Gynecologic Oncology xxx (2018) xxx–xxx

Contents lists available at ScienceDirect

Gynecologic Oncology

journal homepage: www.elsevier.com/locate/ygyno

Inter-pathologist and pathology report agreement for ovarian tumor characteristics in the Nurses' Health Studies

Mollie E. Barnard [a], Alexander Pyden [b], Megan S. Rice [c,d], Miguel Linares [c], Shelley S. Tworoger [a,e], Brooke E. Howitt [f,g], Emily E. Meserve [g,h], Jonathan L. Hecht [b,*]

[a] Department of Epidemiology, Harvard T.H. Chan School of Public Health, 677 Huntington Ave, Boston, MA 02115, United States of America
[b] Department of Pathology, Beth Israel Deaconess Medical Center, 330 Brookline Ave, Boston, MA 02215, United States of America
[c] Clinical and Translational Epidemiology Unit, Department of Medicine, Massachusetts General Hospital, Bartlett Ext., 9th Floor 55 Fruit St., Boston, MA 02114, United States of America
[d] Channing Division of Network Medicine, Department of Medicine, Brigham and Women's Hospital, 75 Francis St, Boston, MA 02115, United States of America
[e] Department of Cancer Epidemiology, Moffitt Cancer Center, 12902 USF Magnolia Drive, Tampa, FL 33612, United States of America
[f] Department of Pathology, Stanford University School of Medicine, 300 Pasteur Drive, Lane 235, Stanford, CA 94305-5324, United States of America
[g] Department of Pathology, Brigham and Women's Hospital, 75 Francis St, Boston, MA 02115, United States of America
[h] Department of Anatomic & Clinical Pathology, Maine Medical Center, Tufts University School of Medicine, 22 Bramhall Street, Portland, ME 04102-3175, United States of America

HIGHLIGHTS

- Ovarian cancer registries often rely on morphologic classification of histotype/grade from local pathology reports.
- Central review pathologists' reproducibility and agreement with original pathology reports is good.
- Agreement with original reports varied by histologic features, but not by slide quality.

ARTICLE INFO

Article history:
Received 30 May 2018
Received in revised form 28 June 2018
Accepted 1 July 2018
Available online xxxx

Keywords:
Ovarian cancer
Grade
Histotype
Agreement

ABSTRACT

Background. Grade and histotype of ovarian carcinomas are often used as surrogates of molecular subtypes. We examined factors affecting pathologists' reproducibility in two prospective studies.

Methods. Two pathologists independently reviewed slides from 459 incident ovarian cancers in the Nurses' Health Study (NHS) and NHSII. We described agreement on tumor characteristics using percent agreement and Cohen's standard kappa (κ) coefficients. We used logistic regression, with disagreement as the outcome, to evaluate the contribution of case and tumor characteristics to agreement.

Results. Inter-rater agreement was 95% (κ = 0.81) for carcinoma versus borderline, 89% (κ = 0.58) for grade and 85% (κ = 0.71) for histotype. Inter-rater grading disagreement was higher for non-serous histotypes (OR = 4.66, 95% CI 2.09–10.36) and lower for cancers with bizarre atypia (OR = 0.13, 95% CI 0.04–0.38). Agreement with original pathology reports was 94% (κ = 0.73) for carcinoma versus borderline, 78% (κ = 0.60) for histotype, and 79% (κ = 0.24) for grade. Grading disagreement was significantly lower for tumors with 'solid, pseudoendometrioid or transitional' (SET) architecture (OR = 0.08, 95%CI 0.01–0.84). Date of original diagnosis, hospital type, number of slides available for review, tumor stage, and slide quality were not related to agreement.

Conclusion. Overall, inter-rater agreement for tumor type and grade for archival tissue specimens was good. Agreement between the consensus review and original pathology reports was lower. Factors contributing to grading disagreement included non-serous histotype, absence of bizarre atypia, and absence of SET architecture.

1. Introduction

Ovarian cancer is a heterogeneous disease [1]. Although sub-classification using molecular diagnostics is an emerging trend [2–4], existing data repositories used in clinical and epidemiologic research largely rely on the morphologic classifications of histotype (e.g., serous, endometrioid, clear cell, mucinous) and grade reported in

* Corresponding author.
E-mail addresses: mollie.barnard@mail.harvard.edu (M.E. Barnard), apyden@bidmc.harvard.edu (A. Pyden), mrice1@mgh.harvard.edu (M.S. Rice), miguel.linares@ucsf.edu (M. Linares), Shelley.Tworoger@moffitt.org (S.S. Tworoger), bhowitt@stanford.edu (B.E. Howitt), EMeserve@mmc.org (E.E. Meserve), jlhecht@bidmc.harvard.edu (J.L. Hecht).

https://doi.org/10.1016/j.ygyno.2018.07.003
0090-8258/

pathology reports or tumor registries [5,6]. Morphologic classification of ovarian cancer, however, is subjective.

Reproducibility studies evaluating inter-observer agreement among expert pathologists have shown excellent agreement for histotype ($\kappa = 0.77–0.97$) [5,7], and Köbel et al. reported that adding a panel of immunohistochemical stains can further improve inter-observer agreement [8]. In contrast, agreement for grade is only fair. For example, reported agreement on International Federation of Gynecology and Obstetrics (FIGO) grade ranges from $\kappa = 0.25–0.26$ [5,6].

Inter-observer agreement and agreement with the original pathology report in the setting of an unselected population or central review during clinical trials or epidemiologic studies have not been extensively reported. Kommoss et al. included a central pathology review during a phase 3 drug trial and cited 65% agreement on histotype with the greatest discrepancy for serous misclassified as endometrioid type. Agreement was not significantly better for cases diagnosed at academic versus private hospitals [9]. Similar results were observed by Lopez-Guerrero et al. for a central pathology review of early-stage ovarian carcinoma in the Spanish Group for Ovarian Cancer Research (GEICO), with concordance of 76% ($\kappa = 0.50$; $p < 0.0001$) [10]. In contrast, Young et al. and Ozols et al. reported little to no misclassification of histotype between central pathology review and the original report [11,12]. In general, borderline tumors were rarely misclassified as carcinoma versus borderline [9,11]. In the only study to assess grade, an evaluation in the Surveillance Epidemiology and End Results (SEER) Residual Tissue Repository, Matsuno et al. found only fair grading agreement between expert review and original pathologist diagnosis (49% agreement, $\kappa = 0.25$) [6].

In addition, relatively few studies have evaluated predictors of grading agreement. Matsuno et al. commented that grading agreement between expert pathologist review and SEER data (based on the original diagnosis) was better for tumors identified as carcinoma (versus borderline) by the pathology review (agreement in 62% of cases and $\kappa = 0.32$), but was not improved when restricted to cases where the reviewer agreed with SEER on tumor histotype [6]. There have been no reports on the effects of variables related to data collection, such as time since original diagnosis, hospital type, slide quality and number of slides available. Further, while high-grade tumors tend to have greater mitotic activity and multinucleated cells [13], the effects of tumor characteristics on grading agreement, such as extensive necrosis, bizarre atypia or with 'solid, pseudoendometrioid or transitional' (SET) architecture, have not been analyzed. This may be especially relevant for studies relying on pathology report data because cancers with these features were formerly classified as high-grade endometrioid carcinoma, but are now being diagnosed as serous [14].

The Nurses' Health Studies (NHS/NHSII) are large, prospective cohort studies that request tumor tissue and pathology reports from reported cases of epithelial ovarian cancer during up to 40 years of follow-up. The NHS/NHSII have obtained tumors slides or blocks from over 450 ovarian cancer cases, allowing for assessment of agreement in the evaluation of tumor characteristics. Using data from NHS/NHSII, we examined agreement on ovarian carcinoma versus borderline, histotype and grade [15] independently assigned by two expert gynecologic pathologists with access to the same archival tissue slides. We also quantified tumor histotype and grading agreement between the consensus of expert gynecologic pathologists and information abstracted from the original pathology reports.

To better understand potential predictors of poor inter-rater agreement, we examined factors that may affect grading agreement in an epidemiologic setting, including variables related to data collection (i.e., time since original diagnosis, hospital type, slide quality, number of slides available), and tumor characteristics (i.e., necrosis, bizarre atypia, SET architecture, tumor stage). Further, we anticipated greater inter-pathologist disagreement among cases with fewer or older slides. We also expected that cases with SET morphology or necrosis would be graded higher on the grading scale by the consensus review of pathologists relative to data abstracted from pathology reports [16].

2. Materials and methods

2.1. Study population

As of August 2010, the latest cancer diagnosis date for cases in this study, the NHS/NHSII included 1311 confirmed cases of incident epithelial ovarian cancer with pathology report data. The NHS/NHSII requested tumor tissue from all of these cases, and received specimens from 459 of them. Common reasons that tumor tissue was not available included, death of the patient, destruction of the tissue block, or inability of the hospital to send a tissue sample [17]. Of the 459 cases with slides available for review by two expert gynecologic pathologists, 41 did not have record-based information on histopathologic features, due to confirmation via linkage to the relevant tumor registry or death certificates.

2.2. Slide review

Between October 2014 and May 2016, two expert gynecologic pathologists (JH and EM) reviewed slides from all 459 cases with tumor slides. Both were blinded to morphologic assessment on original pathology reports. The reviewing pathologists assigned values for carcinoma versus borderline, histotype, grade, and grading-related features: gland formation, nuclear atypia and mitotic rate. They also recorded the number of slides available for review, commented on slide quality, and assessed tumor architectural features such as SET (percent solid, pseudoendometrioid or transitional architecture), geographic necrosis, and bizarre atypia. Cases for which the two reviewers disagreed on carcinoma versus borderline, histotype, or grade, were adjudicated by a third gynecologic pathologist (BH). Extreme disagreements on tumor architecture, as defined by >30% difference in percent SET, were also adjudicated by the third pathologist. The agreed upon or arbitrated values were treated as the consensus values of the three reviewing pathologists. For the 418 cases from whom we also obtained pathology reports, we abstracted values for date of original diagnosis, carcinoma versus borderline, histotype, grade, and stage, when reported. The grading system was generally not specified in the original reports, but was likely the 3-tiered FIGO system given the prevalence of its usage in the United States.

The pathologist review of archival tumor specimens included assignment of histopathologic features. Tumors were categorized as either carcinoma or borderline. Tumor histotypes were described as serous, endometrioid, clear cell, mucinous, or other. Grade was reported for all epithelial ovarian carcinomas using a three-tier grading system in which serous carcinomas were assigned high-grade or low-grade (grades 1 or 3), clear cell carcinomas were assigned grade 3, and endometrioid and mucinous carcinoma were graded based on architecture according to the FIGO system (Grade 1 showed <5% of solid tumor growth; Grade 2 with 5%–50%, and Grade 3 with >50%) [18]. Reviewing pathologists were blinded to morphologic assessment on original pathology reports when reviewing the slides.

Our grading scheme is an modification of the WHO 2014 grading system [19] to allow for comparison to the original pathology reports, the majority of which reported on a the 3-tiered grading scale. In WHO 2014, low-grade and high-grade serous are distinct diseases, so a numerical grade is not assigned; clear cell carcinomas are all considered high-grade for which a numerical grade is not assigned, and endometrioid and mucinous carcinomas are graded on a 3-tiered system [19]. In contrast, a numerical grade was assigned in the majority of original reports regardless of histotype. The grading system in the original reports was generally not specified, but was likely the 3-tiered FIGO system (including grade 2 serous carcinoma) in older cases, or a modification similar to ours in recent cases. Assignment of a numerical grade was often included to accommodate the data format of existing

Please cite this article as: M.E. Barnard, et al., Inter-pathologist and pathology report agreement for ovarian tumor characteristics in the Nurses' Health Studies, Gynecol Oncol (2018), https://doi.org/10.1016/j.ygyno.2018.07.003