Calculator for ovarian carcinoma subtype prediction

Steve E Kalloger[1,2,5], Martin Köbel[3,5], Samuel Leung[1], Erika Mehl[1], Dongxia Gao[1], Krista M Marcon[1], Christine Chow[1], Blaise A Clarke[4], David G Huntsman[1] and C Blake Gilks[1,2]

[1]*Genetic Pathology Evaluation Centre, Vancouver, BC, Canada;* [2]*Cheryl Brown Ovarian Cancer Outcomes Unit of Vancouver General Hospital and the British Columbia Cancer Agency, Vancouver, BC, Canada;* [3]*Department of Pathology and Laboratory Medicine, University of Calgary, Calgary Laboratory Services, AB, Canada and* [4]*Department of Pathology, University of Toronto, Toronto, ON, Canada*

With the emerging evidence that the five major ovarian carcinoma subtypes (high-grade serous, clear cell, endometrioid, mucinous, and low-grade serous) are distinct disease entities, management of ovarian carcinoma will become subtype specific in the future. In an effort to improve diagnostic accuracy, we set out to determine if an immunohistochemical panel of molecular markers could reproduce consensus subtype assignment. Immunohistochemical expression of 22 biomarkers were examined on tissue microarrays constructed from 322 archival ovarian carcinoma samples from the British Columbia Cancer Agency archives, for the period between 1984 and 2000, and an independent set of 242 cases of ovarian carcinoma from the Gynaecologic Tissue Bank at Vancouver General Hospital from 2001 to 2008. Nominal logistic regression was used to produce a subtype prediction model for each of these sets of cases. These models were then cross-validated against the other cohort, and then both models were further validated in an independent cohort of 81 ovarian carcinoma samples from five different centers. Starting with data for 22 markers, full model fit, backwards, nominal logistic regression identified the same nine markers (CDKN2A, DKK1, HNF1B, MDM2, PGR, TFF3, TP53, VIM, WT1) as being most predictive of ovarian carcinoma subtype in both the archival and tumor bank cohorts. These models were able to predict subtype in the respective cohort in which they were developed with a high degree of sensitivity and specificity (κ statistics of 0.88 ± 0.02 and 0.86 ± 0.04, respectively). When the models were cross-validated (ie using the model developed in one case series to predict subtype in the other series), the prediction equation's performances were reduced (κ statistics of 0.70 ± 0.04 and 0.61 ± 0.04, respectively) due to differences in frequency of expression of some biomarkers in the two case series. Both models were then validated on the independent series of 81 cases, with very good to excellent ability to predict subtype ($\kappa = 0.85 \pm 0.06$ and 0.78 ± 0.07, respectively). A nine-marker immunohistochemical maker panel can be used to objectively support classification into one of the five major subtypes of ovarian carcinoma.
Modern Pathology (2011) **24,** 512–521; doi:10.1038/modpathol.2010.215; published online 3 December 2010

Keywords: biomarker expression; immunohistochemistry; ovarian cancer; ovarian carcinoma; subtype

The majority of women diagnosed with ovarian carcinoma who are treated according to standard protocols will die from their disease,[1] and there has been minimal improvement in outcome over the past two decades.[2] There has been progress in refining the morphological criteria for subtyping of ovarian carcinomas, and diagnostic reproducibility has increased recently, reaching an excellent level of reproducibility.[3,4] The five major subtypes of ovarian carcinoma (high-grade serous, clear cell, endometrioid, mucinous, and low-grade serous) that can be identified based on morphological criteria differ with respect to genetic risk factors, precursor lesions, molecular alterations, stage at presentation, and clinical behavior,[3,5–7] and are best considered to be distinct disease entities.[8] Because of the differences in chemosensitivity between ovarian carcinoma subtypes, there has been a call for

subtype-specific clinical trials, to identify more effective therapies for those subtypes (ie clear cell, mucinous) resistant to conventional platinum/taxane chemotherapy.[9–11]

It has recently been shown that neoadjuvant chemotherapy followed by debulking surgery yields similar outcome results to chemotherapy administered after surgery for patients with advanced stage ovarian carcinoma, with the advantage of decreased morbidity, reduced surgical procedural time, and more rapid post-operative recovery.[12] Neoadjuvant chemotherapy is effective for high-grade serous carcinomas but not for others.[9,10] If neoadjuvant chemotherapy is administered only to specific ovarian carcinoma subtypes, it raises questions about how reliably subtype can be diagnosed on limited samples, such as core biopsy or fine needle aspiration biopsy samples. Although excellent reproducibility is possible when diagnosis is based on examination of multiple slides from a well-sampled tumor,[4] there is no evidence that similar reproducibility is possible based on a small tumor sample. We have recently shown that the major ovarian carcinoma subtypes significantly differ with respect to their biomarker expression profile,[8] suggesting that biomarker expression can be of use in subtype diagnosis. We thus searched for the smallest panel of immunohistochemical biomarkers that showed highest sensitivity and specificity for diagnosis of the five major subtypes of ovarian carcinoma.

Materials and methods

Ethics Statement

Approval for the use of these cohorts for this study was obtained from the Research Ethics Board of the British Columbia Cancer Agency and University of British Columbia.

Study Cohorts and Inclusion Criteria

This study utilized three independent cohorts, two for the generation of prediction equations and one for the validation of those prediction equations (Figure 1).

The first cohort, hereafter referred to as the archival cohort, contained a total of 500 formalin-fixed paraffin-embedded tissues from ovarian carcinoma from a previously described retrospective population-based cohort from British Columbia diagnosed between 1984 and 2000. These samples were from >20 hospitals in British Columbia, with no standardization of specimen fixation or processing. The primary eligibility criterion was the diagnosis of chemonaive ovarian carcinoma, with surgical debulking leading to patients being free of macroscopic residual disease after primary cytoreductive surgery. As a result of this case definition, there were a relatively large number of non-serous
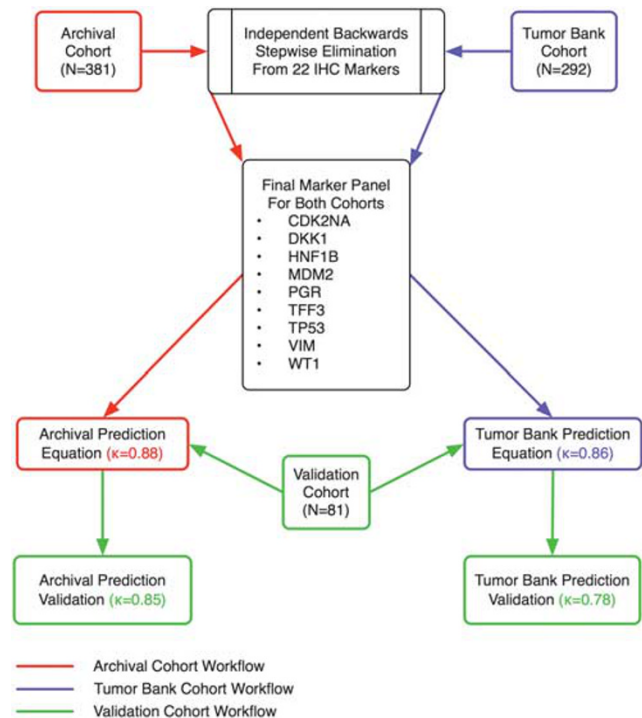

Figure 1 Experimental design for the study detailing independent discovery of the immunohistochemical marker panel from two unrelated cohorts, prediction accuracy represented by the κ statistic (κ), and internal validation.

carcinomas, compared to what would be expected from a series including all patients with ovarian carcinoma.[6]

The second group, hereafter referred to as the tumor bank cohort, consisted of a single hospital-based set of cases from the Gynaecologic Tissue Bank at Vancouver General Hospital and consisted of 292 samples from patients diagnosed with ovarian carcinoma between 2001 and 2008. These samples represent high-quality tissue with short devitalization times and standardized fixation and tissue processing.

A third cohort, hereafter referred to as the validation cohort, was assembled from consecutively diagnosed ovarian carcinoma cases seen at five centers in Canada, and included samples collected from 2006 to 2009 that were part of a recent histomorphological study.[4]

Tumor Classification

All histological slides underwent pathological review (CBG and MK) and were assigned to one of the five major subtypes (high-grade serous, clear cell, endometrioid, mucinous, low-grade serous) or other (including mixed carcinomas and rare types such as undifferentiated, squamous, malignant Brenner tumor, unclassified) according to modified WHO criteria, as recently described.[3,7] Carcinomas with a

transitional growth pattern on low-power examination, but other features of high-grade serous carcinoma, for example, slit-like spaces or microcystic pattern, were classified as a variant of high-grade serous carcinoma ie high-grade serous carcinoma (with transitional features). All cases diagnosed as 'other' (ie not one of the five main subtypes of ovarian carcinoma) were excluded from this study. Furthermore, inter-observer agreement for the histological subtype diagnosis between the two review pathologists was required. Cases with disagreement were excluded ($N = 98$ of 873 cases from the three case series (11.2%)). Finally, only cases where complete data were available for all immunostains were subjected to statistical analysis ($N = 138$ of 775 (17.8%) did not have complete data for all 22 markers and were excluded).

Marker Selection

We searched for biomarkers that are differentially expressed between histological subtypes. The ideal characteristics of these candidate markers for the panel included: availability of a commercial antibody suitable for immunohistochemistry on formalin-fixed, paraffin-embedded tissue, as well as ease of use and interpretation. From our previous studies, we selected 11 markers: Wilms Tumor 1 (WT1), tumor protein p53 (TP53), hepatocyte nuclear factor 1-β (HNF1B), estrogen receptor (ESR), progesterone receptor (PGR), vimentin (VIM), epithelial cell adhesion molecule (EPCAM), mesothelin (MSLN), insulin-like growth factor 2 mRNA binding protein 3, IMP3 (IGF2BP3), cadherin 6, type 2, K-cadherin (CDH6), cyclin-dependent kinase inhibitor 1A (p21, Cip1) (CDKN1A)[8,13,14] and added eight markers reported as differentially expressed in the literature: cyclin-dependent kinase inhibitor 2A (p16) (CDKN2A), mouse double minute 2 (MDM2), catenin (cadherin-associated protein), beta 1, 88 kDa (CTNNB1), phosphatase and tensin homolog (PTEN), mucin 5AC, oligomeric mucus/gel-forming (MUC5AC), paired box 2 (PAX2), secretoglobin, family 1A, member 1 (SCGB1A), CD44 molecule, splice variant 6 (CD44v6).[15–21] Three additional markers: growth differentiation factor 15 (GDF15), trefoil factor 3 (intestinal) (TFF3), dickkopf homolog 1 (DKK1) were derived from comprehensive gene expression profiling data generated using Human Exonic Evidence Based Oligonucleotide microarray (Stanford, CA, USA) or whole transcriptome sequencing, as recently described.[22,23]

Immunohistochemistry

Three tissue microarrays were constructed, as described previously[13] from representative tumor areas using a tissue arrayer (Beecher Instruments, Silver Spring, MD, USA and Pathology Devices, Westminster, MD, USA). Two cores of 0.6 mm

diameter were taken from each donor block and transferred to the recipient block.

Sections of 4 μm thickness were cut and stained within 2 weeks after sectioning. Details of antibodies from the final set of nine markers, and staining procedures, are listed in Table 1. Independently, two pathologists (MK or BG) visually scored these biomarkers from digitalized images (BLISS scanner, Bacus Laboratories/Olympus America, Lombard, IL, USA). All biomarkers were scored as positive or negative, using a cutoff of >1% of cells staining positively (Figure 2), except for CDKN2A for which positive staining was defined as >75% of tumor cells showing strong cytoplasmic and nuclear staining, as defined previously,[24] and TP53, where a three-tier scoring system was used (complete absence—score 0, 1–50%—score 1, and >50%—score 2) because of evidence that scores of 2 and 0 both correlate with underlying mutations, while a score of 1 correlates with wild-type TP53[25] (Figure 3). Less than 5% of cases showed discrepant results between the two scoring pathologists, and for these cases the higher score was used.

Statistical Analysis

Contingency analysis and Pearson's χ^2 statistic (with unadjusted P-values) were used to test for expression heterogeneity for each biomarker between the archival and tumor bank cohorts within each subtype. Full model fit nominal logistic regression modeling, utilizing a manual, iterative backwards elimination process, was used to generate prediction equations starting with the full panel of 22 markers. The criterion for the exclusion of a particular marker was based on the highest P-value in the effect likelihood ratio test.[26] For the model predictions, a receiver operator characteristic area under the curve (AUC) >0.95 was defined as a required result for each histological type for the purpose of this study.

Validation of the prediction equations was accomplished by the application of the final immunohistochemical panel to the validation cohort in order to generate predicted histological subtypes, which would be compared to the consensus histopathological subtype assignment. The subtype considered most likely by the model, even if the probability was <50%, was considered to be the subtype prediction for that case by the model.

In order to address an inherent weakness we identified in the design of this study arising from the possible lack of reliability of prediction equations derived from such a small number of low-grade serous carcinomas in both the archival ($N = 5$) and the tumor bank ($N = 7$) cohorts, we subjected 10 serous borderline tumors from the tumor bank at Vancouver General Hospital to both prediction equations. SEK was responsible for the generation of all statistical analyses and prediction equations.

Table 1 Final antibody list and staining procedures

| Nf marker	Name	Supplier	Cat#	Host	Clone	Antigen retrieval	Primary incubation time	Dilution	Detection system
CDKN2A	Cyclin-dependent kinase inhibitor 2A	CellMarque	CMC802	Mouse	16P04	Mild CC1	32 min with heat	1:20	iView DAB
DKK1	Dickkopf homolog 1	R&D Systems	AF1096	Goat	Poly	Protease 2 8 min	2 h no heat	1:25	UltraMap
HNF1B	Hepatic nuclear factor 1B	Santa Cruz	Sc-7411	Goat	Poly	Standard CC1	2 h no heat	1:200	UltraMap
MDM2	Mouse double minute 2	Neomarkers	MS-291	Mouse	SMP14	Standard CC1	1 h no heat	1:300	UltraMap
PGR	Progesterone receptor	Ventana	790-2223	Rabbit	1E2	Standard CC1	8 min with heat	Prediluted	DABMap
TFF3	Trefoil factor 3	Abnova	H000007033-M01	Mouse	3D9	Standard CC2	1 h with heat	1:50	DABMap
TP53	Tumor protein p53	Dako	M7001	Mouse	D0-7	Mild CC1	32 min with heat	1:400	iView DAB
VIM	Vimentin	Biogenex	MU074-UC	Mouse	V-9	Mild CC1	32 min with heat	1:10 000	iView DAB
WT1	Wilm's tumor 1	Dako	M3561	Mouse	6F-H2	Mild CC1	32 min with heat	1:100	iView DAB
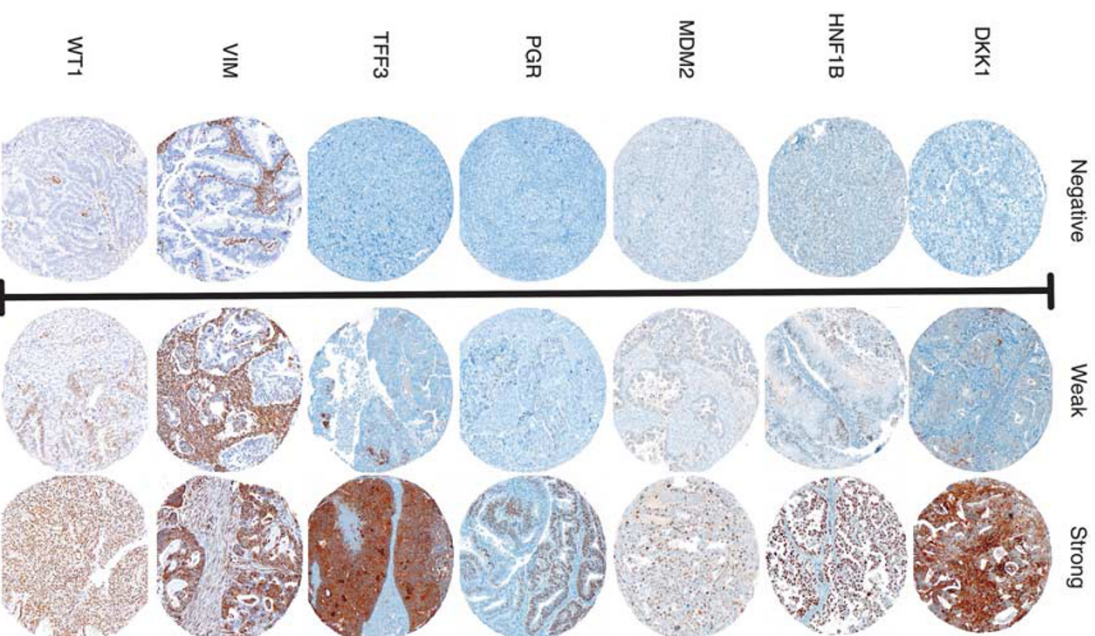
Figure 2 Representative immunostains for DKK1, HNF1B, MDM2, PGR, TFF3, VIM, WT1. Cutoffs between negative (score 0) and positive (score 1) is indicated by the bold vertical line.
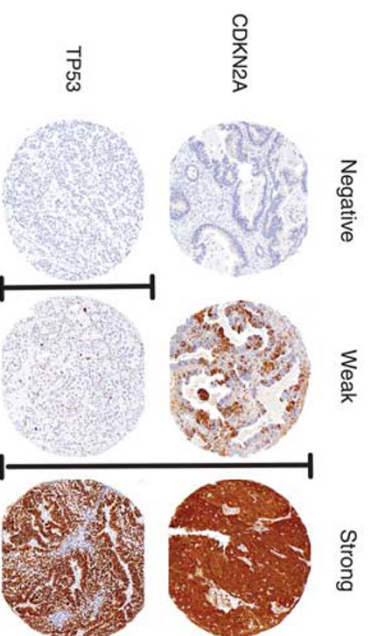
Figure 3 Representative immunostains for CDKN2A and TP53. Cutoffs are indicated with bold lines. For CDKN2A, score 0 is negative or weak positive staining, while strong positive staining is score 1. TP53 was scored using a three part scoring system: 0-negative, 1-weak positive, 2-strong positive. For details of scoring see Materials and Methods.

Table 2 Study populations

Clinicopathological variable	Category	Original archival cohort	Final archival cohort[a]	Original tumor bank cohort	Final tumor bank cohort[a]	Validation cohort
	N	500	314	292	242	81
Mean age (range)		58 (25–89)	58 (25–88)	60 (22–100)	60 (22–100)	61 (35–90)
Subtype	Endometrioid	125 (25%)	85 (27%)	29 (10%)	24 (10%)	6 (7%)
	Mucinous	31 (6%)	17 (5%)	8 (3%)	7 (3%)	6 (7%)
	Clear cell	132 (26%)	79 (25%)	29 (10%)	26 (11%)	12 (15%)
	High-grade serous	200 (40%)	128 (41%)	217 (74%)	178 (73%)	55 (68%)
	Low-grade serous	12 (2%)	5 (2%)	9 (3%)	7 (3%)	2 (3%)
Stage	I	205 (41%)	123 (39%)	36 (13%)	32 (14%)	24 (27%)
	II	211 (42%)	135 (43%)	38 (14%)	32 (14%)	6 (7%)
	III	84 (17%)	56 (18%)	168 (61%)	139 (61%)	56 (64%)
	IV	0 (0%)	0 (0%)	34 (12%)	25 (11%)	4 (2%)

[a]'Final' refers to those cases with complete immunostaining data for all nine markers; derivation of prediction equations was based on these cases.

For all analyses, $P < 0.05$ was considered statistically significant. All statistical analyses were performed using JMP v 8.0.1 (SAS Institute, Cary, NC, USA). Our study is fully compliant with the STARD reporting guidelines that apply to this study.[27]

Results

Basic clinical parameters and morphologically assessed subtype diagnoses for the archival, tumor bank, and validation sets are depicted in Table 2. Due to the requirement of an interpretable staining result for all 22 immunomarkers, the final numbers for the prediction equation generation were 314 for the archival cohort and 242 for the tumor bank cohort. There was no difference between the included and excluded cases, based on tumor subtype frequency.

For both the archival and tumor bank cohorts, the original panel of 22 markers, which were differentially expressed across the subtypes of ovarian carcinoma, was subjected to a full model fit, stepwise nominal logistic regression utilizing backward elimination. The stopping criterion for the backward elimination process was ensuring that the receiver operating characteristic area under the curve (ROC AUC) was ≥ 0.95 for each subtype. This iterative process yielded the same panel of nine markers independently for both the archival and tumor bank cohorts (Figure 1). Both equations were used to predict histological subtype in their respective cohorts utilizing sensitivity, specificity, the receiver operating characteristic area under the curve (ROC AUC), and Cohen's κ coefficient (κ). Overall, the equations did an excellent job predicting the consensus subtype assignment for both the archival ($\kappa = 0.88 \pm 0.02$) and tumor bank ($\kappa = 0.86 \pm 0.04$) cohorts (Tables 3 and 4).

The accuracy of the low-grade serous prediction was tested by the application of both prediction equations to a set of 10 serous borderline tumors,

Table 3 Quality of archival prediction equation for the archival case series

Subtype	Sensitivity (%)	Specificity (%)	ROC AUC	κ
Clear cell	94.9	97.9	0.994	0.88 ± 0.02
Endometrioid	90.6	96.2	0.978	
Mucinous	94.1	98.3	0.993	
High-grade serous	90.6	97.4	0.981	
Low-grade serous	100	98.3	0.999	

Table 4 Quality of tumor bank prediction equation for the tumor bank case series

Subtype	Sensitivity (%)	Specificity (%)	ROC AUC	κ
Clear cell	84.6	99.5	0.995	0.86 ± 0.04
Endometrioid	75.0	99.5	0.991	
Mucinous	100	99.6	0.999	
High-grade serous	98.3	85.3	0.982	
Low-grade serous	85.7	100	0.981	

and showed that the archival prediction equation classified 9 of 10 cases as low-grade serous with one being classified as high-grade serous, while the tumor bank equation classified 1 of 10 cases as low-grade serous carcinoma, 3 of 10 as high-grade serous, and 6 of 10 cases as endometrioid. These results led us to re-derive the tumor bank prediction equation and generate a new prediction equation for this cohort based on only four subtypes (serous, clear cell, endometrioid, mucinous) ie not trying to distinguish between low-grade serous and high-grade serous. This had little impact on the accuracy of the prediction for the remaining four subtypes (Table 5).

The models were then cross-validated against the other cohort (Table 6). In order to investigate the difference between the prediction equations for the archival and tumor bank cohorts, we performed histological subtype-specific tests of heterogeneity for each marker across the test and validation cohort. This showed a trend toward

reduced antigenicity in the archival cohort compared to the clinical cohort and is exemplified best by *HNF1B* staining frequencies in the endometrioid, clear cell, and mucinous subtypes (Table 7). Conversely, *PGR* in the high-grade serous subtype displayed different staining frequencies in the opposite direction with 35.8 and 19.1% of cases in the archival and tumor bank cohorts expressing *PGR*, respectively ($P < 0.001$).

Validation was accomplished by the application of both the archival and tumor bank prediction equations to the validation set (Figure 1). The archival equation yielded a $\kappa = 0.85 \pm 0.06$ and the revised tumor bank equation yielded a $\kappa = 0.78 \pm 0.07$ (Table 8). The archival equation had poor sensitivity for the low-grade serous (50%) and mucinous (66.7%) subtypes, possibly due to the low numbers of these types in the validation cohort ($N = 2$ and $N = 3$, respectively). Overall, the archival prediction equation misclassified five cases; the diagnostic probability according to the model for

Table 5 Quality of revised tumor bank prediction equation for the tumor bank case series

Subtype	Sensitivity (%)	Specificity (%)	ROC AUC	κ
Clear cell	84.6	99.5	0.995	0.86 ± 0.04
Endometrioid	75.0	99.5	0.992	
Mucinous	100	99.6	0.999	
Serous	98.3	85.1	0.986	

This model was not able to distinguish between low-grade and high-grade serous carcinoma; all low-grade serous carcinomas in the cohort were predicted to be serous by the model.

Table 6 Comparison of the archival and tumor bank equations on cross-validation

Subtype	Sensitivity (%)	Specificity (%)	ROC AUC	κ
Archival prediction equation on tumor bank cohort				0.70 ± 0.04
Clear cell	88.5	94.3	0.940	
Endometrioid	79.2	96.0	0.943	
Mucinous	85.7	97.1	0.972	
High-grade serous	86.0	95.5	0.914	
Low-grade serous	71.4	98.3	0.910	
Revised tumor bank prediction equation on archival cohort				0.61 ± 0.04
Clear cell	54.43	98.3	0.899	
Endometrioid	58.82	97.0	0.886	
Mucinous	94.12	98.0	0.933	
Serous	95.5	75.4	0.846	

Table 8 Quality of archival and tumor bank prediction equations for validation case series

Subtype	Sensitivity (%)	Specificity (%)	ROC AUC	κ
Archival prediction equation				0.85 ± 0.06
Clear cell	91.7	97.2	0.982	
Endometrioid	83.3	98.7	0.931	
Mucinous	66.7	98.7	0.967	
High-grade serous	98.2	92.9	0.959	
Low-grade serous	50.0	100	0.826	
Tumor bank prediction equation				0.78 ± 0.07
Clear cell	66.7	97.1	0.851	
Endometrioid	66.7	100	0.872	
Mucinous	100	98.6	0.993	
Serous	96.4	82.8	0.886	

Table 7 Marker expression across subtypes expressed as percent positive

Marker	HGSC			EC			CCC			MC			LGSC		
	A = 128	TB = 178	χ² P	A = 85	TB = 24	χ² P	A = 79	TB = 26	χ² P	A = 17	TB = 7	χ² P	A = 5	TB = 7	χ² P
CDKN2A	64.1	57.9	0.274	15.2	8.3	0.382	6.3	11.5	0.385	0.0	0.0	NR	40.0	0.0	0.067
DKK1	0.8	1.1	0.764	42.4	66.7	**0.035**	0.0	3.8	0.079	17.7	14.3	0.841	0.0	0.0	NR
HNFB1	5.5	7.3	0.522	11.8	33.3	**0.012**	84.8	100	**0.035**	35.3	71.4	0.106	0.0	0.0	NR
MDM2	5.5	1.1	**0.027**	22.4	25.0	0.785	49.4	65.4	0.156	0.0	0.0	NR	60.0	42.9	0.558
PGR	40.6	19.1	**<0.001**	85.5	66.7	0.069	3.8	0.0	0.313	0.0	0.0	NR	60.0	42.3	0.558
TFF3	19.5	16.3	0.464	77.7	70.8	0.489	17.7	15.4	0.784	100	85.7	0.111	60.0	57.1	0.921
TP53 score 0	69.6	77.3	0.321	19.0	0.0	**0.027**	19.2	8.3	0.210	23.1	40.0	0.473	0.0	14.3	0.428
TP53 score 2	80.9	85.8	0.339	8.6	8.3	0.971	1.6	8.3	0.119	28.6	40.0	0.637	20.0	0.0	0.251
VIM	23.4	26.4	0.555	67.1	66.7	0.971	25.3	38.5	0.198	5.9	0.0	0.512	100	85.7	0.377
WT1	82.0	82.0	0.998	3.5	4.2	0.883	1.3	3.6	0.404	0.0	0.0	NR	100	85.7	0.377

A, archival; TB, tumor bank.
The χ² P-values represent tests for heterogeneity for each biomarker across both cohorts, by histological type. For TP53, positivity is based on immunoscore 2 or 0; all other markers were scored as positive or negative, ie staining in >1% of tumor cells was considered as positive. Values in bold are $P < 0.05$.

these cases was of 58.8% (s.d. 9.0) compared to 90.8% (s.d. 14.3) for the correctly classified cases. For 80 cases where the probability of subtype assignment was ≥80%, 68 of 80 (85%) were correctly classified by the tumor bank model. For the archival model, 64 cases had a subtype prediction probability of ≥80% with 60 of 64 (94%) being correctly classified (Figure 4).

We examined the prediction profiles for the excluded cases of undifferentiated and unclassified carcinomas. Both prediction equations classified 9 of 9 cases of undifferentiated carcinoma as high-grade serous. The five cases of unclassified carcinoma were predicted as high-grade serous (3) and mucinous (2) by the archival prediction equation and endometrioid (2), high-grade serous (1), and
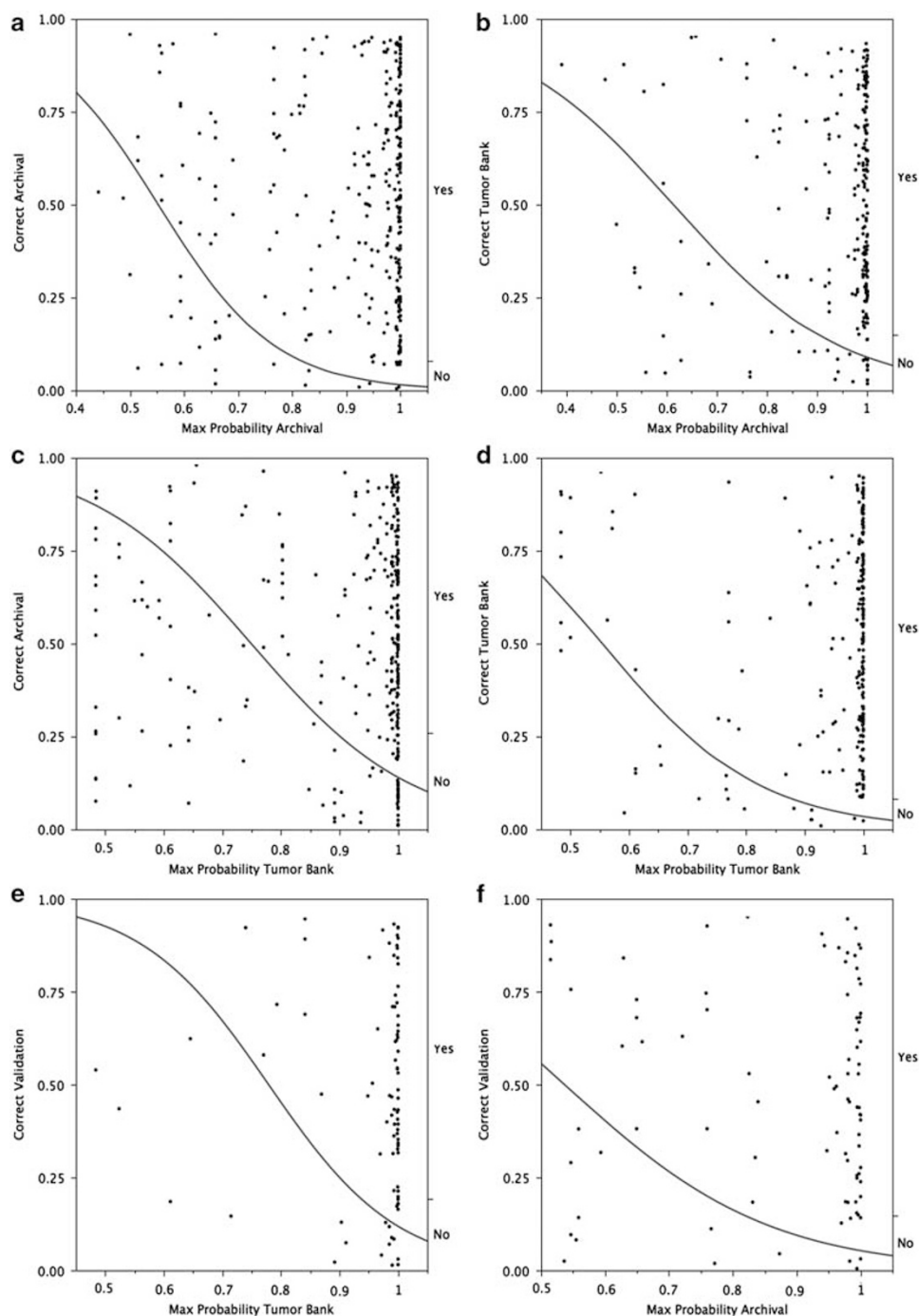

Figure 4 Logistic regression graphs indicating increasing proportion of correct clinical diagnoses with increased predicted probabilities. (**a**) Archival prediction on the archival cohort, (**b**) archival prediction on the tumor bank cohort, (**c**) tumor bank prediction on the archival cohort, (**d**) tumor bank prediction on the tumor bank cohort, (**e**) tumor bank prediction on the validation cohort, (**f**) archival prediction on the validation cohort.

mucinous (2) by the tumor bank prediction equation. All four cases of serous carcinoma with transitional features (as defined in the Materials and methods section) were classified as high-grade serous by both prediction equations.

Discussion

In this study, we developed an immunomarker panel that can accurately predict ovarian carcinoma subtype. This marker panel can be easily applied in practice, providing objective molecular-based support for classification of ovarian carcinomas, enhancing traditional histopathology-based classification.

Ovarian carcinoma types differ with respect to precursor lesions, clinical behavior, molecular alterations, and biomarker expression.[7,8] As for breast carcinomas, where immunomarkers contribute to accurate diagnosis of tumor subtypes and can be used to guide therapeutic decision-making,[28] tumor type-specific management has been also suggested for ovarian carcinoma.[11] For example, in women diagnosed with mucinous or clear cell carcinomas, chemotherapy might be ineffective, whereas, women with advanced high-grade serous carcinomas might benefit from neoadjuvant chemotherapy.[9,10] The integration of molecular findings into the classification of ovarian carcinomas has the potential to enhance reproducibility and accuracy of diagnosis, which will become increasingly important in the future. For example, with neoadjuvant therapy becoming a treatment option for advanced stage ovarian carcinoma[12] accurate typing on core biopsies or cytology specimens prior to treatment will be essential.

Strengths of the study design include the use of two large, independent cohorts, one of which had a larger proportion of minority subtypes (clear cell, endometrioid, and mucinous); the same nine marker panel was derived from 22 candidate makers independently in these two cohorts. As well, a third, independent validation cohort, consisting of cases representative of the frequency of subtypes seen in practice, was used.[4,6] The morphological 'gold standard', which was achieved by the selection of cases in which two expert gynecological pathologists independently agreed on subtype diagnosis, ensured that the diagnoses of the cases in all three case series was uniform and as accurate as possible. Further, in an effort to ensure reproducibility, we restricted our IHC candidate markers to those where commercially available antibodies and simple scoring systems could be used.

Limitations of this study are the fact that the archival and tumor bank cohorts are fundamentally different in terms of tissue handling and fixation. This was the main reason we decided to keep the cohorts separate. As shown in Table 7, there were instances of reduced staining in the older archival cohort compared to the more recent tumor bank cohort. Whether the discordant staining results are caused by the age of the specimens, the tissue handling and fixation, or a combination of the two is impossible to determine with our materials. As one example, HNF1B showed was expressed in 22% more endometrioid carcinomas from the tumor bank cohort compared to the archival cohort and this trend towards increased frequency of expression was seen across all the subtypes in the tumor bank cohort. Logistic regression with sample age revealed that this reduced expression in the archival series becomes more pronounced as a function of specimen age, suggesting antigen degradation as the cause. With attention to tissue handling and advances in antigen retrieval techniques, immunohistochemistry has become a more robust technique that has been shown in quality assurance studies to deliver consistent results in diagnostic laboratories.[29] This will particularly be true for small samples such as core biopsies or cytology specimens, which will be required as part of the triage for neoadjuvant therapy in patients with advanced stage disease, and the tumor bank prediction equation, may prove more accurate for those samples.

Another limitation of this study is the lack of ability of the tumor bank model to distinguish between low-grade serous and high-grade serous carcinoma. Low-grade serous carcinomas account for 3% of ovarian carcinomas,[6] making it difficult to accrue sufficient cases for development of the model. Low-grade serous carcinomas arise from serous borderline tumors in many cases, rather than serous tubal intra-epithelial carcinoma, are not associated with germ line or somatic *BRCA1/2* mutations, or *TP53* mutations, frequently have *BRAF* or *KRAS* mutations, and do not respond as well to platinum-based chemotherapy as high-grade serous carcinomas.[5] Although low-grade and high-grade serous carcinomas can be reproducibly diagnosed when large tissue samples are available for histopathological examination, distinction between low-grade and high-grade serous carcinoma on small samples is problematic, and further work must be done to develop biomarkers to aid in this particular differential diagnosis.

Immunohistochemistry has several advantages over other molecular tests; it is readily available in pathology laboratories, can be implemented as part of the normal diagnostic routine, with rapid turn around time, and is subject to established quality management programs. The interpretation by pathologists in morphological context avoids contradictory results between standard histopathology assessment and independent molecular tests and retains the advantage of assessing the protein expression in the context of exact localization, eg due to the extensive stromal expression VIM would not be a useful marker when assessed at the mRNA level. Immunohistochemical marker panels have been suggested as surrogate markers for expression

CALCULATOR FOR OVARIAN
SUBTYPE PREDICTION (COSP)

[single mode ⬍]

MATERIAL TYPE AND BIOMARKER EXPRESSION

- tissue type: [archival grade material ⬍]
- biomarker expression:

WT1 [1 ⬍]	VIMENTIN [0 ⬍]	TFF3 [0 ⬍]
p16 [0 ⬍]	p53 [1 ⬍]	HNF1B [1 ⬍]
DKK1 [1 ⬍]	PR [1 ⬍]	MDM2 [1 ⬍]

- (reset all)

SUBTYPE PROBABILITY

Subtype	Probability
Clear Cell	0.00%
Endometrioid	72.51%
High-Grade Serous	27.49%
Low-Grade Serous	0.00%
Mucinous	0.00%

Figure 5 Screenshot of the web-based calculator of ovarian carcinoma subtype prediction (COSP). COSP can be used for what we have deemed to be 'archival' or 'tumor bank' grade material. It will be at the users discretion to determine which equation is most applicable to their specimen(s). COSP can be used in single mode or as batched mode in order to allow analysis of large retrospective tissue microarray cohorts with data uploaded in table file format.

profiling in subtyping of breast cancer.[30] The complexity of data generated through use of panels of immunostains requires development of new approaches to interpretation. Nine immunomarkers, with the scoring scheme specified (eight 2-level variables and one 3-level variable), create 768 possible combinations for interpretation. Methods such as recursive partitioning can be used; however, a significant downside to that approach is the potential for overfitting[31] while the nominal logistic regression model avoids overfitting and generates subtype-specific probabilities, which are easy to interpret. To allow pathologists and researchers the immediate use of our findings we launched a calculator of subtype probability (COSP) on our website. http://www.gpec.ubc.ca/index.php? content = papers/ovcasubtype.php (Figure 5). The IHC panel and COSP are designed to be used as a full set of data for all nine markers, as there is no way to reliably impute missing results if there is not a full set of immunostaining results. This approach does not take into account any subjective input from the user and is therefore unbiased. The COSP should currently be considered as a tool for research purposes (eg validation of historical diagnoses for a study group, ensuring comparability of cohorts). Translation for clinical use would require internal validation by institutions planning to apply the COSP to clinical material. Future plans for the tool include the development of an imputation algorithm for missing data and incorporation of new subtype-specific biomarkers. We feel that this resource will complement traditional histopathological classification, and is suitable for use in cases with equivocal morphological features, or where morphological classification is not possible with certainty because of small sample size.

Refined typing, reflecting the underlying tumor biology, and correlating with response to therapy, will contribute to improve the clinical management of women diagnosed with ovarian carcinoma.

Acknowledgements

Disclosure/conflict of interest

MK received fellowship support from Eli Lilly Canada. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

References

1 Jemal A, Siegel R, Ward E, *et al.* Cancer statistics, 2009. CA Cancer J Clin 2009;59:225–249.

2 Chan J, Fuh K, Shin J, *et al.* The treatment and outcomes of early-stage epithelial ovarian cancer: have we made any progress? Br J Cancer 2008;98:1191–1196.

3 Gilks CB, Ionescu DN, Kalloger SE, *et al.* Tumor cell type can be reproducibly diagnosed and is of independent prognostic significance in patients with maximally debulked ovarian carcinoma. Hum Pathol 2008;39:1239–1251.

4 Köbel M, Kalloger SE, Baker PM, *et al.* Diagnosis of ovarian carcinoma cell type is highly reproducible: a Transcanadian study. Am J Surg Pathol 2010;34:984–993.

5 Kurman RJ, Shih IeM. The origin and pathogenesis of epithelial ovarian cancer: a proposed unifying theory. Am J Surg Pathol 2010;34:433–443.

6 Köbel M, Kalloger SE, Huntsman DG, *et al.* Differences in tumor type in low-stage *versus* high-stage ovarian carcinomas. Int J Gynecol Pathol 2010;29:203–211.

7 Gilks CB, Prat J. Ovarian carcinoma pathology and genetics: recent advances. Hum Pathol 2009;40:1213–1223.

8 Köbel M, Kalloger SE, Boyd N, *et al.* Ovarian carcinoma subtypes are different diseases: implications for biomarker studies. PLoS Med 2008;5:e232.

9 Hess V, A'Hern R, Nasiri N, *et al.* Mucinous epithelial ovarian cancer: a separate entity requiring specific treatment. J Clin Oncol 2004;22:1040–1044.

10 Sugiyama T, Kamura T, Kigawa J, *et al.* Clinical characteristics of clear cell carcinoma of the ovary: a distinct histologic type with poor prognosis and resistance to platinum-based chemotherapy. Cancer 2000;88:2584–2589.

11 Gershenson DM. The heterogeneity of epithelial ovarian cancer: getting it right. Cancer 2010;116: 1400–1402.

12 Vergote I, Tropé CG, Amant F, *et al.* Neoadjuvant chemotherapy or primary surgery in stage IIIC or IV ovarian cancer. N Engl J Med 2010;363:943–953.

13 Prentice LM, Klausen C, Kalloger S, *et al.* Kisspeptin and GPR54 immunoreactivity in a cohort of 518 patients defines favourable prognosis and clear cell subtype in ovarian carcinoma. BMC Med 2007;5:33.

14 Köbel M, Kalloger SE, Carrick J, *et al.* A limited panel of immunomarkers can reliably distinguish between clear cell and high-grade serous carcinoma of the ovary. Am J Surg Pathol 2009;33:14–21.

15 O'Neill CJ, McBride HA, Connolly LE, *et al.* High-grade ovarian serous carcinoma exhibits significantly higher p16 expression than low-grade serous carcinoma and serous borderline tumour. Histopathology 2007;50: 773–779.

16 Nenutil R, Smardova J, Pavlova S, *et al.* Discriminating functional and non-functional p53 in human tumours by p53 and MDM2 immunohistochemistry. J Pathol 2005;207:251–259.

17 Geyer JT, López-García MA, Sánchez-Estevez C, *et al.* Pathogenetic pathways in ovarian endometrioid adenocarcinoma: a molecular study of 29 cases. Am J Surg Pathol 2009;33:1157–1163.

18 Ji H, Isacson C, Seidman JD, *et al.* Cytokeratins 7 and 20, Dpc4, and MUC5AC in the distinction of metastatic mucinous carcinomas in the ovary from primary ovarian mucinous tumors: Dpc4 assists in identifying metastatic pancreatic carcinomas. Int J Gynecol Pathol 2002;21:391–400.

19 Tong GX, Chiriboga L, Hamele-Bena D, *et al.* Expression of PAX2 in papillary serous carcinoma of the ovary: immunohistochemical evidence of fallopian tube or secondary Müllerian system origin? Mod Pathol 2007;20:856–863.

20 Tassi RA, Bignotti E, Falchetti M, *et al.* Mammaglobin B expression in human endometrial cancer. Int J Gynecol Cancer 2008;18:1090–1096.

21 Hoshimoto K, Yamauchi N, Takazawa Y, *et al.* CD44 variant 6 in endometrioid carcinoma of the uterus: its expression in the adenocarcinoma component is an independent prognostic marker. Pathol Res Pract 2003;199:71–77.

22 Shah SP, Köbel M, Senz J, *et al.* Mutation of FOXL2 in granulosa-cell tumors of the ovary. N Engl J Med 2009;360:2719–2729.

23 Köbel M, Xu H, Bourne PA, *et al.* IGF2BP3 (IMP3) expression is a marker of unfavorable prognosis in ovarian carcinoma of clear cell subtype. Mod Pathol 2009;22:469–475.

24 Klaes R, Benner A, Friedrich T, *et al.* p16INK4a immunohistochemistry improves interobserver agreement in the diagnosis of cervical intraepithelial neoplasia. Am J Surg Pathol 2002;26:1389–1399.

25 Köbel M, Reuss A, Bois AD, *et al.* The biological and clinical value of p53 expression in pelvic high-grade serous carcinomas. J Pathol 2010;222:191–198.

26 Harrell FE. *Regression modeling strategies: with applications to linear models, logistic regression, and survival analysis.* Springer Verlag: Heidelberg, 2001, pp 189–190.

27 Bossuyt PM, Reitsma JB, Bruns DE, *et al.* Towards complete and accurate reporting of studies of diagnostic accuracy: the STARD initiative. Standards for Reporting of Diagnostic Accuracy. Clin Chem 2003; 49:1–6.

28 Hayes DF, Thor AD, Dressler LG, *et al.* HER2 and response to paclitaxel in node-positive breast cancer. N Engl J Med 2007;357:1496–1506.

29 Terry J, Torlakovic EE, Garratt J, *et al.* Implementation of a Canadian external quality assurance program for breast cancer biomarkers: an initiative of Canadian quality control in immunohistochemistry (cIQc) and Canadian association of pathologists (CAP) national standards committee/immunohistochemistry. Appl Immunohistochem Mol Morphol 2009;17:375–382.

30 Hugh J, Hanson J, Cheang MC, *et al.* Breast cancer subtypes and response to docetaxel in node-positive breast cancer: use of an immunohistochemical definition in the BCIRG 001 trial. J Clin Oncol 2009;27: 1168–1176.

31 Harrell FE, Lee KL, Matchar DB, *et al.* Regression models for prognostic prediction: advantages, problems, and suggested solutions. Cancer Treat Rep 1985;69:1071–1077.