ORIGINAL PAPER

Agreement for tumor grade of ovarian carcinoma: analysis of archival tissues from the surveillance, epidemiology, and end results residual tissue repository

Rayna K. Matsuno · Mark E. Sherman · Kala Visvanathan · Marc T. Goodman ·
Brenda Y. Hernandez · Charles F. Lynch · Olga B. Ioffe · David Horio ·
Charles Platz · Sean F. Altekruse · Ruth M. Pfeiffer · William F. Anderson

Abstract

Background Emerging data suggest that ovarian cancers differ by tumor grade. However, the reliability of microscopic grade from paraffin tissue in the general medical community and as reflected in population-based cancer registries is unknown.

Methods We examined grade agreement between two gynecologic pathologists and the Surveillance Epidemiology and End Results Residual Tissue Repository (SEER). Grade agreement was assessed with percent observer agreement and kappa coefficients for 664 invasive ovarian carcinomas, using previously defined three-tier and two-tier grading systems. A random subset of ovarian carcinomas was selected to compare intra- and inter-pathologist agreement.

Results Five hundred and eighty-six of SEER's 664 tumors were confirmed invasive. Percent agreement was 49 % with fair kappa coefficient = 0.25 (95 % CI: 0.20–0.30) for the 664 tumors. Agreement improved slightly when restricted to the 586 confirmed invasive cancers; it was better for high grade than low grade tumors, for two-tier than three-tier grading systems, and within (66 %) than between study pathologists (43 %). Grade was not a robust independent predictor of ovarian cancer-specific survival.

Conclusions Grade agreement was fair between SEER and study pathologists irrespective of grading system. Recorded grade in SEER should be used with caution and is probably not a reliable metric for ovarian cancer epidemiology.

R. K. Matsuno · M. E. Sherman · R. M. Pfeiffer ·
W. F. Anderson (✉)
DHHS/NIH/National Cancer Institute/Division of Cancer
Epidemiology and Genetics, 6120 Executive Blvd (EPS 8036),
Rockville, MD 20852, USA
e-mail: wanderso@mail.nih.gov

R. K. Matsuno · M. T. Goodman · B. Y. Hernandez · D. Horio
Epidemiology Department, University of Hawaii Cancer Center,
Honolulu, HI, USA

R. K. Matsuno · K. Visvanathan
Department of Epidemiology, Johns Hopkins Bloomberg School
of Public Health, Baltimore, MD, USA

C. F. Lynch
Department of Epidemiology, The University of Iowa,
Iowa City, IA, USA

O. B. Ioffe
Department of Pathology, University of Maryland School
of Medicine, Baltimore, MD, USA

C. Platz
Department of Pathology, University of Iowa Hospitals
and Clinics, Iowa City, IA, USA

S. F. Altekruse
DHHS/NIH/National Cancer Institute/Division of Cancer
Control and Population Sciences, Rockville, MD, USA

Introduction

The molecular classification of ovarian epithelial carcinomas as low grade (type I) or high grade (type II) identifies two sets of cancers with contrasting incidence, molecular characteristics, and clinical outcomes [1–7]. The importance of ovarian cancer grade also transcends relevance for individuals with implications for cancer epidemiology and surveillance. It, therefore, would be important to establish the reliability of microscopic (i.e., histological) ovarian cancer grading in the general medical community and as reflected in population-based cancer registries such as the National Cancer Institute's SEER database. We, therefore, examined the agreement between recorded ovarian carcinoma grade in SEER's Residual Tissue Repository (RTR) and two independent gynecologic pathologists using three previously defined grading schemes; (1) the International Federation of Gynecology and Obstetrics grading system (FIGO) [8, 9], (2) Shimizu and Silverberg system [9–11], and (3) Malpica et al. [12–15] scheme.

Community-based pathologists commonly use the FIGO system; a three-tier grading scheme (low, intermediate, or high grade) that is modeled after the system for endometrial (uterine) carcinoma, which reflects the level of cellular organization into differentiated structures such as glands and papillae as opposed to solid sheets of tumor cells. Shimizu and Silverberg (herein referred to as SS) also devised a three-tier grading scheme that is similar to microscopic grading for breast carcinoma, incorporating histological architecture, nuclear cytology, and mitotic index. Malpica et al. at the M.D. Anderson Cancer Center (herein referred to as MDACC) proposed a two-tier system (low or high grade) for serous ovarian carcinomas [12–15], which is based upon a dualistic conceptual framework where low grade and high grade carcinomas proceed along two separate cancer pathways [1–7].

Materials and methods

The National Cancer Institute's SEER program established its Residual Tissue Repository (RTR) in 2003 to facilitate population-based cancer research using archival biospecimens [16, 17]. SEER's RTR included Tumor Registries in Hawaii, Iowa, and Los Angeles County. The Los Angeles County Tumor Registry did not participate in this study. We retrieved the available formalin-fixed and paraffin-embedded tissue blocks for primary invasive ovarian carcinomas in the Hawaii and Iowa Tumor Registries, excluding tubal and peritoneal tumors. There were 664 ovarian tumors; 516 were from the Hawaii Tumor Registry that were diagnosed from 1983 through 2004, which represented 38 % of all ovarian tumors in the Hawaii catchment area during that time period. The remaining 148 ovarian cases were derived from the Iowa Tumor Registry diagnosed from 1987 through 2003, representing 4 % of all ovarian tumors in the Iowa catchment area during that time period. Because SEER's RTR data were anonymized, the National Institutes of Health's Office of Human Subjects Research designated the project as exempt from IRB approval; nonetheless, IRB approvals were provided at the Universities of Hawaii and Iowa.

Demographic data included age at diagnosis, year of diagnosis, and race [White, Asian or Pacific Islander (API), and other/unknown]. Tumor characteristics consisted of the American Joint Committee on Cancer (AJCC) tumor, node, and metastasis (TNM) stage [18], and histological type, behavior, and grade according to the International Classification of Diseases for Oncology 3rd edition (ICD-O-3) [19]. AJCC ovarian cancer stages were stage I (tumors limited to one or both ovaries), stage II (involvement of one or both ovaries with pelvic extension and/or implants), stage III (involvement of one or both ovaries with microscopically confirmed peritoneal metastasis), and stage IV (distant metastasis, excluding peritoneal metastasis). AJCC guidelines specify 5 histological codes for the microscopic assessment of grade (G) that are independent of TNM stage: GX = unknown, G1 = well differentiated, G2 =moderately differentiated, G3 = poorly differentiated, and G4 = undifferentiated. ICD-O-3 morphology codes have six digits; the 1st four digits are for histological type, the fifth is for behavior (benign or malignant), and the sixth is for tumor grade. Ovarian carcinoma histological codes were serous (8441, 8460, and 8461), mucinous (8470, 8471, 8480, and 8481), endometrioid (8380, 8560, 8570, and 8381), clear cell (8310), and other (8010–8580 (excluding previously listed ICD-O-3 codes), 9000, and 9014). SEER abstracted tumor grade from the 6th ICD-O-3 digit as grade 1 (well differentiated), grade 2 (moderately differentiated, moderately well differentiated, or intermediate differentiation), grade 3 (poorly differentiated), and grade 4 (undifferentiated or anaplastic).

Pathology review

The primary study pathologist (MES) reviewed approximately three H&E stained slides per case (all designated as invasive carcinoma in SEER) to independently re-assess behavior (benign, borderline, or malignant), histological type, and grade for all 664 cases retrieved from SEER's RTR. A set of 19 % of the tumors (128 of 664) was selected for repeat pathology panel review. This set was constructed by taking a random sample of cancers stratified

by histological type, with oversampling of rarer types. Sampling fractions for each histological type were serous (10 %, 30 of 298), mucinous (40 %, 30 of 75), endometrioid (20 %, 20 of 97), clear cell (45 %, 28 of 62), and other carcinomas (15 %, 20 of 132). The selected ovarian cancers were reexamined by MES to evaluate intrapathologist agreement and reviewed by the second pathologist (OBI) to assess inter-pathologist agreement between MES and OBI. The pathologists had access to the gross pathologic descriptions but were masked to SEER's recorded behavior and AJCC stage since histological grade is meant to provide a microscopic assessment of ovarian cancer prognosis that is independent of stage [18].

For a complete description of the three ovarian cancer grading systems see supplemental table 1 for: (1) the International Federation of Gynecology and Obstetrics (FIGO) system [8, 9], (2) the Shimizu and Silverberg (SS) [9–11] system, and (3) the MD Anderson (MDACC) system [14]. In brief, the FIGO/SS grading schemes are three-tier systems that assign all histological types to "low," "intermediate," or "high" grade. The MDACC grading system is a two-tier system that assigns serous types to "low" or "high" grade.

Statistical analysis

We assessed the representativeness of the SEER RTR ovarian tumors with chi square tests for heterogeneity, comparing demographic and tumor characteristics for the recovered ovarian carcinomas in SEER's RTR with the ovarian carcinomas in the corresponding Hawaii and Iowa Tumor Registries. To compare the three-tier FIGO/SS grades to SEER grades 1 through 4, we reclassified SEER grade 1 as low, SEER grade 2 as intermediate, and SEER grades 3–4 as high (Table 1A). To compare the two-tier MDACC low and high grades with SEER grades 1–4, we dichotomized SEER grade 1 as low and SEER grades 2–4 as high (Table 1B). Finally, the three-tier FIGO/SS schemes were further collapsed to low and high (intermediate + high) grades for survival analyses.

Agreement was assessed as percent observer agreement (p_o) and Cohen's standard kappa coefficients (κ) [20]. Kappa coefficients ranged from 0.00 to 1.00 and were interpreted descriptively as poor $\kappa < 0.20$, fair $\kappa = 0.20$–0.40, moderate $k = 0.40$–0.60, good $\kappa = 0.60$–0.80, and very good $\kappa = 0.80$–1.00. The Kaplan–Meier estimator [21] was used to calculate ovarian cancer-specific survival by low or high grade for all AJCC stages combined and then by early stage (AJCC I + II) or late stage (AJCC III + IV). The log-rank test was used to assess survival differences by low and high grade [22].

Results

Descriptive statistics

The 664 ovarian carcinomas in SEER's RTR are shown in Table 2. Approximately three-quarters of the tumors were contributed by the Hawaii RTR (77 %, 516 of 664). Mean age at diagnosis was 59.6 years. Serous carcinomas accounted for 45 % of the ovarian tumors (298 of 664), 59 % were late stage and 40 % were high grade. Clear cell carcinomas were more common among APIs (13 %, 51 of 379) than among Whites (4 %, 11 of 282), $p < 0.01$. Women with serous carcinomas were diagnosed at older age, later stage, and higher grade than women with other histological types ($p < 0.05$). Compared to the 664 tumors in the Hawaii and Iowa RTRs, the ovarian cancers (5347) reported to the full Hawaii and Iowa Tumor Registries demonstrated a higher percentage of White women, slightly older ages at diagnosis, a lower proportion of serous tumors, and lower stage at diagnosis.

Pathology review

MES classified SEER's 664 invasive ovarian tumors as primary invasive ovarian carcinoma ($n = 586$), benign ($n = 3$), borderline ($n = 45$), and other ($n = 30$). The other category included ovarian cancers diagnosed at

Table 1 Graphical representation between the reclassified categories of SEER's 4-tier grading system to the 3-tier and 2-tier grading schemes

A: 4- to 3-tier grade comparison			B: 4- to 2-tier grade comparison			
SEER—4 tier	FIGO—3 tier	SS—3 tier	SEER—4 tier	FIGO—2 tier	SS—2 tier	MDAAC—2 tier
SEER grade 1	Low	Low	SEER grade 1	Low	Low	Low
SEER grade 2	Intermediate	Intermediate	SEER grade 2	High	High	High
SEER grade 3	High	High	SEER grade 3	High	High	High
SEER grade 4	High	High	SEER grade 4	High	High	High
Unknown	Unknown	Unknown	Unknown	Unknown	Unknown	Unknown

SEER grade 1 = well differentiated, SEER grade 2 = moderately differentiated, SEER grade 3 = poorly differentiated, SEER grade 4 = undifferentiated, FIGO the International Federation of Gynecology and Obstetrics, SS Shimizu and Silverberg, MDACC M. D. Anderson Cancer Center

Table 2 Distribution of demographic and tumor characteristics for all 664 ovarian tumors in the Hawaii (1983–2004) and Iowa (1987–2003) Surveillance, Epidemiology and End Results Residual Tissue Repositories (SEER RTR)

Histological type[a]	All cases		Serous		Mucinous		Endometrioid		Clear cell		Other	
	n	%[b]	n	%	n	%	n	%	n	%	n	%
Overall	664	100[c]	298	44.9	75	11.3	97	14.6	62	9.3	132	19.9
SEER RTR												
Hawaii	516	77.7[d]	230	77.2	58	77.3	80	82.5	60	96.8	88	66.7
Iowa	148	22.3	68	22.8	17	22.7	17	17.5	2	3.2	44	33.3
Year of diagnosis												
1983–1987	33	5.0	10	3.4	7	9.3	2	2.1	4	6.5	10	7.6
1988–2000	487	73.3	232	77.9	54	72.0	70	72.2	43	69.4	88	66.7
2001–2004	144	21.7	56	18.8	14	18.7	25	25.8	15	24.2	34	25.8
Age at diagnosis												
Mean age (SE)	59.6 (0.5)		61.4 (0.7)		57.4 (1.6)		55.6 (1.4)		55.2 (1.6)		62.1 (1.3)	
< 40 years	46	6.9	14	4.7	7	9.3	10	10.3	5	8.1	10	7.6
40–49 years	122	18.4	41	13.8	17	22.7	31	32.0	14	22.6	19	14.4
50–59 years	157	23.6	71	23.8	17	22.7	18	18.6	26	41.9	25	18.9
60–69 years	170	25.6	90	30.2	18	24.0	20	20.6	7	11.3	35	26.5
70 + years	169	25.5	82	27.5	16	21.3	18	18.6	10	16.1	43	32.6
Race												
White	282	42.5	137	46.0	27	36.0	31	32.0	11	17.7	76	57.6
Asian or Pacific Islander	379	57.1	160	53.7	47	62.7	66	68.0	51	82.3	55	41.7
Other/unknown	3	0.5	1	0.3	1	1.3	0	0.0	0	0.0	1	0.8
AJCC stage												
Early (I–II)	238	35.8	57	19.1	51	68.0	58	59.8	40	64.5	32	24.2
Late (III–IV)	391	58.9	231	77.5	17	22.7	37	38.1	18	29.0	88	66.7
Other/unknown	35	5.3	10	3.4	7	9.3	2	2.1	4	6.5	12	9.1
SEER grade												
Grade 1	50	7.5	8	2.7	23	30.7	18	18.6	0	0.0	1	0.8
Grade 2	165	24.8	56	18.8	32	42.7	38	39.2	11	17.7	28	21.2
Grade 3	265	39.9	145	48.7	5	6.7	30	30.9	15	24.2	70	53.0
Grade 4	97	14.6	56	18.8	2	2.7	5	5.2	17	27.4	17	12.9
Other/unknown	87	13.1	33	11.1	13	17.3	6	6.2	19	30.6	16	12.1

[a] ICD-O-3 codes for histological type: serous (8441, 8460, and 8461), mucinous (8470, 8471, 8480, and 8481), endometrioid (8380, 8560, 8570, and 8381), clear cell (8310), and other (8010-8580 (excluding previously listed ICD-O-3 codes), 9000, and 9014); [b] %, percentage of all cases; [c] row %; [d] column %

distant metastatic sites (i.e., primary carcinoma in the ovary was unavailable for microscopic examination), non-epithelial ovarian cancers, and non-ovarian carcinomas that were metastatic to the ovary. Grade agreement between the pathologist and SEER was similar for the FIGO (Table 3) and the SS systems (Table 4). Percent agreement with FIGO ranged from 24 % for clear cell carcinoma to 57 % for serous carcinoma with poor to fair kappa coefficients ranging from 0.00 to 0.29 (Table 3A).

Bar graphs along with an inserted contingency table are used in Fig. 1 to supplement FIGO grade agreement between MES and SEER (Table 3A). Percent observer agreement was $p_o = 49$ % between MES and SEER with 327 of 664 tumors in the diagonal of the contingency table

(Fig. 1). MES tended to grade lower than SEER. For example, MES moved 8 % of SEER-assigned high grade to MES low grade cancers (30 of 362) but did not move any SEER-assigned low grade to MES high grade tumors (0 of 362), Fig. 1. Consequently, p_o rose from low to high grade, for example, $p_o = 23$ % for low grade (26 for SEER and 112 for MES), 37 % for intermediate grade (57 for SEER and 153 for MES), and 77 % for high grade (219 for SEER and 284 for MES). Tumor grade was unknown for 177 of the ovarian carcinomas either because grade was not recorded by SEER or MES could not classify the tumor grade because of insufficient microscopic tissue (87 for SEER and 115 for MES with 25 mutually unknown for both SEER and MES). Grade agreement improved when

Table 3 Percent agreement and standard Kappa coefficients for MES versus SEER using FIGO three-tier grading schemes for: (A) all 664 ovarian tumors and (B) restricted to tumors with assigned grade (low, intermediate, or high) and also classified as invasive by the study pathologist

| | A: FIGO, all SEER tumors | | | | B: FIGO, restricted | | | |
| | Low/intermediate/high | | | | Low/intermediate/high | | | |
	N	p_o	κ (95 % CI)		n	p_o	κ (95 % CI)
Overall	664	0.49	0.25 (0.20, 0.30)		477	0.62	0.32 (0.26, 0.39)
Registry							
Hawaii	516	0.49	0.24 (0.18, 0.30)		382	0.61	0.31 (0.23, 0.38)
Iowa	148	0.51	0.29 (0.18, 0.39)		95	0.65	0.41 (0.28, 0.55)
p			0.43				0.18
AJCC stage							
Early (I–II)	238	0.45	0.26 (0.18, 0.34)		171	0.56	0.34 (0.23, 0.44)
Late (III–IV)	391	0.53	0.17 (0.10, 0.24)		292	0.65	0.21 (0.11, 0.30)
Other/unknown	35	0.37	0.15 (−0.07, 0.37)		14	0.71	0.53 (0.14, 0.91)
p			0.25				0.09
Histological type							
Serous	298	0.57	0.21 (0.13, 0.30)		235	0.66	0.19 (0.08, 0.31)
Mucinous	75	0.44	0.26 (0.12, 0.40)		37	0.59	0.37 (0.13, 0.61)
Endometrioid	97	0.46	0.26 (0.14, 0.39)		82	0.55	0.34 (0.20, 0.49)
Clear cell	62	0.24	0.00 (−0.11, 0.12)		39	0.36	0.04 (−0.12, 0.20)
Other	132	0.49	0.16 (0.05, 0.27)		84	0.71	0.31 (0.13 0.48)
p			0.01				0.04

MES Mark E. Sherman, SEER Surveillance, Epidemiology, and End Results, FIGO International Federation of Gynecology and Obstetrics grading scheme, *n* number, p_o observed agreement, κ standard Kappa coefficient, *p* two-sided *p* value for χ^2 test of heterogeneity across strata, AJCC American Joint Committee on Cancer

restricted to those tumors with known grade (low, intermediate, or high) and also labeled as invasive by MES (Table 3B), that is, $p_o = 62$ % and fair kappa coefficient = 0.32 (95 % CI: 0.26–0.39). Similar improvement was observed for SS grade (Table 4A compared to Table 4B). Grade agreements for the three-tier FIGO/SS systems did not improve substantively even when there was histological type agreement (serous, mucinous, endometrioid, or clear cell) between MES and SEER.

Percent agreement but not the kappa coefficient was generally better with the two-tier MDACC than three-tier FIGO system (Table 5). For example, overall agreement between the study pathologist and SEER grade with the MDACC system was $p_o = 64$ % with a poor kappa coefficient = 0.10 (95 % CI: 0.01–0.19), Table 5A, and improved when restricted to cases that were classified as invasive by the study pathologist ($p_o = 95$ %), Table 5B.

The randomly selected ovarian carcinomas (19 %, 128 of 664) were reviewed a second time by MES for intra-pathologist agreement and reviewed by OBI for inter-pathologist agreement between MES and OBI. Inter-pathologist agreement was similar to the agreement between MES and SEER, $p_o = 43$ % and fair kappa = 0.25 (95 % CI: 0.13 to 0.35). Intra-observer agreement for the 1st and 2nd review by MES

yielded $p_o = 66$ % and moderate $\kappa = 0.52$ (95 % CI: 0.41 to 0.64).

Ovarian cancer-specific survival

With 20 years of follow-up (Fig. 2), ovarian cancer-specific survival for the 586 confirmed invasive ovarian carcinomas was better among low grade than high grade tumors for MES (Fig. 2a, log-rank test $p < 0.001$) and for SEER (Fig. 2d, log-rank test $p < 0.001$). Long-term ovarian cancer survival was worse for MES low grade than for SEER low grade, for example, cumulative cancer-specific survival after 15 years of follow-up for MES low grade was 64 % (95 % CI: 55–75 %) and for SEER low grade was 90 % (95 % CI: 80–100 %). On the other hand, short-term cancer survival was similar for MES high grade and SEER high grade, for example, cancer-specific survival after 5 years of follow-up for MES high grade was 48 % (95 % CI: 44–54 %) and for SEER high grade was 51 % (95 % CI: 47–55 %). Re-categorizing SEER low grade to include SEER grade 1 + grade 2 and SEER high grade to include SEER grade 3 + grade 4 did not substantively alter the survival analysis (graphs available upon request).

Table 4 Percent agreement and standard Kappa coefficients for MES versus SEER using SS three-tier grading schemes for: (A) all 664 ovarian tumors and (B) restricted to tumors with assigned grade (low, intermediate, or high) and also classified as invasive by the study pathologist

	A: SS, all SEER tumors			B: SS, restricted		
	Low/intermediate/high			Low/intermediate/high		
	n	p_o	κ (95 % CI)	n	p_o	κ (95 % CI)
Overall	664	0.43	0.20 (0.15, 0.24)	447	0.57	0.27 (0.21, 0.34)
Registry						
Hawaii	516	0.42	0.18 (0.13, 0.24)	356	0.56	0.25 (0.17, 0.32)
Iowa	148	0.46	0.24 (0.14, 0.34)	91	0.62	0.37 (0.23, 0.52)
p			0.32			0.12
AJCC stage						
Early (I–II)	238	0.37	0.17 (0.09, 0.24)	159	0.48	0.25 (0.15, 0.35)
Late (III–IV)	391	0.48	0.15 (0.09, 0.22)	274	0.62	0.19 (0.10, 0.28)
Other/unknown	35	0.37	0.15 (−0.07, 0.36)	14	0.64	0.39 (−0.04, 0.81)
p			0.96			0.50
Histological type						
Serous	298	0.53	0.20 (0.12, 0.28)	224	0.63	0.18 (0.08, 0.29)
Mucinous	75	0.39	0.19 (0.05, 0.32)	34	0.56	0.29 (0.02, 0.56)
Endometrioid	97	0.41	0.21 (0.09, 0.33)	77	0.52	0.31 (0.16, 0.46)
Clear cell	62	0.11	−0.04 (−0.12, 0.04)	37	0.16	0.00 (−0.08, 0.08)
Other	132	0.41	0.11 (0.01, 0.20)	75	0.65	0.26 (0.10, 0.42)
p			<0.001			<0.001

MES Mark E. Sherman, SEER Surveillance, Epidemiology, and End Results, SS Shimizu and Silverberg grading scheme, n number, p_0 observed agreement, κ standard Kappa coefficient, p two-sided p value for χ^2 test of heterogeneity across strata, AJCC American Joint Committee on Cancer

When stratified by AJCC early and late stage, grade was not a robust independent predictor of cancer-specific survival (Fig. 2b, c, e, f). For those tumors that were designated as benign by MES, that is, benign ($n = 3$) or borderline ($n = 45$), ovarian cancer-specific survival was 90 % with only 5 of 48 recorded ovarian cancer deaths during follow-up.

Discussion

Our study demonstrated several interesting findings regarding ovarian carcinoma grade agreement between SEER and two independent gynecological pathologists. First, similar to other clinical studies [23], grade agreement was only fair irrespective of grading system and histological type. For example, Gilks et al. [23] reported inter-observer kappa coefficients of 0.26 and 0.40 for FIGO and SS grading systems, respectively. Second, agreement improved when restricted to tumors with known grade (low, intermediate, or high) and also classified as invasive by the study pathologist. Third, agreement was better for high grade than low grade tumors, two-tier than three-tier grading systems, and intra- than inter-pathologist comparison. Finally, tumor grade was not a strong independent prognostic factor apart from AJCC stage.

Several factors may have affected the generalizability of our results but not the internal validity of agreement for grade. The 664 ovarian tumors from SEER's RTR represented only 38 and 4 % of the ovarian tumors in the Hawaii and Iowa Tumor Registries, respectively. More than 75 % of the data were contributed by the Hawaii Tumor Registry, enriching the study with APIs and clear cell carcinomas, a histological type that is more common among Japanese than White women [24, 25]. There were differences between patient characteristics in the RTR and the complete Hawaii and Iowa Tumor Registries, but grade agreement did not differ by any of these factors.

Percent observer agreement was generally higher than kappa coefficients, reflecting two limitations of the kappa statistic [26, 27]. First, though the kappa statistic attempts to measure the amount of non-random agreement [28], one limitation occurs when the categories for a given variable are not equally distributed [27]. Given that high grade is proportionately more dominant than low or intermediate grade, high grade ovarian tumors would be more likely by chance alone. The second limitation arises with imbalance of the row and column totals of a contingency table (e.g., Fig. 1 insert) [27]. As shown in Fig. 1, there is an inequity in the totals for MES low, intermediate, high, and unknown grade of 16.9, 23.0, 42.8, and 17.3 % compared to the corresponding totals for SEER grade of 7.5, 24.8, 54.5, and 13.1 %.
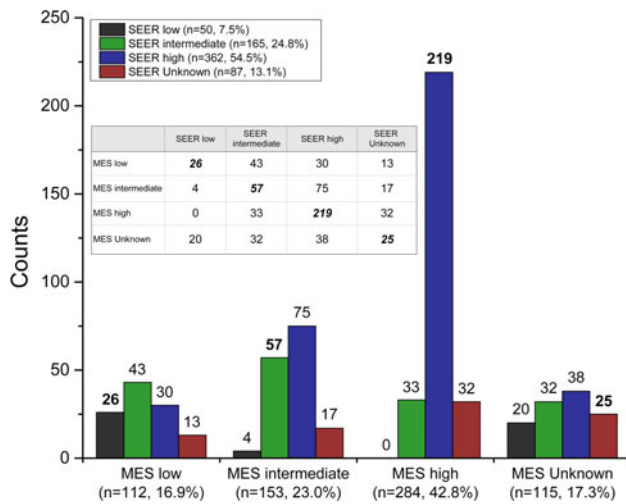
Fig. 1 Contingency table for grade agreement between the study pathologist (MES) and SEER for all 664 ovarian tumors in the Hawaii and Iowa RTRs. The crosstab or contingency table (*insert*) shows percent observer agreement between MES and SEER in the diagonals with disagreements in the off diagonals. *Bold fonts* in the *bar graph* show greater percent agreement between MES and SEER for high grade (77 %, 219 of 284) than low grade (23 %, 26 of 112) or intermediate grade (37 %, 57 of 153). Grading was unknown and/or missing for 177 ovarian tumors

Of note, percent observer agreement increased from low to high grade, possibly reflecting the fact that SEER grade comes from community-based pathologists with more clinical information than was available to the study pathologists, for example, AJCC stage. Even when conditioned upon early and late stage, we observed better agreement between MES and SEER for high than low grade

(Table 3A). More specifically, for early stage cancers, percent observer agreement was 29 % for low grade and 71 % for high grade. For late-stage cancers, percent observer agreement was 11 % for low grade and 75 % for high grade. The knowledge of stage along with a heightened awareness of poor outcomes for advanced stage ovarian carcinomas may have influenced SEER's pathologists to avoid classifying late-stage tumors as low grade. If true, this would tend to yield more conservative low grade carcinomas because of their association with early stage disease. Figure 2 supports this conjecture since cancer-specific survival was better for SEER early stage than for MES early stage, whereas survival was similar for SEER and MES late stage. Indeed, prior reports from individual pathology laboratories have found lower survival for low grade tumors than reported in SEER [29, 30]. Admixing benign and borderline tumors with low grade carcinomas [31] also would tend to improve prognosis for low grade. In fact, 90 % cancer-specific survival for the reclassified benign and borderline tumors is clearly better than otherwise would be expected for typical invasive ovarian cancers.

Percent observer p_o agreement was generally better for the two-tier MDACC than the three-tier staging schemes. Though we cannot exclude better agreement by chance alone, improvement with the two-tier scheme might possibly reflect the dualistic nature of ovarian cancer. Contemporary clinicopathologic and molecular models implicate two main carcinogenic pathways by type I (low grade) or type II (high grade) [1–7]. Type I low grade cancers are believed to arise through a stepwise sequence from adenoma to borderline tumor to invasive cancer and are associated with oncogenic

Table 5 Percent agreement and standard Kappa coefficients for MES versus SEER using MDACC two-tier grading schemes for: (A) all 298 ovarian *serous* carcinomas recorded in SEER and (B) restricted to *serous* tumors with assigned grade (low or high) and also classified as invasive by the primary study pathologist

	(A): MDACC, all serous tumors			(B): MDACC, restricted		
	Low/high			Low/high		
	n	p_0	κ (95 % CI)	n	p_0	κ (95 % CI)*
Overall						
Registry	298	0.64	0.10 (0.01, 0.19)	182	0.95	0.00 ($\sim$)
Hawaii	230	0.69	0.13 (0.02, 0.24)	156	0.96	0.00 ($\sim$)
Iowa	68	0.47	−0.01 (−0.17, 0.14)	26	0.88	0.00 ($\sim$)
p			0.18			$\sim$
AJCC stage						
Early (I–II)	57	0.56	0.17 (0.00, 0.35)	28	0.93	0.00 ($\sim$)
Late (III–IV)	231	0.65	0.08 (−0.02, 0.18)	146	0.95	0.00 ($\sim$)
Other/unknown	10	0.80	0.00 (0.00, 0.00)	8	1.00	0.00 ($\sim$)
p			0.35			$\sim$

MES Mark E. Sherman, SEER Surveillance, Epidemiology, and End Results, MDACC MD Anderson Cancer Center grading scheme, *n* number, p_o observed agreement, κ standard Kappa coefficient, * all cases were designated as high grade by SEER, *p* two-sided *p* value for χ^2 test of heterogeneity across strata, AJCC American Joint Committee on Cancer

Fig. 2 Ovarian cancer-specific survival with 95 % confidence limits by low and high grades for the study pathologist (MES) and SEER: all AJCC stages combined (**a**, **d**), early AJCC stages (**b**, **e**), and late AJCC stages (**c**, **f**)
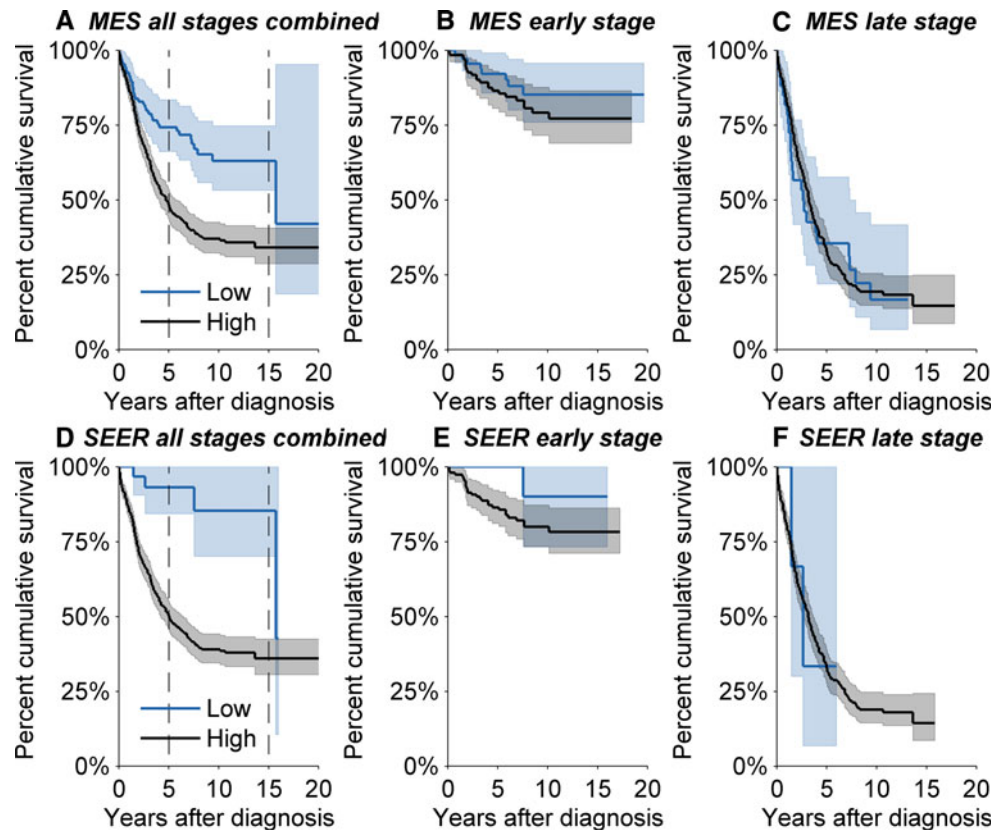
mutations that impact cell proliferation (*KRAS* and/or *BRAF*) [3, 32]. Type II high grade tumors constitute the majority of invasive ovarian cancers in the general population and typically show molecular changes that are associated with genetic instability [10, 11].

In sum, grade agreement was fair to moderate between SEER and two independent gynecological pathologists. Agreement improved with higher grade, a two-tier grading scheme for serous tumors, and when restricted to tumors that were re-classified as invasive by the study pathologist. Additionally, grade also was not a robust independent predictor of survival. Consequently, though molecular studies and individual clinical outcomes differ by grade, recorded grade in SEER should be used with caution and may not be a reliable metric for population-based cancer epidemiology. Nonetheless, given the compelling molecular evidence for type I and II ovarian cancers, the results of this study suggest that epidemiologists may need to supplement the microscopic assessment of grade for ovarian cancer with additional biological information such as protein and/or gene expression profiles similar to "genomic grade" for breast cancer [33, 34].

References

1. Singer G, Kurman RJ, Chang HW, Cho SK, Shih Ie M (2002) Diverse tumorigenic pathways in ovarian serous carcinoma. Am J Pathol 160(4):1223–1228 (Epub 2002/04/12)
2. ShihIe M, Kurman RJ (2004) Ovarian tumorigenesis: a proposed model based on morphological and molecular genetic analysis. Am J Pathol 164(5):1511–1518
3. Kurman RJ, Shih IM (2008) Pathogenesis of ovarian cancer: lessons from morphology and molecular biology and their clinical implications. Int J Gynecol Pathol 27:151–160
4. Landen CN Jr, Birrer MJ, Sood AK (2008) Early events in the pathogenesis of epithelial ovarian cancer. J Clin Oncol 26(6):995–1005
5. Levanon K, Crum C, Drapkin R (2008) New insights into the pathogenesis of serous ovarian cancer and its clinical impact. J Clin Oncol 26(32):5284–5293
6. Kuo KT, Guan B, Feng Y, Mao TL, Chen X, Jinawath N et al (2009) Analysis of DNA copy number alterations in ovarian serous tumors identifies new molecular genetic changes in low-grade and high-grade carcinomas. Cancer Res 69(9):4036–4042
7. Grimley PM, Matsuno RK, Rosenberg PS, Henson DE, Schwartz AM, Anderson WF (2009) Qualitative age interactions between low and high grade serous ovarian carcinomas. Cancer Epidemiol Biomarkers Prev 18(8):2256–2261
8. Classification and staging of malignant tumors in the female pelvis, accepted by the General Assembly of FIGO in New York, April 12, 1970. Acta Obstet Gynecol Scand 50(1):1–7

9. Silverberg SG (2000) Histopathologic grading of ovarian carcinoma: a review and proposal. Int J Gynecol Pathol 19(1):7–15

10. Shimizu Y, Kamoi S, Amada S, Akiyama F, Silverberg SG (1998) Toward the development of a universal grading system for ovarian epithelial carcinoma: testing of a proposed system in a series of 461 patients with uniform treatment and follow-up. Cancer 82(5):893–901

11. Shimizu Y, Kamoi S, Amada S, Hasumi K, Akiyama F, Silverberg SG (1998) Toward the development of a universal grading system for ovarian epithelial carcinoma. I. Prognostic significance of histopathologic features–problems involved in the architectural grading system. Gynecol Oncol 70(1):2–12

12. Malpica A (2008) Grading of ovarian cancer: a histotype-specific approach. Int J Gynecol Pathol 27:175–181

13. Malpica A, Deavers MT, Tornos C, Kurman RJ, Soslow R, Seidman JD et al (2007) Interobserver and intraobserver variability of a two-tier system for grading ovarian serous carcinoma. Am J Surg Pathol 31(8):1168–1174

14. Malpica A, Deavers MT, Lu K, Bodurka DC, Atkinson EN, Gershenson DM et al (2004) Grading ovarian serous carcinoma using a two-tier system. Am J Surg Pathol 28(4):496–504

15. McCluggage WG (2011) Morphological subtypes of ovarian carcinoma: a review with emphasis on new developments and pathogenesis. Pathology (Phila) 43(5):420–432 (Epub 2011/07/01)

16. SEER-Surveillance Epidemiology and End Results (2012) [cited 2012 February 06]; Available from: http://seer.cancer.gov/

17. Goodman MT, Hernandez BY, Hewitt S, Lynch CF, Cote TR, Frierson HF Jr et al (2005) Tissues from population-based cancer registries: a novel approach to increasing research potential. Hum Pathol 36(7):812–820

18. AJCC (2011) Ovary and primary peritoneal carcinoma. In: Edge SB, Byrd DR, Compton CC, Fritz AG, Greene FL, Trotti A (eds) AJCC cancer staging manual, 7th edn. Springer, New York, pp 419–428

19. International Classification of Diseases for Oncology 3rd ed (2000) In: Fritz A, Percy C, Jack A, Shanmugaratnam K, Sobin L, Parkin DM, Whelan S (eds) U.S. Interim Version 2000. World Health Organization, Geneva

20. Cohen CJ (1960) A coefficient of agreement for nominal scales. Educ Psychol Measur 20:37–46

21. Kaplan EL, Meier P (1958) Nonparametric estimation from incomplete observations. J Am Stat Assoc 53:457–481

22. Peto R, Peto J (1972) Asymptomatically efficient rank invariant test procedures. J Roy Stat Soc A 135((A)):185–198

23. Gilks CB, Ionescu DN, Kalloger SE, Kobel M, Irving J, Clarke B et al (2008) Tumor cell type can be reproducibly diagnosed and is of independent prognostic significance in patients with maximally debulked ovarian carcinoma. Hum Pathol 39(8):1239–1251

24. Sugiyama T, Kamura T, Kigawa J, Terakawa N, Kikuchi Y, Kita T et al (2000) Clinical characteristics of clear cell carcinoma of the ovary: a distinct histologic type with poor prognosis and resistance to platinum-based chemotherapy. Cancer 88(11):2584–2589 (Epub 2000/06/22)

25. Soslow RA (2008) Histologic subtypes of ovarian carcinoma: an overview. Int J Gynecol Pathol 27:161–174

26. Cicchetti DV, Feinstein AR (1990) High agreement but low kappa: II. Resolving the paradoxes. J Clin Epidemiol 43(6):551–558 (Epub 1990/01/01)

27. Feinstein AR, Cicchetti DV (1990) High agreement but low kappa: I. The problems of two paradoxes. J Clin Epidemiol 43(6):543–549 (Epub 1990/01/01)

28. Last JM (1995) A dictionary of epidemiology, 3rd edn. Oxford University Press, Oxford

29. Seidman JD, Yemelyanova A, Cosin JA, Smith A, Kurman RJ (2012) Survival rates for international federation of gynecology and obstetrics stage III ovarian carcinoma by cell type: a study of 262 unselected patients with uniform pathologic review. Int J Gynecol Cancer 22(3):367–371 (Epub 2012/01/13)

30. Kobel M, Kalloger SE, Boyd N, McKinney S, Mehl E, Palmer C et al (2008) Ovarian carcinoma subtypes are different diseases: implications for biomarker studies. PLoS Med 5(12):e232

31. Leitao MM Jr, Boyd J, Hummer A, Olvera N, Arroyo CD, Venkatraman E et al (2004) Clinicopathologic analysis of early-stage sporadic ovarian carcinoma. Am J Surg Pathol 28(2):147–159

32. Singer G, Shih Ie M, Truskinovsky A, Umudum H, Kurman RJ (2003) Mutational analysis of K-ras segregates ovarian serous carcinomas into two types: invasive MPSC (low-grade tumor) and conventional serous carcinoma (high-grade tumor). Int J Gynecol Pathol 22(1):37–41

33. Sotiriou C, Pusztai L (2009) Gene-expression signatures in breast cancer. N Engl J Med 360(8):790–800

34. Sotiriou C, Wirapati P, Loi S, Harris A, Fox S, Smeds J et al (2006) Gene expression profiling in breast cancer: understanding the molecular basis of histologic grade to improve prognosis. J Natl Cancer Inst 98(4):262–272