Human Pathology (2019) **94**, 51–57

ELSEVIER

Human PATHOLOGY

www.elsevier.com/locate/humpath

Original contribution

Interobserver variability in breast carcinoma grading results in prognostic stage differences☆,☆☆

Kimmie Rabe MD[a], Olivia L. Snir MD[b], Veerle Bossuyt MD[c], Malini Harigopal MD[a], Romulo Celli MD[a], Emily S. Reisenbichler MD[a,*]

[a]*Department of Pathology, Yale University School of Medicine New Haven, New Haven, CT*
[b]*Department of Pathology, Oregon Health and Science University School of Medicine, Portland, OR*
[c]*Department of Pathology, Massachusetts General Hospital, Boston, MA*

Received 8 July 2019; revised 19 August 2019; accepted 6 September 2019

Keywords:
Breast carcinoma;
Histologic grade;
Interobserver variability;
Prognostic stage;
AJCC

Summary The AJCC Cancer Staging Manual 8th edition included tumor grade in the pathologic prognostic stage for breast carcinomas. Due to the known subjectivity of tumor grading, we aimed to assess the degree of interobserver agreement for invasive carcinoma grade among pathologists and determine its effect on pathologic prognostic stage. One hundred consecutive cases of invasive stage II carcinomas were independently graded twice, with an 4-week intervening wash-out period, by 6 breast pathologists utilizing established Nottingham grading criteria. Inter- and intra-observer variability was determined for overall grade and for each of the 3 scoring components. Interobserver variability was good to very good (κ range = 0.582–0.850) with even better intra-observer variability (mean κ = 0.766). Tubule score was the most reproducible element (κ = 0.588). Complete concordance was reached in 54 cases and 58 cases in rounds 1 and 2 respectively. In round 1 this resulted in different pathologic prognostic stage in only 25 of discordant cases, 18 of which were stage IA versus IB. In conclusion, grading agreement between pathologists was good to very good and discordant grades resulted in small changes to pathologic prognostic stage.

1. Introduction

Over the last decade, key advances in molecular diagnostics have expanded our understanding of the importance of biologic tumor factors. Treatment decisions for patients with breast cancer are dictated by tumor biology and as such, the reporting of prognostic and predictive tumor biomarkers, including estrogen receptor (ER), progesterone receptor (PR), and epidermal growth factor receptor 2 (HER2) is essential to guide patient management. The most recent, 8th edition of the *American Joint Committee on Cancer (AJCC) Staging*

☆ Competing interests: The authors declare no conflict of interest.
☆☆ This work was presented as a platform at the USCAP annual meeting in National Harbor, MA in March 2019 and therefore published in abstract form in the *Modern Pathology* supplement edition but has not otherwise been published previously.

* Corresponding author at: Yale University, PO Box 208023, New Haven, CT 06520-8023.
E-mail address: emily.reisenbichler@yale.edu (E. S. Reisenbichler).

https://doi.org/10.1016/j.humpath.2019.09.006
0046-8177/

Manual [1] emphasizes the importance of these tumor features, incorporating them into the staging of breast cancer. Traditional anatomic evaluation of tumor size, lymph node status and presence of distant metastasis (TNM) is now combined with histologic grade, biomarker status and genomic assays to form two new prognostic stage categories, the pathological prognostic stage (PPS) and the clinical prognostic stage. This represents a significant change in the codification of breast cancer staging and reflects recent advances in our understanding of the biology of the disease.

The concept of tumor grading was originally applied to breast carcinomas over a century ago in an effort to determine "the degree of malignancy" and to predict patient outcomes based on microscopic tumor features. Greenough utilized multiple histologic characteristics to classify tumors into three groups showing "high, medium and low malignancy", associating increasing degree of malignancy with a decreasing potential for surgical cure [2]. This grading concept was further refined by Scarff and Patey in 1928 and Bloom and Richardson in the 1950s to identify three key histologic features, tubule formation, nuclear variation and mitotic activity as prognostic tumor features [3-5]. Further modifications by Elston [6] and Elston and Ellis [7] led to our current understanding of tumor grading. This grading scheme is referred to as the Nottingham modification of the Scarff-Bloom-Richardson system. Use of the Nottingham grading system is recommended by the AJCC, the World Health Organization (WHO), and the College of American Pathologists (CAP).

We currently know that tumor size, histologic grade and lymph node status are the most important prognostic measures in breast carcinoma. Histologic grade gives vital information when combined with tumor size and lymph node status, but has also been proven as an independent prognostic indicator. However, there have been concerns over grading inconsistency between observers [8]. These concerns were cited as a specific reason for omission of tumor grade as a staging component in the 7th edition of the AJCC Staging Manual [9]. In this study, we focus on the inclusion of histologic grade, an important marker of tumor biology, into PPS assignment. Specifically, we aimed to assess the degree of interobserver agreement for tumor grade among a dedicated group of breast pathologists and how variability in grading may affect the PPS.

2. Materials and methods

Following institutional review board approval, a retrospective review of the electronic pathology database from our institution was performed. From the files, sequential excision specimens of invasive breast carcinoma with an anatomic stage of at least II, were identified from 2010–2011. Tumor stage was defined by the 7th Edition AJCC Staging Manual in use at the time of the original case review. Tumors from patients who underwent neoadjuvant chemotherapy were excluded.

Breast specimens, at the time of the primary histologic evaluation, were fixed in 10% buffered formalin and routinely processed. Time to fixation was not known for these specimens. One representative H&E stained slide containing the largest tumor section was selected from each case. The surgical pathology reports were reviewed and pathologic data, including hormone receptor and HER2 status, tumor size and lymph node involvement was recorded.

The cases were independently reviewed by six pathologists from the same academic institution. Each pathologist had special interest in breast pathology with varying levels of subspecialty training and experience. The observers had between 1 and 22 years of practice experience with residency training from 4 different institutions. Four pathologists had completed breast fellowship training at 3 different institutions. Although all observers were employed at the same institution during the study period, two had been at the institution for less than 1 year. All invasive cancer was graded using the Nottingham grading system. The tumors were assigned combined histologic grades (grade I, well differentiated; grade II, moderately differentiated; and grade III, poorly differentiated) based on established scoring for tubule formation, nuclear pleomorphism and calibrated mitotic count. Prior to the first round of grading, pathologists were provided with grading parameters as outlined by Elston and Ellis and adapted in the WHO Classification of Tumors of the Breast [7].

Each pathologist graded the tumors, blinded to patient and clinical information, and recorded the scores on a Microsoft Excel spreadsheet. Individual scores for tubule formation, nuclear grade, and mitotic count as well as the total score and degree of differentiation were recorded. The cases were graded a second time by the same 6 observers following a minimum 4-week washout period. For the second round of grading, the scoresheet provided the hormone receptor status and HER2 status for each case.

Inter- and intra-observer agreement was calculated using Cohen's κ analysis from Vassarstats.net. Cohen's κ was calculated by comparing grade variation between observer pairs and within individual observers for rounds 1 and 2 of grading. Fleiss' κ for overall agreement among all observers was calculated in Microsoft Excel. Overall interobserver variation between all 6 observers for tubule, nuclear, and mitotic scores was determined by Fleiss' κ for each grading component in both rounds 1 and round 2. The strength of agreement for κ was categorized as follows: <0.20: slight; 0.21–0.40: fair; 0.41–0.60: moderate; 0.61–0.80: good; >0.8: very good [10].

3. Results

3.1. Case details

A total of 100 H&E stained sections from 100 invasive carcinomas were identified, representing material from 97 patients, 94 patients with unilateral and 3 with bilateral

disease. The pathologic details obtained from the original pathology reports are detailed in Table 1.

3.2. Round 1 interobserver agreement

In the first round of grading, observers graded the majority of tumors as moderately differentiated (MD), in 49% to 58% of cases; poorly differentiated (PD) in 31–35% and well differentiated (WD) in 9% to 20% (Fig. 1). There was considerable interobserver agreement among the 6 pathologists. For observer pairs, concordance ranged from κ of 0.594 to 0.845 (moderate to very good) (mean = 0.721, Table 2). There was an overall Fleiss' κ of 0.675 (good agreement) with complete agreement across all 6 observers in 54 cases. In the 46 cases where exact concordance was not obtained, 25 resulted in the assignment of a different PPS, most commonly prognostic stage IA to IB (n = 18). This was followed by a change of stage IB to IIA (n = 5), IB to IIB (n = 1), and stage IIIB to IIIC (n = 1) (Table 3). In 44 of the cases lacking complete concordance across all observers, there was only a single step difference in grade (i.e. WD vs MD or MD vs PD). Only 2 cases were given discordant grades ranging from well to poorly differentiated (Fig. 2). Case 63 and 79 were both tumors that showed significant morphologic heterogeneity, resulting in intra- as well as inter-observer variability. Case 63 was graded as MD by 4 observers but WD by 1 observer and PD by the final observer (Fig. 2A-C). This case showed foci of tubule formation and was assigned tubule scores of 2 and 3. Nuclear atypia was evident, receiving a nuclear score of 3 from all but 1 observer. Mitotic activity was low, with a score of 1 assigned by all but 1 observer. The second discordant case, number 79, demonstrated a 1/3/2 split with 1 WD, 3 MD and 2 PD assignments given to this tumor (Fig. 2D-F). In this case, the interpretation of tubule formation in a tumor with a cribriform growth pattern resulted in varying tubule scores of 2 and 3. Dramatic heterogeneity in the proliferative activity within this single section also resulted in mitotic scores of 1, 2, and 3.

Overall Fleiss' κ was calculated for the individual components of the histologic grade (Table 4). There was only moderate agreement for each variable with a κ range of 0.461 to 0.588 in round 1. Interobserver agreement was best for the tubule score and worst for nuclear score.

Table 1 Pathologic characteristics of cases utilized for tumor grading

Characteristic	n
Mean tumor size (range)	4.5 cm (0.95-11)
Tumor grade	
I (well differentiated)	20
II (moderately differentiated)	55
III (poorly differentiated)	25
Histologic subtype	
Invasive ductal	55
Invasive lobular or ductal with lobular features	36
Other special type	9
Tumor ER/PR/HER2 status	
ER+/PR+/HER2-	77
ER+/PR-/HER2-	8
ER+/PR+/HER2+	3
ER-/PR-/HER2-	10
ER-/PR-/HER2 +	2
Lymph node status	
N0	56
N1mi	6
N1a	35
Unknown	3

3.3. Round 2 interobserver agreement

The majority of tumors were again graded as moderately differentiated (Fig. 1). In this round, complete agreement across all 6 observers was achieved in 58 cases. Observer pair agreement ranged from κ of 0.582 to 0.857 (moderate to very good) (mean = 0.723; Table 2). The overall Fleiss' κ was 0.68 (good agreement). Of the 42 cases with discordant grades, 29 resulted in the assignment of a different PPS, most commonly prognostic stage IA to IB (n = 21) (Table 3). Overall Fleiss' κ calculated for the individual components of the histologic grade, showed only moderate agreement among observers for each variable (Table 4). Interobserver agreement was again best for the tubule score, but worst for mitotic count in this round.

Cases 63 and 79 remained problematic with case 63 again showing a 1/4/1 split between WD/MD/PD. Although the same observer designated the case as PD, the WD designation was assigned by a different observer than in the first round. In case 79, the previous 1/3/2 WD/MD/PD split was now 1/2/3 with 4 observers changing the grade of this tumor between rounds 1 and 2. One observer changed their grading assignment 2 steps from WD in round 1 to PD in round 2. In this second grading round, 2 additional cases, cases 57 and 99, resulted in a 2-step grade discordance between observers with a 1/4/1 split of WD/MD/PD assignments seen in each case.

3.4. Intra-observer agreement

Intra-observer agreement was good to very good with a κ range of 0.708 to 0.833 (mean = 0.766) (Table 5). A 2-step intra-observer grading disparity (WD to PD) occurred in only one case with a single observer. Individual observers changed the tumor grade between rounds 1 and 2 in 11–20 cases (mean = 15.8). This intra-observer difference in tumor grade resulted in a change of PPS in 4–14 cases (mean = 9.6).

Reviewing individual components of the grading score revealed that intra-observer consistency was best for tubule formation with a κ of 0.633 to 0.860 (mean = 0.781), followed by mitoses (κ range 0.575 to 0.811; mean = 0.700), and nuclear grade (κ range 0.499 to 0.767; mean = 0.617). A score

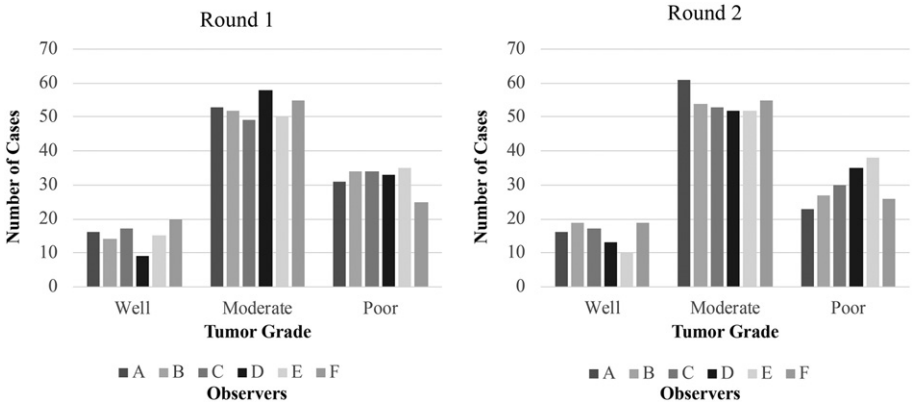

Fig. 1 Distribution of tumor grade assigned by each observer (A-F) in rounds 1 and 2 of independent grading.

of 1 for tubules was assigned infrequently by individual observers (in 3–12 cases; mean = 7). Most cases were given a tubule score of 3 (in 69–79 cases; mean = 75). A nuclear score of 1 was rarely given (range 0–3 cases; mean = 1) with no observer assigning a nuclear score of 1 in both rounds for a single case. The most common scoring change, between rounds 1 and 2, was between nuclear scores 2 and 3. The most common score for mitoses was 1 (range 29–59 cases; mean 53). Two-step scoring changes within a single observer occurred most frequently with the mitotic score (9 cases).

4. Discussion

Histologic tumor grade in breast carcinoma has long been shown to be associated with patient outcome. As a result of the many studies linking tumor grade and patient outcomes in breast cancer, the 8th edition of the AJCC Cancer Staging Manual incorporated tumor grade in prognostic stage groups. Many studies have examined the consistency of tumor grading with variable results. Gilchrist et al found that the level of agreement decreased as the number of observers increased, arguing that the level of agreement on histologic features in stage II breast carcinomas did not reach a clinically acceptable level for prognostic use [8]. Agreement levels have improved with refinement to the grading system. Elston and Ellis modified the Bloom and Richardson grading method to improve its accuracy, further refining the system to address tissue preparation and give explicit details for assessing the tubule, nuclear and mitotic scoring components [6,7]. The study showed that reproducibility is good, achieving >90% agreement between 2 independent pathologists, if the grading protocol is followed consistently. Subsequent studies have confirmed that this modified grading system remains a powerful, independent prognostic indicator [11-14]. Published studies using the Nottingham modification of the Bloom and Richardson system have shown levels of interobserver agreement ranging from fair to moderate (κ = 0.5-0.7) [15-18], similar to concordance rates found in our current study.

We found more consistent grading within individual observers than when comparing the grading of different observers, suggesting the possibility of personal biases that may have been taught in early training or acquired with practice experience. One tumor feature might influence our opinion of another scoring component, such as the formation of tubular architecture, often the most readily identified feature when grading tumors, causing an observer to lower the score

Table 2 Pairwise κs for interobserver variability

	Pathologist B	Pathologist C	Pathologist D	Pathologist E	Pathologist F
			Round 1		
Pathologist A	0.726	0.829	0.772	0.645	0.743
Pathologist B		0.845	0.754	0.629	0.730
Pathologist C			0.778	0.709	0.734
Pathologist D				0.607	0.716
Pathologist E					0.594
			Round 2		
Pathologist A	0.805	0.850	0.745	0.593	0.818
Pathologist B		0.815	0.715	0.627	0.756
Pathologist C			0.727	0.666	0.857
Pathologist D				0.611	0.671
Pathologist E					0.582

Table 3 Prognostic stage of cases with discordant tumor grades

Prognostic Stage of Cases with Discordant Tumor Grades							
Round 1							
	I A	I B	II A	II B	III A	III B	III C
I A	16	18[a]					
I B			5[a]	1[a]			
II A			5				
II B				1			
III A							
III B							1[a]
Round 2							
	I A	I B	II A	II B	III A	III B	IIIC
I A	9	21[a]	1[a]				
I B			6[a]	1[a]			
II A			2				
II B				1			
III A							
III B						1	

[a] The change in tumor grade resulted in a change in prognostic stage.

for nuclear or mitotic scores; however, a prior study has shown that tubule formation in breast tumors does not influence nuclear scores [19]. While tubule score was the most reproducible individual scoring element in our study, two of the most discrepant cases reflect clear differences in definition of tubule formation as it relates to grading. Some observers do not include the complex gland formation of cribriform architecture when quantifying tubule formation while others do include this, leading to wide grading differences in tumors with this growth pattern, like case 79. Differences in interpretation of this pattern have been previously reported as a source of interobserver variability [20]. Improved definitions regarding the inclusion or exclusion of cribriform spaces could lead to improved grading agreement. Additionally, in the current study, observers avoided nuclear scores of 1, only assigning this score in up to 3 cases. One observer never assigned a nuclear score of 1. Despite this tendency to give only scores of 2 and 3, nuclear pleomorphism is reportedly the least reproducible component of the grading system [16], as it was in the first round of this grading study. Dalton et al focused specifically on nuclear scoring differences between observers and found that the level of discordance in assigned nuclear scores was a predictive tool for patient outcome [21]. Those cases with discordant nuclear grades represented "in-between" tumors with a more favorable prognosis than those in which there was high concordance of high nuclear grade. In our study, we found high inter- and intra-observer variation in mitotic scores. As the seemingly most objective component of tumor grading, various tools have been suggested to increase the reproducibility of mitotic scoring, including lowering the mitotic index or utilizing an inmmunohistochemical labeling index [22,23]. Improved definitions for scoring tumors may improve interobserver agreement and help individual pathologists to be aware of their biases.

Several studies have suggested consensus review as a method to improve concordance, especially for borderline cases, with reinforcement by frequent review of criteria [12,20,22]. This is difficult in daily practice and one study showed no improvement in concordance between "training" and "test" sets [20]. Dalton et al found high levels of grading agreement of 25 pathologists from 6 separate institutions when ensuring specified guidelines are utilized [24]. Nearly all observer pairs in our study showed at least "good" agreement ($\kappa = 0.61–0.80$), and up to four pairs reached the "very good" threshold of interobserver agreement ($\kappa > 0.80$). Without comparison to outcomes in these patients, however, this is a measure of reproducibility only, not accuracy of

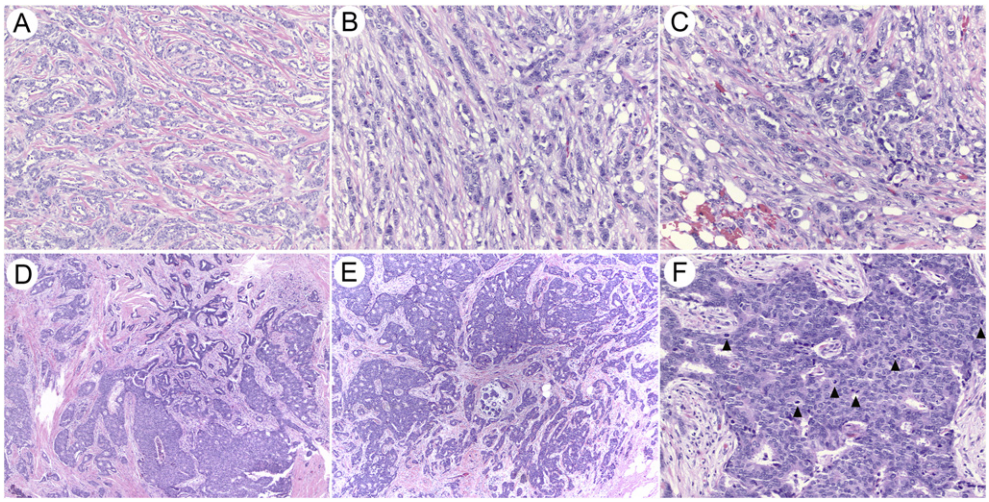

Fig. 2 Case 63 (A-C): A, Focal tubule formation (H&E, original magnification ×100). B, Admixed with areas of linear growth (H&E, original magnification ×200). C, Nuclear atypia was noted (H&E, original magnification ×200). Case 78 (D-F): D, An area of clear tubule formation (H&E, original magnification ×40). E, With predominantly cribriform architecture (H&E, original magnification ×40). F, Mitotic activity was heterogeneous throughout the tissue section but scattered foci showed high mitotic counts (mitotic figures highlighted by arrowheads (H&E, original magnification ×200).

Table 4 Interobserver variability of individual grading components

Component	Interobserver Fleiss' κ Round 1	Interobserver Fleiss' κ Round 2
Tubule formation	0.588	0.596
Nuclear pleomorphism	0.461	0.530
Mitotic count	0.550	0.521

outcome prediction. With the advent of digital microscopy, analysis of microscopic digital images has added another variable to tumor grading. Rakha et al compared grading by light microscopy to that of whole slide imaging, finding both methods to be comparable, with similar ability to predict outcome [25]. Potential next steps would be to utilize image analysis and deep learning to perform tumor grading for added reproducibility [26,27].

The notion of observer variability in tumor grading is not new, but understanding its impact on the new prognostic staging groups is important. In our study we found that intra-observer variability resulted in PPS changes in up to 14% of cases. As expected, comparing the grading variability across 6 observers resulted in even more frequent PPS differences (up to 29% of cases) and were most frequently between stages IA and IB. This PPS already represents down-staging from the anatomic stage of II, as a result of the ER+/HER2- status of most tumors. Most PPS variances were a result of observer grading discordance between MD and PD (19 of 25 cases). In a recent study of tumors with discordant grade assignment, these borderline tumors typically showed slightly worse outcomes, suggesting behaviors may be more in keeping with the higher assigned grade [28]. With the incorporation of costly multigene testing into the new PPS system, cases with an Oncotype score of <11 would receive a PPS of IA regardless of tumor grade. Although most cases in this study were ER+/Her2- tumors without axillary nodal metastases, most did not have an Oncotype DX score available to determine how this test result would influence the final PPS. Future work may determine if interobserver variability in grading is mitigated through the use of multigene testing, such as Oncotype DX.

Table 5 Intra-observer variability between round 1 and round 2 grading

	κ
Pathologist A	0.820
Pathologist B	0.744
Pathologist C	0.833
Pathologist D	0.721
Pathologist E	0.708
Pathologist F	0.770

In summary, we found that while there is grading variability between pathologists, agreement remained good across 6 different pathologists and very good within individual observers. Two- step grading discrepancies were rare and most differences in tumor grade resulted in only minor PPS changes within the stage I group. The increasing use of Oncotype DX may further reduce discordances due to tumor grading within the IA and IB PPS groups. This level of grading and staging discordance is therefore likely clinically acceptable, particularly when combined with tumor biomarker status.

Author contributions

KR wrote the paper; collected the data, contributed data; performed analysis.

OS, VB, MH, RC: contributed data.

ER wrote the paper; conceived and designed the analysis; collected data; contributed data; performed analysis.

References

[1] Amin MB, American Joint Committee on Cancer., American Cancer Society. AJCC cancer staging manual. Eight ed. New York: Springer; 2017. xvii, 1024 pages p.589–628.
[2] Greenough RB. Varying degrees of malignancy in cancer of the breast. J Cancer Res 1925;9:452-63.
[3] Bloom HJ. Prognosis in carcinoma of the breast. Br J Cancer 1950;4 (3):259-88.
[4] Bloom HJ. Further studies on prognosis of breast carcinoma. Br J Cancer 1950;4(4):347-67.
[5] Bloom HJ, Richardson WW. Histological grading and prognosis in breast cancer; a study of 1409 cases of which 359 have been followed for 15 years. Br J Cancer 1957;11(3):359-77.
[6] Elston CW. The assessment of histological differentiation in breast cancer. Aust N Z J Surg 1984;54(1):11-5.
[7] Elston CW, Ellis IO. Pathological prognostic factors in breast cancer. I. the value of histological grade in breast cancer: experience from a large study with long-term follow-up. Histopathology 1991;19(5): 403-10.
[8] Gilchrist KW, Kalish L, Gould VE, et al. Interobserver reproducibility of histopathological features in stage II breast cancer. An ECOG study Breast Cancer Res Treat 1985;5(1):3-10.
[9] Edge SB, American Joint Committee on Cancer. AJCC Cancer Staging Manual. 7th ed. New York: Springer; 2010 xiv, 648 p.363.
[10] Fleiss JL. Statistical methods for rates and proportions. 2nd ed. New York: Wiley; 1981 xviii, 321 p. 230.
[11] Carriaga MT, Henson DE. The histologic grading of cancer. Cancer 1995;75(1 Suppl):406-21.
[12] Dalton LW, Pinder SE, Elston CE, et al. Histologic grading of breast cancer: linkage of patient outcome with level of pathologist agreement. Mod Pathol 2000;13(7):730-5.
[13] Page DL, Gray R, Allred DC, et al. Prediction of node-negative breast cancer outcome by histologic grading and S-phase analysis by flow cytometry: an eastern cooperative oncology group study (2192). Am J Clin Oncol 2001;24(1):10-8.
[14] Rakha EA, El-Sayed ME, Lee AH, et al. Prognostic significance of Nottingham histologic grade in invasive breast carcinoma. J Clin Oncol 2008;26(19):3153-8.

[15] Boiesen P, Bendahl PO, Anagnostaki L, et al. Histologic grading in breast cancer—reproducibility between seven pathologic departments. South Sweden Breast Cancer Group Acta Oncol 2000;39(1):41-5.
[16] Frierson Jr HF, Wolber RA, Berean KW, et al. Interobserver reproducibility of the Nottingham modification of the Bloom and Richardson histologic grading scheme for infiltrating ductal carcinoma. Am J Clin Pathol 1995;103(2):195-8.
[17] Liyanage T. Effect of experience of pathologist on the inter-observer reproducibility of the Nottingham grade. Galle Medical Journal 2011;16(1):6-10.
[18] Longacre TA, Ennis M, Quenneville LA, et al. Interobserver agreement and reproducibility in classification of invasive breast carcinoma: an NCI breast cancer family registry study. Mod Pathol 2006; 19(2):195-207.
[19] Mora B, Bombari D, Schaefer SC, et al. Tumor architecture exerts no bias on nuclear grading in breast cancer diagnosis. Virchows Arch 2012;461(4):399-403.
[20] Robbins P, Pinder S, de Klerk N, et al. Histological grading of breast carcinomas: a study of interobserver agreement. Hum Pathol 1995;26(8):873-9.
[21] Dalton LW, Gerds TA. The advantage of discordance: an example using the highly subjective nuclear grading of breast cancer. Am J Surg Pathol 2017;41(8):1105-11.
[22] Meyer JS, Alvarez C, Milikowski C, et al. Breast carcinoma malignancy grading by Bloom-Richardson system vs proliferation index: reproducibility of grade and advantages of proliferation index. Mod Pathol 2005;18(8):1067-78.
[23] O'Shea AM, Rakha EA, Hodi Z, Ellis IO, Lee AH. Histological grade of invasive carcinoma of the breast assessed on needle core biopsy—modifications to mitotic count assessment to improve agreement with surgical specimens. Histopathology 2011;59(3):543-8.
[24] Dalton LW, Page DL, Dupont WD. Histologic grading of breast carcinoma. A reproducibility study. Cancer 1994;73(11):2765-70.
[25] Rakha EA, Aleskandarani M, Toss MS, et al. Breast cancer histologic grading using digital microscopy: concordance and outcome association. J Clin Pathol 2018;71(8):680-6.
[26] Dimitropoulos K, Barmpoutis P, Zioga C, Kamas A, Patsiaoura K, Grammalidis N. Grading of invasive breast carcinoma through Grassmannian VLAD encoding. PLoS One 2017;12(9):e0185110.
[27] Spanhol FA, Oliveira LS, Petitjean C, Heutte L. A dataset for breast cancer histopathological image classification. IEEE Trans Biomed Eng 2016;63(7):1455-62.
[28] Rakha EA, Aleskandarany MA, Toss MS, et al. Impact of breast cancer grade discordance on prediction of outcome. Histopathology 2018; 73(6):904-15.