Research Article

Biomarker-Based Ovarian Carcinoma Typing: A Histologic Investigation in the Ovarian Tumor Tissue Analysis Consortium

Martin Köbel[1], Steve E. Kalloger[3], Sandra Lee[1], Máire A. Duggan[1], Linda E. Kelemen[2], Leah Prentice[3], Kimberly R. Kalli[4], Brooke L. Fridley[9], Daniel W. Visscher[5], Gary L. Keeney[5], Robert A. Vierkant[6], Julie M. Cunningham[7], Christine Chow[3], Roberta B. Ness[10], Kirsten Moysich[11], Robert Edwards[12,13,14], Francesmary Modugno[12,13,14], Clareann Bunker[15], Eva L. Wozniak[16], Elizabeth Benjamin[17], Simon A. Gayther[18], Aleksandra Gentry-Maharaj[16], Usha Menon[16], C. Blake Gilks[3], David G. Huntsman[3], Susan J. Ramus[18], and Ellen L. Goode[8], on behalf of the Ovarian Tumor Tissue Analysis consortium

Abstract

Background: Ovarian carcinoma is composed of five major histologic types, which associate with outcome and predict therapeutic response. Our aim was to evaluate histologic type assessments across the centers participating in the Ovarian Tumor Tissue Analysis (OTTA) consortium using an immunohistochemical (IHC) prediction model.

Methods: Tissue microarrays (TMA) and clinical data were available for 524 pathologically confirmed ovarian carcinomas. Centralized IHC was conducted for ARID1A, CDKN2A, DKK1, HNF1B, MDM2, PGR, TP53, TFF3, VIM, and WT1, and three histologic type assessments were compared: the original pathologic type, an IHC-based calculated type (termed TB_COSPv2), and a WT1-assisted TMA core review.

Results: The concordance between TB_COSPv2 type and original type was 73%. Applying WT1-assisted core review, the remaining 27% discordant cases subdivided into unclassifiable (6%), TB_COSPv2 error (6%), and original type error (15%). The largest discordant subgroup was classified as endometrioid carcinoma by original type and as high-grade serous carcinoma (HGSC) by TB_COSPv2. When TB_COSPv2 classification was used, the difference in overall survival of endometrioid carcinoma compared with HGSC became significant [RR 0.60; 95% confidence interval (CI), 0.37–0.93; $P = 0.021$], consistent with previous reports. In addition, 71 cases with unclear original type could be histologically classified by TB_COSPv2.

Conclusions: Research cohorts, particularly those across different centers within consortia, show significant variability in original histologic type diagnosis. Our IHC-based reclassification produced more homogeneous types with respect to outcome than original type.

Impact: Biomarker-based classification of ovarian carcinomas is feasible, improves comparability of results across research studies, and can reclassify cases which lack reliable original pathology. *Cancer Epidemiol Biomarkers Prev; 22(10); 1677–86. ©2013 AACR.*

Authors' Affiliations: [1]Department of Pathology and Laboratory Medicine; [2]Department of Population Health Research, Alberta Health Services-Cancer Care and Department of Medical Genetics, University of Calgary, Calgary, Alberta; [3]Department of Pathology and Laboratory Medicine, University of British Columbia, Vancouver, British Columbia, Canada; [4]Department of Medical Oncology; [5]Division of Anatomic Pathology, Department of Laboratory Medicine and Pathology; [6]Division of Biomedical Statistics and Informatics, Department of Health Science Research; [7]Division of Experimental Pathology and Laboratory Medicine, Department of Laboratory Medicine and Pathology; [8]Division of Epidemiology, Department of Health Science Research, Mayo Clinic, Rochester, Minnesota; [9]Biostatistics and Informatics Shared Resource, University of Kansas Medical Center, Kansas City, Kansas; [10]University of Texas School of Public Health, Houston, Texas; [11]Roswell Park Cancer Institute, Buffalo, New York, New York; [12]Women's Cancer Research Center, Magee-Womens Research Institute, University of Pittsburgh Cancer Institute; [13]Department of Obstetrics, Gynecology, and Reproductive Sciences, University of Pittsburgh School of Medicine, [14]Department of Epidemiology, [15]University of Pittsburgh Graduate School of Public Health, Pittsburgh, Pennsylvania; [16]Gynaecological Cancer Research Centre, Women's Cancer, UCL EGA Institute for Women's Health; [17]Department of Pathology, Cancer Institute, UCL, London, United Kingdom; and [18]Department of Preventive Medicine, Keck School of Medicine, USC/Norris Comprehensive Cancer Center, University of Southern California, Los Angeles, California

Introduction

Studies in recent years have revealed that ovarian carcinoma is not a single disease entity and that histologic type is a reasonable first stratification (1–3). The 5 major histologic/morphologic types of ovarian carcinoma are: high-grade serous carcinoma (HGSC) accounting for 68%, clear cell carcinoma (CCC) for 12%, endometrioid carcinoma for 11%, mucinous carcinoma for 3%, and low-grade serous carcinoma (LGSC) for 3% (4). Histologic types of ovarian carcinoma are characterized by distinct precursor lesions, such as the recently identified serous tubal intraepithelial carcinoma for HGSC and atypical endometriosis for CCC and endometrioid carcinoma, which are associated with distinct molecular alterations (5–7). Because carcinomas of different histologic types originate from different precursor cells, they retain their cell lineage characteristics, which, together with the acquired molecular alterations during oncogenesis, result in specific gene and biomarker expression profiles, as well as a distinct morphologic phenotype (1, 3, 8). Histologic type is a prognostic marker independent of stage (HGSC is associated with the lowest 5-year survival rate) and is a predictive marker for response to standard platinum/paclitaxel chemotherapy, as well as targeted therapies (9–15). Recent studies also provide evidence that several epidemiologic and inherited risk factors are specific to histologic types (16–19).

Although misclassification of histologic type would not affect current clinical management, it is important to achieve a robust histologic classification for ovarian carcinomas according to histologic type for a number of reasons; principally for research into the etiology and prognostic factors in ovarian carcinoma. If pathologists are trained to use refined criteria, interobserver agreement for histologic type with available full slide sets can attain Cohen's κ values of 0.89 (3). However, without specific training, which better reflects current pathology practice, the interobserver agreement is only moderate with Cohen's κ varying between 0.54 and 0.67 (20). A major diagnostic shift has occurred in recent years affecting the classification of carcinomas with glandular architecture and high-grade nuclear features that were formerly classified as high-grade endometrioid carcinoma to now be diagnosed as HGSC (Supplementary Fig. S1; refs. 21, 22). This shift is justified by the fact that those carcinomas are molecularly indistinguishable from the morphologically typical HGSC (23).

We recently suggested an alternative approach to standard morphology-based typing using a nine-marker immunohistochemical (IHC) model termed Calculator for Ovarian Subtype Probability (COSP; ref. 24). COSP incorporated IHC-derived protein expression data, from formalin-fixed paraffin-embedded tissue (FFPE) assembled on tissue microarrays (TMA). Nine markers (CDKN2A, DKK1, HNF1B, MDM2, PGR, TFF3, TP53, VIM, and WT1) were used to predict ovarian carcinoma type in 2 cohorts with differences in tissue handling (24). One COSP model was developed for an archival cohort (A_COSP) assem-

bled from samples collected from 1984 to 2000 when tissue fixation procedures were less standardized than today; this showed very good agreement with expert reviewed morphologic histotype (Cohen's κ = 0.85; ref. 24) and was subsequently validated on archival clinical trial material (25). Another COSP model was developed on the corresponding FFPE tissue of a tumor bank cohort (TB_COSP) consisting of more contemporarily handled samples diagnosed between 2001 and 2008; this also showed substantial agreement with expert reviewed morphologic histotype (Cohen's κ = 0.78; ref. 24). Performance of the latter TB_COSP suffered from a low sensitivity to detect endometrioid carcinoma, CCC, and LGSC (24). Despite this shortcoming, the use of TB_COSP is greater for modern research cohorts because most contemporary cases are processed according to the standardized tissue handling and fixation procedures that are common across pathology departments (26).

The recently formed international Ovarian Tumor Tissue Analysis (OTTA) consortium includes a variety of tissue-based ovarian cancer research studies combined with the goal of understanding the factors related to etiology and outcome of the disease (27). Recently, we have shown that reclassification of histologic type in an effort to improve disease homogeneity can strengthen risk associations for some ovarian cancer histologic types (28). This improved classification may be useful for future research efforts, as current studies often suffer from heterogeneity in diagnostic criteria, adherence to an outdated World Health Organization standard (29), and do not reflect modern classification systems of the 5 major histologic types (3). The main objective of this study was to reclassify TMA cohorts in OTTA using a combined biomarker and morphology approach. Without having access to the review of full pathology slide sets, we were limited to the TMA resource, where we assessed a 10-marker IHC classifier (TB_COSP) and a morphologic review of TMA cores, which was done with the knowledge of WT1 expression status (WT1-assisted core review). The serous cell lineage WT1 marker was selected because it represents the most informative sole biomarker for the most anticipated problem of distinguishing endometrioid carcinoma from HGSC (21). The specific aims of the study were: (i) to refine TB_COSP, (ii) to compare the internal validity of the newly refined TB_COSP version 2 (TB_COSPv2) with the previously reported TB_COSPv1; (iii) to evaluate agreement of the original type with TB_COSPv2; (iv) to arbitrate disagreement of TB_COSPv2 with the original type using WT1-assisted core review, and (v) to compare associations with overall survival across type assessment strategies.

Materials and Methods

Study design

To refine TB_COSP, we used an expanded set of cases with nonmissing clinicopathologic, IHC, and outcome data from the previous FFPE tumor bank cohort (24) as a training set and then used OTTA cases as a testing set.

Local research ethics committees approved all aspects of this study.

Training set analysis

We supplemented the previously published tumor bank cohort (24) with 32 additional cases identified from a similar resource using the consultation files of one author (M.A.D.). These additional cases were composed of: CCC ($N = 7$), endometrioid carcinoma ($N = 6$), HGSC ($N = 13$), LGSC ($N = 4$), and mucinous carcinoma ($N = 2$). These cases were diagnosed during the same time period (2001–2008), and fixation procedures can be expected to adhere to the same standards as the original tumor bank cohort. As in the original, only cases in which 2 gynecological pathologists independently agreed on histologic type were included, and this histologic type assignment was considered the "reference standard" for the training set. The new training set consisted of 253 cases representing the 5 major histologic types (Table 1).

Testing set analysis

The testing set was composed of cases from 3 studies participating in OTTA, including cases enrolled at the Mayo Clinic (MAY; Rochester, MN; $N = 544$; ref. 30), in the UK Ovarian Cancer Population Study (UKO, $N = 119$; ref. 31), and the Hormones and Ovarian Cancer Prediction study ($N = 49$; ref. 32) totaling 712 cases. As with the training set, only cases with nonmissing clinicopathologic and IHC data were included, which resulted in the exclusion of 117 cases which failed IHC for at least one marker (described below; Supplementary Table S1). A further 71 cases with uncertain original histology were excluded from comparisons between IHC-based prediction models and original type. This resulted in a final testing set of 524 cases, for which demographics and outcome data are shown in Table 1 and Supplementary Table S1.

Immunohistochemistry

TMAs containing duplicate to quadruplicate 0.6 mm tissue cores were used; 10 sections (4 microns in thickness)

were stained for ARID1A, CDKN2A, DKK1, HNF1B, MDM2, PGR, TFF3, TP53, VIM, and WT1 (which were the nine markers from the previously published COSP panel; ref. 24) to which we added ARID1A because of its high specificity for clear cell and endometrioid carcinomas (6, 33). Centralized immunohistochemistry was conducted using the Ventana XT platform (Ventana Medical Systems) according to the standard procedures (protocol details are given in Supplementary Table S2), and scoring was conducted by a single pathologist (M. Köbel). The highest score for a given case was used for analysis. Guidelines for categorizing staining as positive or negative are given in Supplementary Table S2, with some refinement made for scoring cutoffs for HNF1B, TFF3, and VIM compared with the previous study (24). All markers were categorized as positive and negative with the exception of TP53, which was kept as 3 tiers: complete absence, wild-type pattern, and overexpression (3). These data were used to create TB_COSPv2 (described below). The testing set was subjected to A_COSP using the publicly available online calculation (under: http://www.gpec.ubc.ca/index.php?content=papers/ovcasubtype.php) as well as TB_COSPv2. A_COSP and TB_COSPv2-assigned probabilities for each of the 5 major histologic types. The histologic type with the highest probability, even if it was less than 50%, was assigned as the predicted type by A_COSP and TB_COSPv2.

WTI-assisted core review

To address what has been previously found to be the predominant misclassification in the typing of ovarian carcinoma, the difference between endometrioid carcinoma and HGSC (21), the testing set was assessed by WT1-assisted core review. Tumor cores on TMA slides that were stained for hematoxylin and eosin (H&E) were reviewed by a single pathologist (M. Köbel) in combination with a corresponding TMA slide that was stained for WT1, and were assigned to one of the 5 major histologic types. In cases with discrepant WT1-assisted core review, resolution was achieved using the WT1-assisted core

Table 1. Characteristics of study populations

		Training set	Testing set	P
N		253	524	N/A
Diagnosis year	Mean (range)	2006 (2001–2008)	2005 (2000–2009)	0.021
Age, year	Mean (range)	60.2 (28–99)	61.5 (21–93)	0.17
Original histologic type	HGSC	176 (69%)	336 (64%)	
	EC	30 (12%)	97 (18%)	
	CCC	31 (12%)	47 (9%)	0.080
	MC	7 (3%)	22 (4%)	
	LGSC	9 (4%)	22 (4%)	
Stage	I/II	71 (28%)	147 (28%)	0.97
	III/IV	182 (72%)	377 (72%)	

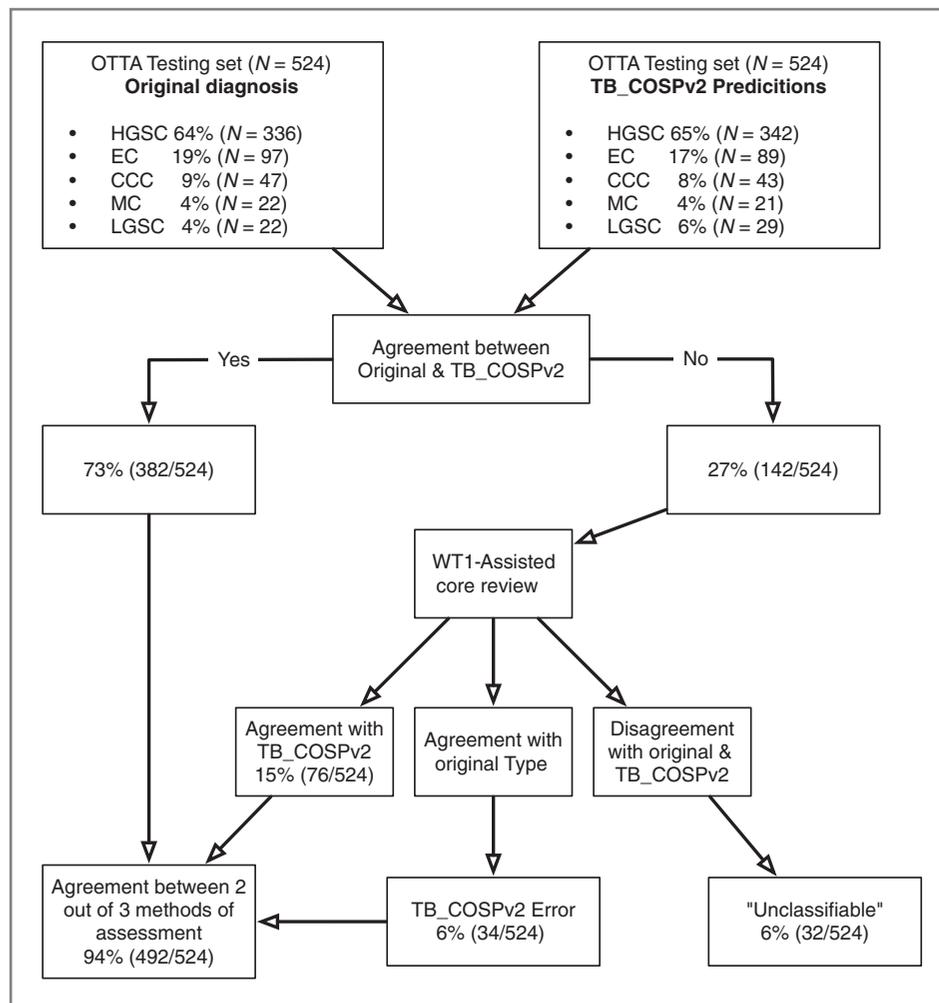Abbreviations: EC, endometrioid carcinoma; MC, mucinous carcinoma.

Figure 1. Histologic type assignments in the testing set. EC, endometrioid carcinoma; MC, mucinous carcinoma.

review for the majority. For example, if 2 of the 3 cores were called HGSC, the WT1-assisted core review was HGSC. A sixth category, "other," was used for cases that could not be histologically classified on the core, for cases where a common assessment could not be reached on a majority of the cores assessed for any single case.

Evaluation of agreement between type assessments

Given the lack of a clear gold standard, we created 2 internal references for histologic type. The first was based on the assumption that, in the cases we studied, agreement between TB_COSPv2 and the original type is likely correct. Second, WT1-assisted core review was used as a "tie-breaker" to arbitrate the cases in the testing set with disagreement between original type and TB_COSPv2. For each case, there were 3 possible outcomes: (i) WT1-assisted core review agreed with original type (TB_COSPv2 error assumed), (ii) WT1-assisted core review agreed with TB_COSPv2 (original type diagnosis error assumed), (iii) WT1-assisted core review did not agree with either (declared as "unclassifiable"). Thus, by combining the 3 methods (original type, TB_COSPv2 and WT1-assisted

core review), each case was either assigned to a certain histologic type by at least 2 out of the 3 methods or was unclassifiable because the 3 methods disagreed (Fig. 1).

Statistical analysis

Nominal logistic regression modeling was used to generate prediction equations, as previously described (24), using the 10-marker panel on the training set. For model predictions, a receiver operator characteristic area under the curve (ROC AUC) for each histologic type category was calculated. The model predictions were then tested for external validity by application to the testing set. To determine agreement between the 3 assessments of histologic type, Cohen's κ statistics were calculated (34). Qualitatively, a k value of 0.2 to 0.4 indicates minimal agreement, 0.4 to 0.6 indicates moderate agreement, 0.6 to 0.8 indicates substantial agreement, and 0.8 to 1.0 indicates excellent agreement (34). Two-way unsupervised hierarchical clustering was conducted using the Wald algorithm for histologic type assignments by different methods. To evaluate histologic type assessments with overall survival, survival curves were generated using the

Table 2. Pairwise agreement of histologic types in the testing set between original type and TB_COSPv2

		TB_COSPv2 Prediction						Concordance rate	K (95% CI)
		HGSC	EC	CCC	MC	LGSC	Total		
Original type	HGSC	**288**	17	7	9	15	336	86%	0.497 (0.432–0.561)
	EC	38	**49**	5	3	2	97	51%	
	CCC	7	7	**29**	1	3	47	62%	
	MC	4	8	1	**8**	1	22	36%	
	LGSC	5	8	1	0	**8**	22	36%	
	Total	342	89	43	21	29	**524**	73%	

NOTE: Ns are shown in each cell, except where % is indicated. Bold indicates cases with agreement.
Abbreviations: EC, endometrioid carcinoma; MC, mucinous carcinoma.

Kaplan–Meier Method and compared with the Wilcoxon test. We used the Cox proportional hazards model to estimate the HRs and 95% confidence intervals (CI) for overall survival, accounting for left truncation. The covariates included in the Cox model were study site (MAY vs. UKO), FIGO stage (stage I and II vs. stage III and IV), and age (older than median vs. median and younger). All statistical analyses were computed with JMP version 10.0 (SAS Institute). This study adhered to the REMARK guidelines for the reporting of biomarker studies (35).

Results

Training set

In an effort to improve upon our previous work in ovarian carcinoma histologic type classification, we examined an expanded version of a previously used training set (24) and improved the internal validity of the immunohistochemistry-based type prediction by adding ARID1A and refining the scoring categories of HNF1B, TFF3, and VIM, the details of which are shown in Supplementary Table S2. The new prediction algorithm (TB_COSPv2) was compared with both of the previous algorithms (TB_COSPv1 and A_COSP). Supplementary Table S3 shows the overall superiority of the TB_COSPv2 algorithm compared with the previous algorithms produced from our earlier efforts as quantified by ROC AUC. This internal validation showed a 98% concordance between TB_COSPv2 and original type, yielding only 5 misclassifications within the training set.

Testing set

External validity was assessed by the application of TB_COSPv2 to the OTTA testing set. The testing did not significantly differ from the training set with respect to patient age, histologic type, and stage distribution. The mean diagnosis year of the tumor blocks of the training set was 2006, which was significantly older by one year compared with the testing set ($P = 0.021$). Still, all training set cases were diagnosed within the time range of the testing set cases suggesting no differences in tissue handling due to overlapping time periods. Supplementary Table S4 shows the categorized expression results for each

marker by histologic type. Statistically significant differential marker expression across the training and testing sets was seen for two markers in HGSC: DKK1 and MDM2; and for MDM2 alone in EC. Table 2 shows cross-tabulations and agreement between original type and TB_COSPv2. The overall agreement with original diagnosis was 73% but varied across the types. This agreement was highest for HGSC with 86% and lowest for mucinous carcinoma and LGSC both with 36%. Most reclassifications by TB_COSPv2 were made in endometrioid carcinomas, which were reclassified primarily to HGSC (HGSC, $N = 38$; non-HGSC, $N = 10$) or were reclassified from other histologic types to endometrioid carcinoma ($N = 40$). We also applied the previous A_COSP to the OTTA training set, and cross-tabulations of A_COSP with original type showed a similar agreement rate with original diagnosis (75%; Supplementary Table S5). To objectively assess which prediction equation was superior (A_COSP vs. TB_COSPv2) in the testing set, we chose to use the prognostic difference between HGSC and endometrioid carcinoma as a measure of proper classification based on the fact that the latter should have a superior prognosis to the former. In the examination of the prognostic significance, we had to exclude cases derived from the Hormones and Ovarian Cancer Prediction study due to a lack of outcome data. This exclusion resulted in the removal of 10 cases with original type of CCC ($N = 1$) and endometrioid carcinoma ($N = 9$). Supplementary Table S6 shows the HRs for histologic types in univariate analysis. The HR for endometrioid carcinoma compared with the HGSC reference was smaller for TB_COSPv2 (HR = 0.35; 95% CI, 0.22–0.52) compared with A_COSP (HR = 0.43; 95% CI, 0.28–0.64). Because of the superior survival difference, the following analyses were restricted to TB_COSPv2.

TB_COSPv2 had a 73% agreement with the original type. For the 27% ($N = 142$) discordant cases, the WT1-assisted core review was used as an arbiter to resolve the discrepancy (Fig. 1). By doing so, 54% of cases discordant between the original type and TB_COSPv2 were determined to have an error in original type ($N = 76$ or 15% of all cases). Twenty-four percent of discordant cases were

Table 3. Pairwise agreement of histologic types in the testing set between original type and agreement between 2 out of 3 methods of agreement (original type, TB_COSPv2, and WT1-assisted core review)

		Agreement between 2 out of 3 methods of assessment						Concordance rate	K (95% CI)
		HGSC	EC	CCC	MC	LGSC	Total		
Original type	HGSC	**300**	6	5	1	9	321	94%	0.697 (0.636–0.758)
	EC	32	**54**	3	2	0	91	59%	
	CCC	5	3	**37**	1	0	46	80%	
	MC	2	4	0	**13**	0	19	68%	
	LGSC	3	0	0	0	**12**	15	80%	
	Total	342	67	45	17	21	**492**	85%	

NOTE: Ns are shown in each cell, except where % is indicated. Bold indicates cases with agreement.
Abbreviations: EC, endometrioid carcinoma; MC, mucinous carcinoma.

found to have an error for TB_COSPv2 prediction ($N = 34$ or 6% of all cases), and 22% of discordant cases were declared as "unclassifiable" ($N = 32$ or 6% of all cases; Fig. 1). The probabilities derived by TB_COSPv2 were not significantly different for cases with original type error, TB_COSPv2 error, or unclassifiable ($P = 0.51$). The most common TB_COSPv2 error was the prediction of endometrioid carcinoma ($N = 16/34$), which were equally distributed across the types and predicted as HGSC ($N = 4$), LGSC ($N = 4$), mucinous carcinoma ($N = 4$), and CCC ($N = 4$) by the other 2 methods.

Table 3 shows the patterns of agreement between the original type and after the WT1-assisted core review was applied to resolve discrepancies between original type and TB_COSPv2. After excluding the 32 unclassifiable cases, the agreement rate with original type increased to 85% ($N = 416/492$) compared with 73% ($N = 382/524$) for

TB_COSPv2 alone because 34 cases were added where the WT1-assisted core review agreed with original type. The agreement rate was highest for HGSC with 94% and intermediate for CCC and LGSC with 80%. Endometrioid carcinoma, however, showed the lowest level of agreement with 59%, which was due to a large group ($N = 32$) with systematic disagreement between original endometrioid carcinoma type and HGSC by both the other methods (TB_COSPv2 and WT1-assisted core review).

To provide justification for the reclassification, outcome analysis were conducted with 4 type assessments: original type (100% of cases), TB_COPSv2 (100%), agreement between original type and TB_COSPv2 (73%), and agreement between 2 out of 3 methods of assessment (94% of cases; Fig. 1). Kaplan–Meier curves are displayed in Fig. 2 and the 5-year overall survival rates (5y-OS) for each histologic type are shown in Supplementary Table S7.
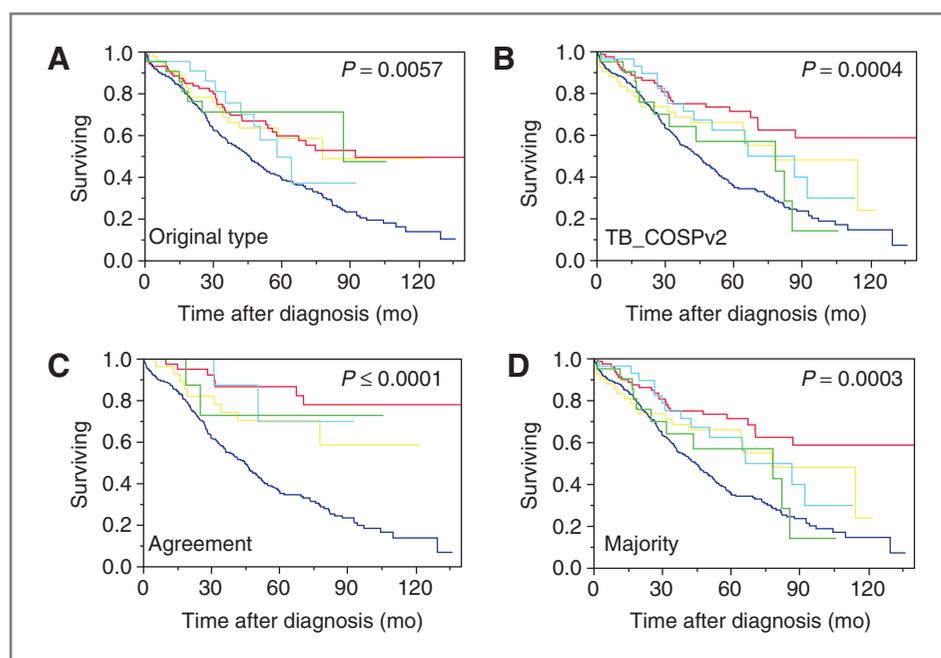
Figure 2. Overall survival of the testing set shown in univariate Kaplan–Meier analysis. Red is endometrioid carcinoma, dark blue is HGSC, yellow is CCC, green is mucinous carcinoma, light blue is LGSC. P values represent Wilcoxon test of survival difference across type assessments: A, original type; B, TB_COSPv2; C, agreement between original type and TB_COSPv2; and D, agreement between 2 out of 3 methods of assessment (original type, TB_COSPv2, and WT1-assisted core review).

Table 4. Multivariate Cox model in the testing set using different methods of type assessment

		Original type HR (95% CI), P	TB_COSPv2 HR (95% CI), P	Agreement between original type and TB_COSPv2 HR (95% CI), P	Agreement between 2 out of 3 methods of original type, TB_COSPv2, and WT1-assisted core review HR (95% CI), P
FIGO stage	Stage III/IV versus I/II	4.39 (2.81–7.12), **P < 0.0001**	3.83 (2.50–5.99), **P < 0.0001**	4.47 (2.47–8.67), **P < 0.0001**	4.82 (3.00–8.08), **P < 0.0001**
Site	MAY versus UKO	1.02 (0.70–1.53), P = 0.91	1.06 (0.75–1.60), P = 0.76	1.24 (0.76–2.14), P = 0.40	1.07 (0.73–1.64), P = 0.71
Age	Older than median versus median or younger	1.45 (1.14–1.87), **P = 0.0026**	1.42 (1.11–1.82), **P = 0.0053**	1.42 (1.06–1.81), **P = 0.016**	1.49 (1.16–1.93), **P = 0.0019**
Histologic type	HGSC (Reference) P value overall	1.00, P = 0.67	1.00, P = 0.054	1.00, **P = 0.026**	1.00, **P = 0.032**
	EC	0.85 (0.57–1.23), P = 0.41	0.60 (0.37–0.93), **P = 0.021**	0.36 (0.14–0.80), **P = 0.0098**	0.56 (0.29–0.99), **P = 0.046**
	CCC	1.25 (0.72–2.07), P = 0.41	1.06 (0.62–1.70), P = 0.82	1.00 (0.46–1.92), P = 1.00	1.31 (0.78–2.06), P = 0.29
	MC	1.46 (0.42–3.83), P = 0.50	0.95 (0.45–1.75), P = 0.87	1.30 (0.20–5.60), P = 0.73	1.51 (0.52–3.47), P = 0.40
	LGSC	0.85 (0.40–1.58), P = 0.64	0.57 (0.31–0.97), **P = 0.037**	0.31 (0.05–0.98), **P = 0.045**	0.50 (0.21–0.99), **P = 0.046**

NOTE: Significant P value highlighted in bold.
Abbreviations: EC, endometrioid carcinoma; MC, mucinous carcinoma.

The 5y-OS of HGSC, the largest group, varied only slightly across type assessments, probably because only a small fraction of HGSCs were reclassified by TB_COSPv2. The 5y-OS rates for CCC were also relatively stable across the methods. In contrast, endometrioid carcinoma, the second largest group, showed an improved 5y-OS rate of 87% when both the original type and TB_COSPv2 prediction agreed, compared with only 59% when the original type was used. Reclassified endometrioid carcinoma by the TB_COSPv2 prediction was associated with a 71% 5y-OS rate, and this improved to 80% when agreement between 2 out of 3 methods of assessment predicted the endometrioid carcinoma histology. Results of the Cox proportional hazards models adjusted for stage, age, and study site are depicted in Table 4. The survival difference between endometrioid carcinoma and HGSC was statistically insignificant with the original type (HR 0.85; 95% CI, 0.57–1.23; P = 0.41) but attained statistical significance when the other methods of type assessment were applied, for example, TB_COSPv2 (HR 0.60; 95% CI, 0.37–0.93; P = 0.021), suggesting improved group homogeneity.

When we compared clinicopathologic parameters among cases classified as endometrioid carcinoma according to the different methods of type assessment, those with agreement between the original type and TB_COSPv2 prediction, compared with the original type alone, had lower proportions of high-stage disease (20% vs. 46%; P = 0.0022), grade 3 (25% vs. 50%; P = 0.050), and WT1 expression (6% vs. 40%; P < 0.0001; Supplementary Table S8), which is consistent with expected endometrioid carcinoma characteristics (21).

Finally, 71 cases with uncertain original diagnoses were included in TB_COSPv2 predictions (Supplementary Table S9). Of the 35 cases known to be serous but which were ungraded (hence could not be assigned to HGSC or LGSC), 28 were predicted to be HGSC and 7 to be LGSC. Of the remaining 36, the most common predicted types were HGSC (N = 23, 64%) and endometrioid carcinoma (N = 8, 22%). Almost all "other" or undifferentiated carcinomas classified as HGSC. Mixed carcinoma split into HGSC, endometrioid carcinoma, and CCC. The predicted types showed the expected 5y-OS rates of 39% for HGSC and 71% for endometrioid carcinoma, providing an example of the viability of using immunohistochemistry when the original pathology is unclear.

Discussion

This study shows the feasibility of using IHC classifiers such as the TB_COSPv2 prediction model for improved classification of histologic type in research using TMA cohorts. TMA technology has become popular for biomarker interrogation and can overcome the remarkable heterogeneity in histologic type assignment across cohorts that use combinations of original pathology report, local or central pathology review, or full slide or selected slide review. Currently, very few studies rely on tumor biomarkers for histologic classification purposes. However, in light of the recent acceptance that ovarian

carcinoma types are essentially distinct diseases, and that misclassification of histologic type can confound results (1), it is important to achieve a standardized and reproducible system/method of subclassification. Morphologic review of slides suffers from high retrieval costs for slides, constraints of pathologists' time, and interobserver variation between pathologists. Therefore, an approach that is able to directly use TMAs can be advantageous. Our results show that the application of IHC biomarker assessment can be highly accurate and able to uncover misclassified cases. The TB_COSPv2 model showed a 98% concordance with the morphologic "gold standard" type in the training set and a 73% concordance with the original type in the OTTA testing set. In the testing set, WT1-assisted core review revealed that most cases of disagreement (about 15% of all cases) were likely due to original type misclassification, whereas TB_COSPv2 error occurred only in 6% of cases. Another 6% of cases remained unclassifiable with this approach. We acknowledge that the amount of tissue available for WT1-assisted core review, being limited, may not be representative. But the amount of tissue in TMA cores is comparable with diagnostic material available from a cell block obtained by paracentesis or a core biopsy taken from an omental cake in current clinical practice before commencing neoadjuvant chemotherapy (36). Nevertheless, WT1-assisted core review increased the agreement rate with the original type from 73% to 85% of cases, and can therefore be used as a safeguard to avoid TB_COSPv2 errors in discordant cases.

By considering the final reclassification assessment in Table 3, the pattern of disagreement is mostly random. The minor pattern of systematic disagreement could be due to the tendency for pathologists to overcall endometrioid carcinoma in the original type that was consequently reclassified as HGSC by TB_COSPv2 and WT1-assisted core review. Reclassification of endometrioid carcinoma to HGSC is an expected result, which has been shown previously in a morphologic review of a large population-based series in Canada (21). The differential diagnosis of HGSC and endometrioid carcinoma, particularly in high-grade cases, is still a matter of controversy among pathologists (22). But many now advocate that the vast majority of cases that show high-grade nuclear atypia regardless of architectural features should be diagnosed as HGSC (2, 37–40). Hence, the criteria for diagnosing HGSC versus endometrioid carcinoma have evolved over time. This study uses biomarkers for reclassification; therefore, it is difficult to improve on tumor classification systems and to avoid circular reasoning. In the current study, 40% of original endometrioid carcinoma showed WT1 expression, which decreased to 6% ($P < 0.0001$) when any 2 of the 3 assessment methods were in agreement to predict endometrioid carcinoma histology. Although this would be expected, because WT1 expression was a component of both the TB_COSPv2 algorithm and the WT1-assisted core review, a similar shift was observed for reclassified endometrioid carcinoma based on morphology only (21, 41).

Furthermore, reclassified endometrioid carcinoma was also less likely to be categorized as grade 3 and FIGO stage III or higher, suggesting improved homogeneity and consistency with expected patterns.

We considered survival as an objective outcome parameter to provide validity for the specific reclassification of endometrioid carcinoma to HGSC. The 5y-OS rate for the original type of endometrioid carcinoma was only 59%, and this survival became statistically significantly different compared with HGSC following the reclassification of HGSCs to endometrioid carcinomas by TB_COSPv2. Survival-based outcome measures are highly stage dependent and almost half of the original diagnoses of endometrioid carcinoma were high stage compared with approximately 27% when any 2 of the 3 assessment methods agreed. One could argue that our findings for endometrioid carcinoma are due to the exclusion of high-stage disease; however, associations remained statistically significant even when the adjustment was made for stage in multivariate models. Overall, we believe that these outcome changes justified our approach.

Reclassification of histologic types is much more random with respect to the other types. TB_COSPv2 shifted 34% of original CCC into other categories (mostly HGSC) and reclassified mostly HGSC and endometrioid carcinoma to compose 35% of reclassified CCC. The majority of these reclassifications were confirmed by WT1-assisted core review. This shift resulted in no relevant change in the 5y-OS rate (56%–65%), probably due to the intermediate outcome of CCC and the random nature of reclassification events. Therefore, outcome may not be used as a surrogate for histologic type for all instances or in individual cases. The WT1-assisted core review was particularly helpful to increase the agreement rate of mucinous carcinoma and LGSC from 36% each by TB_COSPv2 alone to 68% or 80%, respectively. Further research in larger TMA collections of these minority types is needed.

This exercise not only shows the feasibility of retrospective reclassification of histologic type for accuracy but offers economic efficiencies as well. With an average cost of $40 per IHC slide, the 10-marker assay would cost approximately CAD 400 per TMA block plus scoring and analytic time at a typical medical research institution. For this study, cases were assembled on 8 TMA resulting in a total assay cost of CAD 3,200 (or CAD 6 per case). There are also limitations to this method, including case loss due to the requirement for a complete biomarker dataset for TB_COSPv2. We lost 117 (16%) of the 712 testing set cases for classification. An imputation of missing data would be very desirable to avoid case loss, as would use of triplicate cores and optimization of TMA design to minimize core drop out. On the other hand, 71 (10%) of the test set cases, which did not have a diagnosis of 1 of the 5 major histological types, could be reclassified by TB_COSPv2. Notably, TB_COSPv2 could differentiate between high-grade and low-grade serous, reclassify almost all "other" or undifferentiated carcinomas classified as HGSC, and assign mixed carcinoma as HGSC,

endometrioid carcinoma, or CCC. Another limitation of TMAs is with regard to mixed carcinomas. Because it is often not feasible to take multiple cores from different tumor components of a mixed carcinoma, full sections will be necessary to accurately classify mixed carcinomas. Reproducibility of IHC stains has been stressed in recent years. Although our data show that the majority of IHC markers used in this study showed a constant expression rate within types across sets (some markers even within 1% range), we identified 2 problem markers where more reliable antibodies are needed (MDM2 and DKK1).

This study sheds light on issues of heterogeneity in pathologic subclassification of ovarian carcinomas. Our reclassification exercise shows that outcome of endometrioid carcinoma is largely driven by accurate diagnosis. The issues of heterogeneous type assignment could have a major effect on large-scale tissue-based biomarker studies, particularly within the uncommon types. This study presents a research tool that can be used for immunohistochemistry-based ovarian carcinoma typing. Further advances that will increase robustness of the classification include digital pathology enabling assessment of H&E morphology with simultaneous assessment of biomarker expression on a single screen, multiplex immunofluorescence methods, and use of additional molecular markers such as somatic mutation status. Application of these approaches could provide a more objective frame of reference toward a biologic stratification of ovarian carcinomas.

Disclosure of Potential Conflicts of Interest

M.A. Duggan has commercial research support from Hologic and is a consultant/advisory board member of BD. U. Menon has expert testimony from Abcodia Ltd. No potential conflicts of interest were disclosed by the other authors.

Authors' Contributions

Conception and design: M. Köbel, S. Kalloger, S.J. Ramus, E.L. Goode
Development of methodology: M. Köbel, S. Kalloger, B. Gilks, S.J. Ramus
Acquisition of data (provided animals, acquired and managed patients, provided facilities, etc.): M. Köbel, M.A. Duggan, K.R. Kalli, D.W. Visscher, G.A. Keeney, K.B. Moysich, R.P. Edwards, F. Modugno, C.H. Bunker, E.L. Wozniak, E.Benjamin, S.A. Gayther, A. Gentry-Maharaj, B. Gilks, S.J. Ramus, E.L. Goode
Analysis and interpretation of data (e.g., statistical analysis, biostatistics, computational analysis): M. Köbel, S. Kalloger, K.R. Kalli, B.L. Fridley, D.W. Visscher, R.A. Vierkant, K.B. Moysich, E.Benjamin, A. Gentry-Maharaj, S.J. Ramus, E.L. Goode
Writing, review, and/or revision of the manuscript: M. Köbel, S. Kalloger, S. Lee, M.A. Duggan, L.E. Kelemen, L. Prentice, K.R. Kalli, D.W. Visscher, R.A. Vierkant, J.M. Cunningham, R.B. Ness, K.B. Moysich, R.P. Edwards, F. Modugno, E.Benjamin, S.A. Gayther, A. Gentry-Maharaj, U. Menon, B. Gilks, S.J. Ramus, E.L. Goode
Administrative, technical, or material support (i.e., reporting or organizing data, constructing databases): M. Köbel, S. Kalloger, L. Prentice, D. W. Visscher, C. Chow, E.L. Wozniak, S.J. Ramus
Study supervision: M. Köbel, S. Kalloger, K.B. Moysich, C.H. Bunker, D.G. Huntsman, S.J. Ramus

References

1. Köbel M, Kalloger SE, Boyd N, McKinney S, Mehl E, Palmer C, et al. Ovarian carcinoma subtypes are different diseases: implications for biomarker studies. PLoS Med 2008;5:e232.
2. Kurman RJ, Shih I-M. The origin and pathogenesis of epithelial ovarian cancer: a proposed unifying theory. Am J Surg Pathol 2010;34:433–43.
3. Köbel M, Kalloger SE, Baker PM, Ewanowich CA, Arseneau J, Zherebitskiy V, et al. Diagnosis of ovarian carcinoma cell type is highly reproducible: a transcanadian study. Am J Surg Pathol 2010;34: 984–93.
4. Köbel M, Kalloger SE, Huntsman DG, Santos JL, Swenerton KD, Seidman JD, et al. Differences in tumor type in low-stage versus high-stage ovarian carcinomas. Int J Gynecol Pathol 2010;29: 203–11.
5. Crum CP, McKeon FD, Xian W. BRCA, the oviduct, and the space and time continuum of pelvic serous carcinogenesis. Int J Gynecol Cancer 2012;22 Suppl 1:S29–34.
6. Wiegand KC, Shah SP, Al-Agha OM, Zhao Y, Tse K, Zeng T, et al. ARID1A mutations in endometriosis-associated ovarian carcinomas. N Engl J Med 2010;363:1532–43.
7. Kurman RJ, Shih I-M. Molecular pathogenesis and extraovarian origin of epithelial ovarian cancer-Shifting the paradigm. Hum Pathol 2011;42:918–31.
8. Zorn KK, Bonome T, Gangi L, Chandramouli GVR, Awtrey CS, Gardner GJ, et al. Gene expression profiles of serous, endometrioid, and clear cell subtypes of ovarian and endometrial cancer. Clin Cancer Res 2005;11:6422–30.
9. Köbel M, Kalloger SE, Santos JL, Huntsman DG, Gilks CB, Swenerton KD. Tumor type and substage predict survival in stage I and II ovarian carcinoma: insights and implications. Gynecol Oncol 2010; 116:50–6.
10. Mackay HJ, Brady MF, Oza AM, Reuss A, Pujade-Lauraine E, Swart AM, et al. Prognostic relevance of uncommon ovarian histology in women with stage III/IV epithelial ovarian cancer. Int J Gynecol Cancer 2010;20:945–52.
11. Gelmon KA, Tischkowitz M, Mackay H, Swenerton K, Robidoux A, Tonkin K, et al. Olaparib in patients with recurrent high-grade serous or poorly differentiated ovarian carcinoma or triple-negative breast cancer: a phase 2, multicentre, open-label, non-randomised study. Lancet Oncol 2011;12:852–61.
12. McAlpine JN, Wiegand KC, Vang R, Ronnett BM, Adamiak A, Kobel M, et al. HER2 overexpression and amplification is present in a subset of ovarian mucinous carcinomas and can be targeted with trastuzumab therapy. BMC Cancer 2009;9:433.
13. Anglesio MS, Carey MS, Köbel M, Mackay H, Huntsman DGVancouver Ovarian Clear Cell Symposium Speakers. Clear cell carcinoma of the ovary: a report from the first Ovarian Clear Cell Symposium, June 24th, 2010. Gynecol Oncol 2011;121:407–15.
14. Kelemen LE, Köbel M. Mucinous carcinomas of the ovary and colorectum: different organ, same dilemma. Lancet Oncol 2011;12:1071–80.

15. Diaz-Padilla I, Malpica AL, Minig L, Chiva LM, Gershenson DM, Gonzalez-Martin A. Ovarian low-grade serous carcinoma: a comprehensive update. Gynecol Oncol 2012;126:279–85.

16. Pearce CL, Templeman C, Rossing MA, Lee A, Near AM, Webb PM, et al. Association between endometriosis and risk of histological subtypes of ovarian cancer: a pooled analysis of case-control studies. Lancet Oncol 2012;13:385–94.

17. Song H, Ramus SJ, Tyrer J, Bolton KL, Gentry-Maharaj A, Wozniak E, et al. A genome-wide association study identifies a new ovarian cancer susceptibility locus on 9p22.2. Nat Genet 2009;41:996–1000.

18. Goode EL, Chenevix-Trench G, Song H, Ramus SJ, Notaridou M, Lawrenson K, et al. A genome-wide association study identifies susceptibility loci for ovarian cancer at 2q31 and 8q24. Nat Genet 2010;42:874–9.

19. Chen S, Parmigiani G. Meta-analysis of BRCA1 and BRCA2 penetrance. J Clin Oncol 2007;25:1329–33.

20. Patel C, Harmon B, Soslow RA, Garg K, DeLair D, Hwang S, et al. Interobserver agreement in the diagnosis of ovarian carcinoma types: impact of sub-specialization. Lab Invest 2012;92:292A.

21. Gilks CB, Ionescu DN, Kalloger SE, Köbel M, Irving J, Clarke B, 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;39:1239–51.

22. Gilks CB, Clarke BA, Han G, Köbel M, Longacre T, McCluggage WG, et al. Letter to the editor regarding 'Roh MH, Lassin Y, Miron A et al. High-grade fimbrial-ovarian carcinomas are unified by p53, PTEN and PAX2 expression'. Mod Pathol 2011;24:1281–2.

23. Schwartz DR, Kardia SLR, Shedden KA, Kuick R, Michailidis G, Taylor JMG, et al. Gene expression in ovarian cancer reflects both morphology and biological behavior, distinguishing clear cell from other poor-prognosis ovarian carcinomas. Cancer Res 2002;62: 4722–9.

24. Kalloger SE, Köbel M, Leung S, Mehl E, Gao D, Marcon KM, et al. Calculator for ovarian carcinoma subtype prediction. Mod Pathol 2011;24:512–21.

25. Kommoss S. Specialized pathology review in patients with ovarian cancer: results from a prospective study. Modern Pathol 2012; 25:281A.

26. Hammond MEH, Hayes DF, Dowsett M, Allred DC, Hagerty KL, Badve S, et al. American Society of Clinical Oncology/College of American Pathologists guideline recommendations for immunohistochemical testing of estrogen and progesterone receptors in breast cancer. Arch Pathol Lab Med 2010;134:907–22.

27. Sieh W, Köbel M, Longacre TA, Bowtell DD, Defazio A, Goodman MT, et al. Hormone-receptor expression and ovarian cancer survival: an Ovarian Tumor Tissue Analysis consortium study. Lancet Oncol 2013;14:853–62.

28. Kelemen LE, Goodman MT, McGuire V, Rossing MA, Webb PMon behalf of the Australian Cancer Study (Ovarian Cancer) , et al. Genetic variation in TYMS in the one-carbon transfer pathway is associated with ovarian carcinoma types in the ovarian cancer association consortium. Cancer Epidemiol Biomarkers Prev 2010;19:1822–30.

29. Lee KR, Tavassoli FA, Prat J, Dietel M, Gersell DJ, Karseladze AI , et al. Surface epithelial stromal tumors. In: Tavassoli FA, Deville P, editors. World health organization classification of tumours: Pathology and genetics of tumours of the breast and female genital tract. Lyon: IARC Press; 2003. p. 117–46.

30. Goode EL, Chenevix-Trench G, Hartmann LC, Fridley BL, Kalli KR, Vierkant RA, et al. Assessment of hepatocyte growth factor in ovarian cancer mortality. Cancer Epidemiol Biomarkers Prev 2011;20:1638–48.

31. Balogun N, Gentry-Maharaj A, Wozniak EL, Lim A, Ryan A, Ramus SJ, et al. Recruitment of newly diagnosed ovarian cancer patients proved challenging in a multicentre biobanking study. J Clin Epidemiol 2011;64:525–30.

32. Lo-Ciganic W-H, Zgibor JC, Bunker CH, Moysich KB, Edwards RP, Ness RB. Aspirin, nonaspirin nonsteroidal anti-inflammatory drugs, or acetaminophen and risk of ovarian cancer. Epidemiology 2012;23: 311–9.

33. Jones S, Wang T-L, Shih I-M, Mao T-L, Nakayama K, Roden R, et al. Frequent mutations of chromatin remodeling gene ARID1A in ovarian clear cell carcinoma. Science 2010;330:228–31.

34. Sim J, Wright CC. The kappa statistic in reliability studies: use, interpretation, and sample size requirements. Phys Ther 2005;85: 257–68.

35. Altman DG, McShane LM, Sauerbrei W, Taube SE. Reporting Recommendations for Tumor Marker Prognostic Studies (REMARK): explanation and elaboration. PLoS Med 2012;9:e1001216.

36. Freedman OC, Dodge J, Shaw P, Oza AM, Bernadini M, Klachook S, et al. Diagnosis of epithelial ovarian carcinoma prior to neoadjuvant chemotherapy. Gynecol Oncol 2010;119:22–5.

37. McCluggage WG. My approach to and thoughts on the typing of ovarian carcinomas. J Clin Pathol 2008;61:152–63.

38. Soslow RA. Histologic subtypes of ovarian carcinoma: an overview. Int J Gynecol Pathol 2008;27:161–74.

39. Gilks CB, Prat J. Ovarian carcinoma pathology and genetics: recent advances. Hum Pathol 2009;40:1213–23.

40. Soslow RA, Han G, Park KJ, Garg K, Olvera N, Spriggs DR, et al. Morphologic patterns associated with BRCA1 and BRCA2 genotype in ovarian carcinoma. Mod Pathol 2012;25:625–36.

41. Madore J, Ren F, Filali-Mouhim A, Sanchez L, Köbel M, Tonin PN, et al. Characterization of the molecular differences between ovarian endometrioid carcinoma and ovarian serous carcinoma. J Pathol 2010; 220:392–400.

BLOOD CANCER DISCOVERY

Biomarker-Based Ovarian Carcinoma Typing: A Histologic Investigation in the Ovarian Tumor Tissue Analysis Consortium

Martin Köbel, Steve E. Kalloger, Sandra Lee, et al.

Updated version	Access the most recent version of this article at: doi: 10.1158/1055-9965.EPI-13-0391
Supplementary Material	Access the most recent supplemental material at: http://cebp.aacrjournals.org/content/suppl/2013/07/25/1055-9965.EPI-13-0391.DC1

Cited articles	This article cites 40 articles, 9 of which you can access for free at: http://cebp.aacrjournals.org/content/22/10/1677.full#ref-list-1
Citing articles	This article has been cited by 3 HighWire-hosted articles. Access the articles at: http://cebp.aacrjournals.org/content/22/10/1677.full#related-urls

E-mail alerts	Sign up to receive free email-alerts related to this article or journal.
Reprints and Subscriptions	To order reprints of this article or to subscribe to the journal, contact the AACR Publications Department at pubs@aacr.org.
Permissions	To request permission to re-use all or part of this article, use this link http://cebp.aacrjournals.org/content/22/10/1677. Click on "Request Permissions" which will take you to the Copyright Clearance Center's (CCC) Rightslink site.