapted a semi-supervised prediction method involving principal component analysis, which was developed by Bair and Tibshirani[21] (available through the R package superpc version 1.09) and has been successfully applied to transcriptome data.[22] Briefly, a univariate Cox proportional hazards regression analysis of progression-free survival versus the miRNA log expression level was used to establish the association between each miRNA and progression-free survival of OC179. The miRNAs that we entered into the model were not fixed a priori, but were selected on the basis of their false discovery rate, which defines the expected proportion of false positive results and allows for balancing of the competing demands of sensitivity and specificity to avoid data overfitting. We used a false discovery rate less than 0·1, corresponding to an α less than 0·025. Subsequently, we used the principal component analysis to reduce the dimensionality (number of variables, ie, miRNAs) of the miRNAs included in the model. The first two principal components, which captured 74% of miRNA expression variability, were used to obtain a regression coefficient (weight) for each miRNA and to develop a model to calculate the prognostic risk score. We used a ten-fold cross-validation approach (known as internal validation)[23] to classify the OC179 samples as being at low or high risk of progression or relapse. This approach to risk classification used the median index values obtained from 90% of the cases (training set) to classify the remaining 10% of the omitted cases (test set) according to this median value. All cases were stratified after the entire procedure had been reiterated ten times, with a different 10% of cases omitted each time until each case had been excluded. The miRNAs entered into the different cross-validation sets were used in the model based on the percentage of cross-validation support. The cross-validation support assesses the percentage of iterations in which the specific miRNA was selected in the ten-cross-validated set.

The prognostic risk index for each patient can be calculated with the following formula:

$$\sum_i w_i\, x_i + 3 \cdot 196617$$

where w_i is the weight and x_i is the logarithmically transformed miRNA expression of the i-th miRNA. A sample was predicted to be at high risk if its prognostic index was greater than 0·07359 (low risk if index ≤0·07359), which is the median value obtained in the cross-validation of the OC179 group.

We assessed the ability of the model to predict progression-free survival with Kaplan-Meier curves and the log-rank test. Following a procedure known as random shuffling, we did a 1000-permutation test on the OC179 dataset to assess the degree of overfitting in the development of our prognostic model.[24] The survival data were randomly reassigned among the cases and the entire survival risk prediction process was repeated, to assess the null distribution of the log-rank test. The tail area of the null distribution beyond the log-rank statistic of the real data estimates the significance of the association for each permutation to test the null hypothesis, which is the absence of an association between progression-free survival and miRNA expression.[25] The 1000-permutation test reached p=0·0010. We assessed the prognostic performance of the model with receiver operating characteristic curves (appendix p 2).

Outcomes

The primary endpoint was progression-free survival because the main goal of the predictor was to be able to identify patients at risk of early relapse. Progression-free survival was defined as the time (in months) between the date of random assignment for OC179 (to maintain consistency with trial data) or surgery (OC263 and OC452) and the date of progression or death, whichever occurred first, or the date of last follow-up for patients alive without progression (appendix p 2).

The secondary endpoint was overall survival, which was defined as the time (in months) between the date of random assignment (for OC179) or surgery (for the validation sets) and the date of death or the date of last contact for surviving patients.

Statistical analysis

Progression-free survival and overall survival curves were plotted with the Kaplan-Meier method and were compared with the log-rank test. Median estimates with 95% CIs are also reported. We used a Cox univariate model to estimate the hazard ratio (HR) for each relevant prognostic variable.

We used multivariable analysis with a Cox regression model to assess the prognostic effect of our model in the context of concomitant effects of other known prognostic factors (ie, stage and residual disease). We tested the validity of the proportional hazards assumption for a Cox model fit via evaluation of scaled Schoenfeld residuals. The choice of the covariates used in the model was based on several criteria. First, we tried to select variables with high known prognostic value in terms of progression-free survival, to give a number of covariates that was not too high and adequate to the sample size of the OC179 training set. Second, we tried to avoid variables with very small subgroups (eg, grading, based on current definitions[26]). Third, we avoided subjective variables (eg, performance status). Fourth, we avoided variables that were not available in the validation sets (eg, performance status was not available in OC263 and histology was not informative in OC452, which included only high-grade serous tumours). We did not consider age a prognostic factor.

Consequently, the only covariates that we used for multivariable analyses were FIGO stage and residual disease after primary surgery. Based on the extent of residual disease after primary surgery, patients were divided into three groups: no evident residual disease, minimum residual disease (residual tumour <1 cm), and gross residual disease (residual tumour larger ≥1 cm). Patients were then classified into two categories for further analysis: optimum debulking (includes patients with no evident residual disease or with minimum residual disease) and suboptimum debulking (residual tumour ≥1 cm). We choose to compare optimum debulking *vs* suboptimum debulking to be consistent with the study reporting the final analysis of the MITO-2 clinical trial[13] and to avoid small subgroups that might derive from use of three categories. For both univariate and multivariable analyses, stage and surgical debulking were coded as dichotomous indicator variables (stage III or IV *vs* stage I or II, suboptimum debulking *vs* optimum debulking). However, we also did a multivariable analysis of progression-free survival with residual disease coded as a three-level categorical variable (no evident residual disease, minimum residual disease, and gross residual disease; appendix pp 3 and 6). In the OC263 validation set, we also separately analysed: first, type II tumours,[27] which include high-grade serous and high-grade endometrioid ovarian cancer, undifferentiated tumours, and malignant mixed Müllerian tumours, and second, high-grade serous ovarian cancer, for a better comparison with the OC452 validation set, which included high-grade serous ovarian cancer only. We then grouped patients based on similar clinical and pathological characteristics and we used the χ^2 test to analyse the distribution of high-risk and low-risk patients (as categorised with our model) in relation to clinical and pathological variables. We deemed a p value less than 0·05 to show statistical significance.

We did all statistical analyses with the R statistical language version 3.1.0 and we created graphs with GraphPad PRISM version 5.02.

Role of the funding source

The funder of the study had no role in study design, data collection, data analysis, data interpretation, or writing of the report. MB, SC, MDM, FP, SP, LDC, and DM had access to the raw data and the corresponding author had full access to all the data in the study and had final responsibility for the decision to submit for publication.

Results

The cohorts used to develop our predictive model are summarised in table 1. Table 2 shows the characteristics of the patients included in our study; Kaplan-Meier curves of progression-free survival and overall survival for each cohort are shown in the appendix (p 11).

Compared with the overall MITO-2 trial population, the sub-population included in the OC179 cohort contained slightly fewer patients who were not operated at baseline or had stage IV disease according to the International Federation of Gynecology and Obstetrics (FIGO) criteria at diagnosis (see appendix p 4). Accordingly, the population included in OC179 had longer progression-free survival (22·8 months, 95% CI 18·2–29·4 *vs* 17·1, 16·0–19·4) and overall survival (median not reached, 63–not estimable (NE) *vs* 56·6, 50·0–68·2) than did the overall MITO-2 population.

The proportion of cases with death as the progression-free survival event was four (3%) of 124 events for the OC179 training set, seven (4%) of 195 events for the OC263 validation set, and two (<1%) of 327 events for the OC452 validation set 2. For patients in OC179, the mean time between randomisation and surgery was 1 month (SD 0·46 months, range 0·23–2·56).

Overall, 894 patients with epithelial ovarian cancer of varying stage and histotype were analysed at the time of diagnosis: all patients in the OC452 cohort had serous tumours, the OC179 and OC263 cohorts had patients with various tumour histotypes (table 2). We used data pre-processing to enable comparison among the different platforms and chip arrays used for miRNA profiling of the three cohorts. The three cohorts were separately filtered to exclude miRNAs that were not detectable in all samples and we obtained data matrices of 921 miRNAs from OC179, 706 miRNAs from OC263, and 661 miRNAs from OC452. Since each dataset was designed based on different miRBase releases, each platform was re-annotated on miRBase release 21.0 at the sequence level. Putative miRNAs, sequences

	OC179 (n=179)	OC263 (n=263)	OC452 (n=452)
Age (years)	58 (28–78)	55 (25–85)	59 (26–87)
Histotypes			
Serous	124 (69%)	190 (72%)	452 (100%)
Undifferentiated	10 (6%)	23 (9%)	NA
Endometroid	24 (13%)	26 (10%)	NA
Mucinous	0	1 (<1%)	NA
Clear cells	6 (3%)	7 (3%)	NA
Others and mixed	13 (7%)	15 (6%)	NA
Missing information	2 (1%)	1 (<1%)	NA
Stage (FIGO)			
I	17 (9%)	16 (6%)	11 (2%)
II	15 (8%)	9 (3%)	27 (6%)
III	123 (69%)	212 (81%)	350 (77%)
IV	24 (13%)	26 (10%)	63 (14%)
Missing information	0	0	1 (<1%)
Grade			
Borderline	0	3 (1%)	1 (<1%)
1 (well differentiated)	5 (3%)	7 (3%)	5 (1%)
2 (moderately differentiated)	27 (15%)	51 (19%)	55 (12%)
3 (poorly differentiated)	126 (70%)	177 (67%)	382 (85%)
Undifferentiated	10 (6%)	23 (9%)	0
GX	0	0	8 (2%)
Missing information	11 (6%)	2 (1%)	1 (<1%)
Subtype			
Type I	11 (6%)	17 (6%)	6 (1%)
Type II	144 (81%)	230 (87%)	437 (97%)
Not classified	24 (13)	16 (6%)	9 (2%)
Amount of residual disease			
No evident residual disease	73 (41%)	76 (29%)	102 (23%)
<1 cm (minimum residual disease)	42 (23%)	85 (32%)	208 (46%)
≥1 cm (gross residual disease)	53 (30%)	101 (38%)	100 (22%)
Not operated	11 (6%)	0	0
Missing information	0	1 (<1%)	42 (9%)
Follow-up duration (months)	73 (60–88)	44 (24–71)	56 (25–86)

Data are n (%) or median (IQR). FIGO=International Federation of Gynecology and Obstetrics. GX=grade cannot be established.

Table 2: **Clinical and pathological characteristics of patients included in the three cohorts**

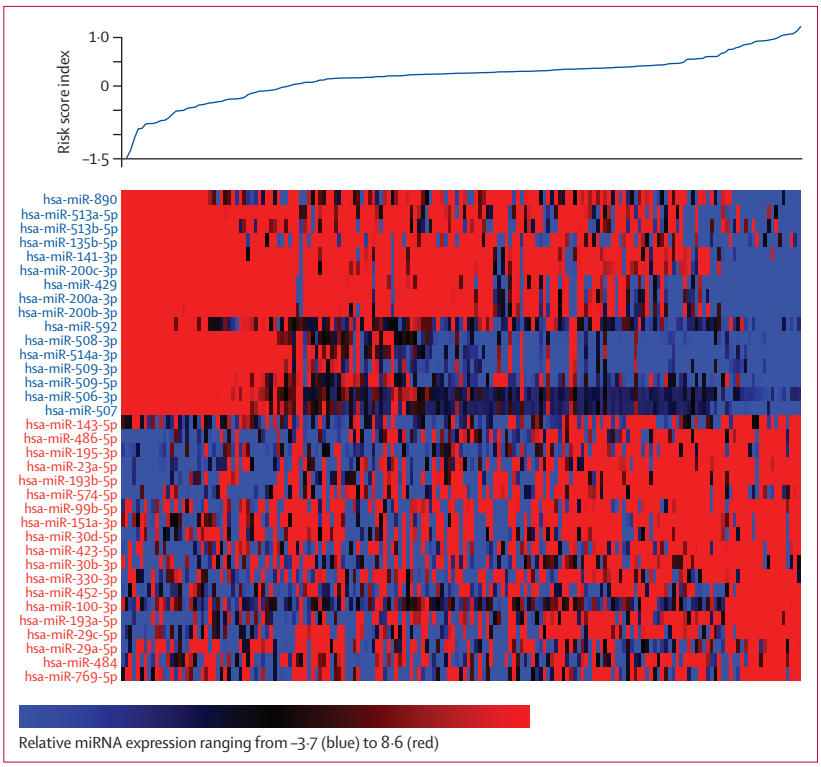

Figure 1: **Expression heat map of miRNAs included in the predictive model**
Each column represents one patient in OC179 (n=179) and each row represents a miRNA included in the model (n=35), sorted on the basis of the established prognostic index. The plot above the heat map shows the specific risk score index for each sample. For the risk score, low index values are associated with low risk and high values are associated with high risk. In the heatmap, blue represents low expression and red represents high expression. Blue text shows miRNAs for which expression is associated with a good prognosis. Red texts shows miRNAs for which expression is associated with a poor prognosis. miRNA=microRNA.

	p value	Cross-validation support	Hazard ratio (95% CI)	Weight (w_i)
Associated with good prognosis				
hsa-miR-141-3p	0·001681	100%	0·819 (0·731–0·918)	–0·032066
hsa-miR-200a-3p	0·003171	100%	0·808 (0·706–0·925)	–0·032221
hsa-miR-200b-3p	0·002689	100%	0·786 (0·678–0·911)	–0·028151
hsa-miR-200c-3p	0·001545	100%	0·793 (0·694–0·905)	–0·027508
hsa-miR-506-3p	<0·0001	100%	0·635 (0·479–0·841)	–0·032425
hsa-miR-507	<0·0001	100%	0·588 (0·429–0·805)	–0·026022
hsa-miR-508-3p	<0·0001	100%	0·747 (0·637–0·874)	–0·045965
hsa-miR-509-3p	<0·0001	100%	0·783 (0·685–0·895)	–0·049717
hsa-miR-509-5p	<0·0001	100%	0·684 (0·555–0·843)	–0·035031
hsa-miR-513a-5p	0·000736	100%	0·766 (0·662–0·886)	–0·021663
hsa-miR-513b-5p	0·000723	100%	0·817 (0·732–0·912)	–0·028496
hsa-miR-514a-3p	<0·0001	100%	0·811 (0·726–0·907)	–0·058425
hsa-miR-592	0·000155	100%	0·255 (0·099–0·661)	–0·002782
hsa-miR-135b-5p	0·008942	80%	0·851 (0·756–0·958)	–0·024577
hsa-miR-429	0·012234	60%	0·835 (0·727–0·958)	–0·030913
hsa-miR-890	0·023114	40%	0·085 (0·010–0·717)	–0·000287
Associated with poor prognosis				
hsa-miR-29c-5p	0·000713	100%	1·595 (1·232–2·065)	0·005566
hsa-miR-193a-5p	<0·0001	100%	1·977 (1·492–2·612)	0·010396
hsa-miR-30b-3p	0·006413	100%	1·983 (1·241–3·165)	0·002938
hsa-miR-486-5p	0·002991	90%	1·345 (1·121–1·612)	0·015239
hsa-miR-423-5p	0·002895	90%	1·765 (1·226–2·537)	0·005948
hsa-miR-100-3p	0·008982	90%	1·958 (1·287–2·974)	0·003563
hsa-miR-484	0·00786	80%	1·6 (1·160–2·206)	0·002136
hsa-miR-23a-5p	0·005207	80%	1·641 (1·181–2·278)	0·006169
hsa-miR-143-5p	0·009584	80%	1·674 (1·183–2·367)	0·00264
hsa-miR-330-3p	0·006058	80%	1·856 (1·206–2·854)	0·004021
hsa-miR-99b-5p	0·00938	70%	1·35 (1·075–1·695)	0·007011
hsa-miR-769-5p	0·008215	70%	1·762 (1·186–2·614)	0·002445
hsa-miR-452-5p	0·017454	60%	1·276 (1·062–1·531)	0·00919
hsa-miR-151a-3p	0·013404	60%	1·363 (1·031–1·512)	0·004522
hsa-miR-193b-5p	0·024076	60%	1·506 (1·059–2·14)	0·005293
hsa-miR-574-5p	0·016105	50%	1·283 (1·049–1·568)	0·005807
hsa-miR-29a-5p	0·017911	50%	1·765 (1·153–2·700)	0·000855
hsa-miR-30d-5p	0·023319	40%	1·253 (1·032–1·52)	0·000766
hsa-miR-195-3p	0·01865	40%	1·629 (1·126–2·356)	0·005412

The miRNAs are identified by their unique miRNA IDs according to miRBase 21.0. Cross-validation support gives an idea of the strength of each miRNA in the signature. Weight (w_i) shows the contribution of the miRNA to the calculation of the risk index, as described in the Methods. miRNA=microRNA.

***Table 3*: miRNAs entered into the prognostic model**

identifying virus miRNAs, non-mature miRNAs, or probe sets unable to distinguish the members of a miRNA family and those discontinued through different miRBase versions were excluded. A list of 385 unique miRNAs (appendix p 5) was obtained and shared among the platforms, then checked with the miRBase Tracker tool[28] to avoid confounding miRNA nomenclature.

On the basis of the defined algorithm and after ten-fold cross-validation on the OC179 data, we developed a model containing 35 unique miRNAs whose expression significantly affected the risk of disease progression (figure 1). Of the 35 miRNAs, 16 were associated with improved prognosis (putative oncosuppressive miRNAs) and 19 with worse prognosis (putative oncogenic miRNAs; figure 1, table 3). We named the 35-miRNA model, miRNA-based predictor of Risk of Ovarian Cancer Relapse or progression (MiROvaR).

When applied to OC179, MiROvaR identified 89 patients at high risk of progression (72 events, median progression-free survival 18 months, 95% CI 15–22) and 90 at low risk (52 events, median progression-free survival 38 months [24–NE]; HR 1·85, 95% CI 1·29–2·64; p=0·00082; figure 2A and table 4). We used ROC analyses to assess the predictive performance of MiROvaR. The average area under the curve (AUC) in the ten-fold cross-validation reached a value of 0·68 (SD 0·02), which supports the good performance of the model in OC179 (appendix p 12).

When tested against the validation sets, MiROvaR was able to stratify patients into risk groups that had significantly different progression-free survival. In OC263, 141 patients were classified as being at high risk of progression (122 events, median progression-free survival 12 months, 95% CI 10–13) and 122 were classified as being at low risk (73 events, median progression-free survival 34 months, 95% CI 26–45; figure 2B). In OC452, 283 patients were classified as high risk (212 events, median progression-free survival 15 months, 95% CI 14–18) and 169 were classified as low risk (115 events, 19 months, 17–27; figure 2C).

We did a univariate analysis of the validation cohorts: in OC263, the HR was 3·16 (95% CI 2·33–4·29, p<0·0001) and in OC452, HR was 1·39 (95% CI 1·11–1·74, p=0·0047; table 4). The AUC was 0·72 (SD 0·01) in OC263 and 0·58 (0·02) in OC452 (appendix p 12). In all three cohorts, advanced stage at diagnosis and suboptimum debulking after primary surgery were significantly associated with progression in univariate analysis (table 4). Importantly, MiROvaR maintained its independent prognostic effect in all cohorts when analysed in multivariable analysis adjusting for these clinical covariates (table 4; appendix p 6). Kaplan-Meier curves of overall survival in patients stratified with MiROvaR are shown for all cohorts in the appendix (pp 7, 13). We found no interactions between MiROvaR and the type of treatment (carboplatin plus taxane *vs* carboplatin plus pegylated doxorubicin) in the OC179 set ($p_{interaction}$=0·62; appendix p 7). The MiROvaR high-risk classification was significantly associated with advanced disease stage (stages III and IV) only in OC179 (p=0·013) and in patients only with residual disease (ie, suboptimum debulking) in OC263 (p=0·0005; appendix p 8).

The OC452 validation set included only high-grade serous ovarian cancer, whereas the OC263 validation set was more heterogeneous in histotype and grading (table 1). We therefore investigated the independent

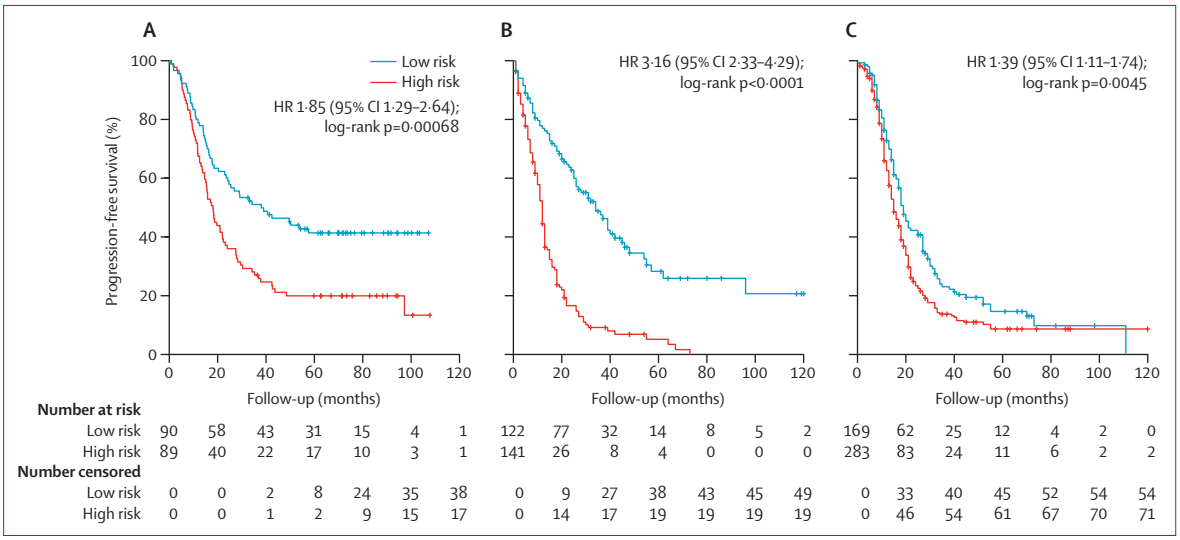

Figure 2: **Progression-free survival stratified by risk according to MiROvaR**
Kaplan-Meier curves show progression-free survival in patients stratified by MiROvaR risk classification in the OC179 training set (A), OC263 validation set (B), and OC452 validation set (C). MiROvaR high-risk and low-risk curves were compared with the log-rank test. miRNA=microRNA. HR=hazard ratio. MiROvaR=miRNA-based predictor of Risk of Ovarian Cancer Relapse or progression.

predictive value of MiROvaR in more homogeneous groups of patients in OC263: the 230 patients with type II tumours, which according to the new proposed classification of epithelial ovarian cancer include high-grade serous, high-grade endometrioid, undifferentiated and malignant mixed Müllerian tumours,[27] (figure 3A and table 5) and the 185 patients with high-grade serous ovarian cancer (figure 3B and table 5). Of the 230 patients with type II epithelial ovarian cancer, 128 were classified as being at high risk of progression (111 events, median progression-free survival 12 months, 95% CI 10–13) and 102 were classified as being at low risk (62 events, 34 months, 95% CI 26–42; figure 3A). Of the 185 patients from OC263 with high-grade serous ovarian cancer, 106 were classified as being at high risk of progression (91 events, median progression-free survival 12 months, 95% CI 11–13) and 79 were classified as being at low risk (50 events, 32 months, 95% CI 25–39; figure 3B). MiROvaR performed well in these two subgroups of patients, with an AUC of 0·72 (SD 0·02) for patients with type II disease and 0·71 (SD 0·01) for patients with high-grade serous ovarian cancer (appendix p 12).

	Univariate analysis		Multivariable analysis	
	HR (95% CI)	p value	HR (95% CI)	p value
OC179 (n=179, events=124)				
Stage III–IV *vs* I–II	4·74 (2·40–9·36)	<0·0001	3·70 (1·83–7·49)	0·00028
Surgical debulking (suboptimum debulking *vs* optimum debulking)	2·10 (1·46–3·00)	<0·0001	1·46 (1·01–2·12)	0·043
miRNA predictor (high *vs* low risk)	1·85 (1·29–2·64)	0·00082	1·48 (1·03–2·13)	0·036
OC263 (n=262, events=194)				
Stage III–IV *vs* I–II	2·16 (1·25–3·73)	0·0057	2·16 (1·21–3·90)	0·0097
Surgical debulking (suboptimum debulking *vs* optimum debulking)	2·23 (1·67–2·97)	<0·0001	1·53 (1·13–2·08)	0·0060
miRNA predictor (high *vs* low risk)	3·16 (2·33–4·29)	<0·0001	3·09 (2·24–4·28)	<0·0001
OC452 (n=409, events=300)				
Stage III–IV *vs* I–II	1·68 (1·02–2·78)	0·04	1·79 (1·04–3·08)	0·035
Surgical debulking (suboptimum debulking *vs* optimum debulking)	1·37 (1·07–1·75)	0·012	1·27 (0·99–1·63)	0·059
miRNA predictor (high *vs* low risk)	1·39 (1·11–1·74)	0·0047	1·41 (1·11–1·79)	0·0047

Univariate and multivariate analyses were done by Cox regression. HR=hazard ratio.

Table 4: **Univariate and multivariable analysis of progression-free survival**

Discussion

A method to identify patients with epithelial ovarian cancer who have very poor prognoses is urgently needed to improve the design of tailored therapy. Our molecular predictor, MiROvaR, was able to stratify patients according to their risk of relapse or progression, identifying groups of patients with significantly different progression-free survival. The development of molecular classifiers such as MiROvaR is based on an a-priori choice of the outcome of interest. Because the main goal of our predictor was to identify patients at risk of early relapse, we chose progression-free survival as the most appropriate endpoint. Progression-free survival is widely accepted as a reasonable endpoint in ovarian cancer[29] both clinically and in drug development, since it is not affected by the heterogeneous mix of second-line treatments used. Although with a less impressive power than in OC179 and OC263, MiROvaR's value was also confirmed in the OC452 dataset, so far the only available public collection of ovarian cancer samples with fully annotated clinical data, which we use as second validation set, supporting MiROvaR's ability to provide clinically significant prognostic information.

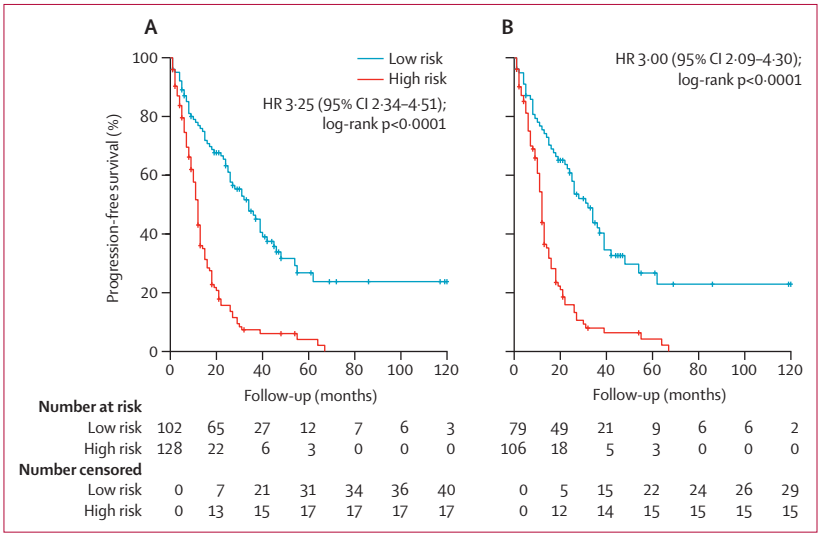

***Figure 3*: Progression-free survival stratified by risk according to MiROvaR in OC263 subpopulations**
Kaplan-Meier curves show progression-free survival in patients in the OC263 validation set stratified by MiROvaR risk classification who have type II epithelial ovarian cancer (A) and high-grade serous ovarian cancer (B). MiROvaR high-risk and low-risk curves were compared with the log-rank test. miRNA=microRNA. HR=hazard ratio. MiROvaR=miRNA-based predictor of Risk of Ovarian Cancer Relapse or progression.

	Univariate analysis		Multivariable analysis	
	HR	p value	HR	p value
All type II epithelial ovarian cancers (n=230, events=172)				
Stage III–IV *vs* I–II	2·45 (1·20–5·00)	0·013	2·37 (1·10–5·12)	0·028
Surgical debulking (suboptimum debulking *vs* optimum debulking)	2·07 (1·53–2·81)	<0·0001	1·50 (1·10–2·06)	0·011
miRNA predictor (high *vs* low risk)	3·25 (2·34–4·51)	<0·0001	3·16 (2·24–4·45)	<0·0001
High-grade serous ovarian cancer (n=185, events=140)				
Stage (III–IV *vs* I–II)	2·81 (1·15–6·90)	0·023	2·67 (0·96–7·38)	0·058
Surgical debulking (suboptimum debulking *vs* optimum debulking)	2·10 (1·50–2·95)	<0·0001	1·62 (1·14–2·29)	0·0071
miRNA predictor (high *vs* low risk)	3·00 (2·09–4·3)	<0·0001	2·96 (2·03–4·31)	<0·0001

Univariate and multivariate analyses were done by Cox regression. HR=hazard ratio.

***Table 5*: Univariate and multivariable analysis of progression-free survival in patients in OC263 with type II ovarian cancer**

MiROvaR retained its independent prognostic effect in multivariable analysis including FIGO stage and residual disease which are generally considered the strongest prognostic clinical variables for prediction of progression-free survival in epithelial ovarian cancer. MiROvaR performed well independent of the criteria adopted for residual disease categorisation and even if applied to heterogeneous populations of patients, thus supporting its value in stratifying patients, irrespective of the clinical–pathological characteristics of their tumours at presentation. With MiROvaR, we aimed to develop a widely useful tool that could encompass the biological and molecular differences among the histological subtypes of ovarian cancer. MiROvaR validation in datasets of patients with various histological characteristics (ie, OC263), strengthens its potential use. Its validity in patients with homogeneous tumour histotypes was confirmed in the OC452 dataset, which relies on high-grade serous ovarian cancers only, and in OC263 when we considered patients with high-grade serous ovarian cancers or type II tumours. Low-grade serous and endometrioid, clear cell, and mucinous ovarian tumours (type I tumours) are generally rare and poorly represented in our study, and would benefit from dedicated studies to test MiROvaR's performance.

The subgroup of patients with a very poor prognosis identified with MiROvaR might be candidate for more aggressive strategies (possibly the addition of bevacizumab to front-line treatment, maintenance treatment, or both). However, we could not predict response to therapeutic treatments in the analysed cohorts since interactions of potential predictive biomarkers and treatment can only be studied in randomised trials. In this study, OC179, which was derived from the MITO-2 randomised clinical trial, represents the only dataset in which such a correlation between treatment and MiROvaR could be done and this yielded negative results. However, the sample size of OC179, is too small to provide sufficient power for any definitive conclusions.

From a molecular point of view, one limit of the functional interpretation of miRNA profiles is their regulatory role, since each miRNA could regulate many genes and the fine-tuning of each gene could be tissue specific. Most of the 35 miRNAs included in MiROvaR have already been identified as having key roles as central nodes in biological processes. Of the 16 miRNAs that gave 100% cross-validation support to the MiROvaR predictor (ie, those contributing most to MiROvaR's ability to stratify patients according to risk of progression), 13 were associated with a favourable prognosis and three with a poor prognosis. In samples classified as high risk by MiROvaR, all 13 of these miRNAs that were individually associated with favourable prognoses were downregulated and the three miRNAs associated with poor prognoses were upregulated. This finding suggests that the maintenance or loss of potentially onco-suppressive miRNAs has a greater effect on the prognosis of epithelial ovarian cancer than does the expression or loss of potentially oncogenic miRNAs, in line with the observation that most miRNAs exert an oncosuppressive role and are consequently mostly downregulated in cancer.[30] Available evidence about these miRNAs' biological roles in epithelial ovarian cancer supports our assumptions, although the roles of the three putative oncogenic miRNAs (miR-193a-5p, miR-30b-3p, and miR-29c-5p) are unclear. Both miR-193a-5p and miR-30b-3p were upregulated in epithelial ovarian cancer refractory to neoadjuvant chemotherapy[31] and in low-grade serous epithelial ovarian cancer compared with healthy fallopian tube tissue.[32] Although miR-29c-5p has been implicated in the regulatory network related to the mesenchymal subtype of high-grade serous ovarian cancer,[33] no information is currently available about its

prognostic role in epithelial ovarian cancer. miR-29c has, however, been described as having an oncosuppressive role in colorectal cancer.[34] In contrast, mature data are available for most of the 13 putative oncosuppressive miRNAs. In particular, we have previously shown that loss of a ChrXq27.3 miRNA cluster, the entirety of which is included in the MiROvaR contributors with 100% cross-validation support (miR-508-3p, miR-509-5p, miR-514a-3p, miR-506-3p, miR-507, miR-509-3p, miR-513b-5p, and miR-513a-5p), is associated with early relapse of epithelial ovarian cancer.[14] This cluster also seems to be downregulated in most of the patients classified as high risk by MiROvaR (figure 1). A deep functional characterisation of miR-506, a key node of the master miRNA regulatory network related to the mesenchymal epithelial ovarian cancer subtype,[33,35] showed that its expression at the tumour level[33] was associated with inhibition of epithelial ovarian cancer proliferation and induction of senescence;[36] suppression of the epithelial-to-mesenchymal transition;[37] and increased response to chemotherapy,[38] suggesting that it has an oncosuppressive role. MiROvaR included most members of the miR-200 family (ie, miR-200a-3p, miR-200b-3p, miR-200c-3p, miR-141-3p, and miR-429), which are regulators of the epithelial-to-mesenchymal transition,[39,40] and loss of miR-200c expression has been associated with relapse in stage I epithelial ovarian cancer.[11] Furthermore, this miRNA family seems to be downregulated in most of the patients classified by MiROvaR as being at high risk of relapse (figure 1). Among the miRNAs contributing with 100% cross-validation support, only miR-592 has no data available about its prognostic role in epithelial ovarian cancer, although its expression was predictive of improved outcome in three different cohorts of colorectal cancers,[41] thus suggesting an oncosuppressive role.

Although definitive data about the biological and prognostic role of all of the miRNAs included in MiROvaR are not yet available in epithelial ovarian cancer, their main effect on the prediction of early recurrence seems to be associated with regulation of the epithelial-to-mesenchymal transition. The large number of miRNAs regulating the epithelial-to-mesenchymal transition that were included in our model (ie, the miR-506 family and miR-200 family) suggests that cellular reprogramming to a more mesenchymal phenotype may be an initiating event during the spread and progression of epithelial ovarian cancer. Furthermore, a study has shown the relevance of the miR-200 family members in a mouse model of breast cancer in which their loss contributed to recurrent lung metastases after chemotherapy, thus suggesting the potential usefulness of an epithelial-to-mesenchymal transition-targeting strategy combined with conventional therapy.[42]

We used three independent datasets for which mature follow-up data were available to develop a strong predictor of risk of progression or relapse in epithelial ovarian cancer. We are aware that our analysis, by relying on the 385 miRNA shared by all the platforms used, might have omitted other relevant miRNAs. However, our work is one of the few attempts to integrate the existing data to build a single model that could be fully validated. This approach attempts to overcome one of the main limitations related to miRNA analyses, which rely on the use of different platforms and annotated lists. Before MiROvaR can be applied as a clinical-grade assay further steps are needed in accordance with the established framework[43] and guidelines,[44] which include, first, identification of an appropriate approach to quantify expression (eg, microarray, RTqPCR, or Nanostring); second, design of specific probes based on the sequences tested in the microarray chips; and third, validation in independent cohorts of patients with full clinical annotation available.

A future opportunity to assess the performance of MiROvaR will be when data become available for translational purposes from samples retrospectively collected in the MITO-7 trial[45] and prospectively collected in the MITO-16 programme (NCT01706120 and NCT01802749); tumour collection has become mandatory in the MITO-16 programme, which might help to avoid attrition bias. Our model is free to the scientific community, and if other key miRNAs are identified in the future, they can be tested and integrated into MiROvaR and potentially be used to help construct a clinically available model for stratification of ovarian cancer risk of relapse or progression.

Contributors

All the authors were involved in the preparation of this manuscript. SC, GB, SP, and DM had the idea for the study. MB, SC, LDC, FP, and DM designed the experiments. DC, SL, GSco, GC, and DR centralised and prepared samples from OC179; FR, DL, MLC, MB, SC, and DM collected samples and clinical data for the OC130 dataset included in the OC263 validation set. GT, EC, RS, VC collected samples and clinical data for the OC133 dataset included into the OC263 validation set. FR, DL, RS, GSca, AS, PS, EB, VM, and SP obtained samples and clinical data. SL, GSco, MLC, VC, GFZ, and MC did the histopathological analysis. LDC did the microarray analysis and submitted the data to the Gene Expression Omnibus. MB, LDC, and MDM did the bioinformatic and statistical analyses. MB, SC, DM, LDC, SP, and FP critically revised all the data. All the authors reviewed the manuscript and approved the final version.

Declaration of interests

We declare no competing interests.

Acknowledgments

The study was partly supported by The Italian Association for Cancer Research (AIRC: IG-10302 to SC, IG-17475 to DM, Special Program 12162 to SC and IG-5776 to SP) and CARIPLO Foundation (2013-0865 to DM). We thank the Biorepository personnel at the Fondazione IRCCS Istituto Nazionale dei Tumori whose activity made this study possible, in particular Silvia Veneroni for her valuable contribution to the tissue sample collection.

References

1 Jayson GC, Kohn EC, Kitchener HC, Ledermann JA. Ovarian cancer. *Lancet* 2014; **384:** 1376–88.

2 Vaughan S, Coward JI, Bast RC Jr, et al. Rethinking ovarian cancer: recommendations for improving outcomes. *Nat Rev Cancer* 2011; **11:** 719–25.

3 Siegel RL, Miller KD, Jemal A. Cancer statistics, 2015. *CA Cancer J Clin* 2015; **65:** 5–29.

4 Tothill RW, Tinker AV, George J, et al. Novel molecular subtypes of serous and endometrioid ovarian cancer linked to clinical outcome. *Clin Cancer Res* 2008; **14:** 5198–208.

5 Kang J, D'Andrea AD, Kozono D. A DNA repair pathway-focused score for prediction of outcomes in ovarian cancer treated with platinum-based chemotherapy. *J Natl Cancer Inst* 2012; **104:** 670–81.

6 Tan TZ, Miow QH, Huang RY, et al. Functional genomics identifies five distinct molecular subtypes with clinical relevance and pathways for growth control in epithelial ovarian cancer. *EMBO Mol Med* 2013; **5:** 983–98.

7 Verhaak RG, Tamayo P, Yang JY, et al. Prognostically relevant gene signatures of high-grade serous ovarian carcinoma. *J Clin Invest* 2013; **123:** 517–25.

8 Riester M, Wei W, Waldron L, et al. Risk prediction for late-stage ovarian cancer by meta-analysis of 1525 patient samples. *J Natl Cancer Inst* 2014; **106:** dju048.

9 Konecny GE, Wang C, Hamidi H, et al. Prognostic and therapeutic relevance of molecular subtypes in high-grade serous ovarian cancer. *J Natl Cancer Inst* 2014; **106:** dju249.

10 Iorio MV, Visone R, Di Leva G, et al. MicroRNA signatures in human ovarian cancer. *Cancer Res* 2007; **67:** 8699–707.

11 Marchini S, Cavalieri D, Fruscio R, et al. Association between miR-200c and the survival of patients with stage I epithelial ovarian cancer: a retrospective study of two independent tumour tissue collections. *Lancet Oncol* 2011; **12:** 273–85.

12 Vecchione A, Belletti B, Lovat F, et al. A microRNA signature defines chemoresistance in ovarian cancer through modulation of angiogenesis. *Proc Natl Acad Sci USA* 2013; **110:** 9845–50.

13 Pignata S, Scambia G, Ferrandina G, et al. Carboplatin plus paclitaxel versus carboplatin plus pegylated liposomal doxorubicin as first-line treatment for patients with ovarian cancer: the MITO-2 randomized phase III trial. *J Clin Oncol* 2011; **29:** 3628–35.

14 Bagnoli M, De Cecco L, Granata A, et al. Identification of a chrXq27.3 microRNA cluster associated with early relapse in advanced stage ovarian cancer patients. *Oncotarget* 2011; **2:** 1265–78.

15 Cancer Genome Atlas Research Network. Integrated genomic analyses of ovarian carcinoma. *Nature* 2011; **474:** 609–15.

16 Callari M, Dugo M, Musella V, et al. Comparison of microarray platforms for measuring differential microRNA expression in paired normal/cancer colon tissues. *PLoS One* 2012; **7:** e45105.

17 Edgar R, Domrachev M, Lash AE. Gene Expression Omnibus: NCBI gene expression and hybridization array data repository. *Nucleic Acids Res* 2002; **30:** 207–10.

18 Gentleman RC, Carey VJ, Bates DM, et al. Bioconductor: open software development for computational biology and bioinformatics. *Genome Biol* 2004; **5:** R80.

19 Heider A, Alt R. virtualArray: a R/bioconductor package to merge raw data from different microarray platforms. *BMC Bioinformatics* 2013; **14:** 75.

20 Johnson WE, Li C, Rabinovic A. Adjusting batch effects in microarray expression data using empirical Bayes methods. *Biostatistics* 2007; **8:** 118–27.

21 Bair E, Tibshirani R. Semi-supervised methods to predict patient survival from gene expression data. *PLoS Biol* 2004; **2:** E108.

22 De Cecco L, Bossi P, Locati L, Canevari S, Licitra L. Comprehensive gene expression meta-analysis of head and neck squamous cell carcinoma microarray data defines a robust survival predictor. *Ann Oncol* 2014; **25:** 1628–35.

23 Taylor JM, Ankerst DP, Andridge RR. Validation of biomarker-based risk prediction models. *Clin Cancer Res* 2008; **14:** 5977–83.

24 Radmacher MD, McShane LM, Simon R. A paradigm for class prediction using gene expression profiles. *J Comput Biol* 2002; **9:** 505–11.

25 Crijns AP, Fehrmann RS, de JS, et al. Survival-related profile, pathways, and transcription factors in ovarian cancer. *PLoS Med* 2009; **6:** e24.

26 Kurman RJ, Carcangiu ML, Herrington CS, Young RH, eds. WHO classification of tumors of female reproductive organs, 4th edn. Lyon: International Agency for Research on Cancer, 2014.

27 Shih I-M, Kurman RJ. Ovarian tumorigenesis: a proposed model based on morphological and molecular genetic analysis. *Am J Pathol* 2004; **164:** 1511–18.

28 Van Peer G, Lefever S, Anckaert J, et al. miRBase Tracker: keeping track of microRNA annotation changes. *Database (Oxford)* 2014; **2014:** bau080.

29 Stuart GC, Kitchener H, Bacon M, et al. 2010 Gynecologic Cancer InterGroup (GCIG) consensus statement on clinical trials in ovarian cancer: report from the Fourth Ovarian Cancer Consensus Conference. *Int J Gynecol Cancer* 2011; **21:** 750–75.

30 Lu J, Getz G, Miska EA, et al. MicroRNA expression profiles classify human cancers. *Nature* 2005; **435:** 834–38.

31 Mariani M, McHugh M, Petrillo M, et al. HGF/c-Met axis drives cancer aggressiveness in the neo-adjuvant setting of ovarian cancer. *Oncotarget* 2014; **5:** 4855–67.

32 Zhang S, Lu Z, Unruh AK, et al. Clinically relevant microRNAs in ovarian cancer. *Mol Cancer Res* 2015; **13:** 393–401.

33 Yang D, Sun Y, Hu L, et al. Integrated analyses identify a master MicroRNA regulatory network for the mesenchymal subtype in serous ovarian Cancer. *Cancer Cell* 2013; **23:** 186–99.

34 Zhang JX, Mai SJ, Huang XX, et al. MiR-29c mediates epithelial-to-mesenchymal transition in human colorectal carcinoma metastasis via PTP4A and GNA13 regulation of beta-catenin signaling. *Ann Oncol* 2014; **25:** 2196–204.

35 Sun Y, Guo F, Bagnoli M, et al. Key nodes of a microRNA network associated with the integrated mesenchymal subtype of high-grade serous ovarian cancer. *Chin J Cancer* 2015; **34:** 28–40.

36 Liu G, Sun Y, Ji P, et al. MiR-506 suppresses proliferation and induces senescence by directly targeting the CDK4/6-FOXM1 axis in ovarian cancer. *J Pathol* 2014; **233:** 308–18.

37 Sun Y, Hu L, Zheng H, et al. MiR-506 inhibits multiple targets in the epithelial-to-mesenchymal transition network and is associated with good prognosis in epithelial ovarian cancer. *J Pathol* 2014; **235:** 25–36.

38 Liu G, Yang D, Rupaimoole R, et al. Augmentation of response to chemotherapy by microRNA-506 through regulation of RAD51 in serous ovarian cancers. *J Natl Cancer Inst* 2015; **107:** djv108.

39 Park SM, Gaur AB, Lengyel E, Peter ME. The miR-200 family determines the epithelial phenotype of cancer cells by targeting the E-cadherin repressors ZEB1 and ZEB2. *Genes Dev* 2008; **22:** 894–907.

40 Mezzanzanica D, Bagnoli M, De Cecco L, Valeri B, Canevari S. Role of microRNAs in ovarian cancer pathogenesis and potential clinical implications. *Int J Biochem Cell Biol* 2010; **42:** 1262–72.

41 Boisen MK, Dehlendorff C, Linnemann D, et al. Tissue microRNAs as predictors of outcome in patients with metastatic colorectal cancer treated with first line capecitabine and oxaliplatin with or without bevacizumab. *PLoS One* 2014; **9:** e109430.

42 Fischer KR, Durrans A, Lee S, et al. Epithelial-to-mesenchymal transition is not required for lung metastasis but contributes to chemoresistance. *Nature* 2015; **527:** 472–76.

43 Micheel CM, Nass SJ, Omenn GS, eds. Commitee on the review of omics-based tests for predicting patient outcomes in clinical trials; Board on health care service; Board on health sciences policy; Institute of medicine; Evolution of translational OMICS: lessons learned and the path foreward. Washington, DC: National Academies Press, 2012.

44 Altman DG, McShane LM, Sauerbrei W, Taube SE. Reporting recommendations for tumor marker prognostic studies (REMARK): explanation and elaboration. *PLoS Med* 2012; **9:** e1001216.

45 Pignata S, Scambia G, Katsaros D, et al. Carboplatin plus paclitaxel once a week versus every 3 weeks in patients with advanced ovarian cancer (MITO-7): a randomised, multicentre, open-label, phase 3 trial. *Lancet Oncol* 2014; **15:** 396–405.