IJC
International Journal of Cancer

Significant inter- and intra-laboratory variation in grading of invasive breast cancer: A nationwide study of 33,043 patients in the Netherlands

Carmen van Dooijeweert[1], Paul J. van Diest[1], Stefan M. Willems[1], Chantal C. H. J. Kuijpers[1], Elsken van der Wall[2], Lucy I. H. Overbeek[3] and Ivette A. G. Deckers[3]

[1]Department of Pathology, University Medical Center Utrecht, Utrecht, The Netherlands
[2]Department of Medical Oncology, University Medical Center Utrecht, Utrecht, The Netherlands
[3]Foundation PALGA (the nationwide network and registry of histo- and cytopathology in The Netherlands), Houten, The Netherlands

Accurate, consistent and reproducible grading by pathologists is of key-importance for identification of individual patients with invasive breast cancer (IBC) that will or will not benefit from adjuvant systemic treatment. We studied the laboratory-specific grading variation using nationwide real-life data to create insight and awareness in grading variation. Synoptic pathology reports of all IBC resection-specimens, obtained between 2013 and 2016, were retrieved from the nationwide Dutch Pathology Registry (PALGA). Absolute differences in laboratory-proportions of Grades I–III were compared to the national reference. Multivariable logistic regression provided laboratory-specific odds ratios (ORs) for high- *vs.* low-grade IBC. 33,792 IBC pathology reports of 33,043 patients from 39 laboratories were included, of which 28.1% were reported as Grade I (range between laboratories 16.3–43.3%), 47.6% as Grade II (38.4–57.8%), and 24.3% as Grade III (15.5–34.3%). Based on national guidelines, the indication for adjuvant chemotherapy was dependent on histologic grade in 29.9% of patients. After case-mix correction, 20 laboratories (51.3%) showed a significantly deviant OR. Significant grading differences were also observed among pathologists within laboratories. In this cohort of 33,043 breast cancer patients, we observed substantial inter- and intra-laboratory variation in histologic grading. It can be anticipated that this has influenced outcome including exposure to unnecessary toxicity, since choice of adjuvant chemotherapy was dependent on grade in nearly a third of patients. Better standardization and training seems warranted.

Introduction

About one in seven women in the Netherlands will develop breast cancer during her life,[1] which makes breast cancer the most common type of cancer in Dutch women with approximately 15,000 new diagnoses per year.[2] Histologic grade is one of the best established prognostic factors in breast cancer and is strongly and independently associated with both breast cancer-specific and disease-free survival.[3,4] Studies even suggest that histologic grade can predict tumor behavior more accurately than other "time-dependent" prognostic factors like tumor size.[3–7] Hence, histologic grade is an important clinical contributor and is widely used to guide therapeutic breast cancer management.[3,4,8] Furthermore, since breast cancer screening programs resulted in earlier detection and thereby a greater proportion of both smaller tumors and

Key words: invasive breast cancer, prognostic factor, patient management, histologic grade, pathology

Abbreviations: CI: confidence interval; ER: estrogen; GDPR: General Data Protection Regulation Act; HER2: Her2neu/ErbB-2/ERBB2; IBC: invasive breast cancer; IHC: immunohistochemistry; ISH: in situ hybridization; OR: odds ratio; PALGA: nationwide Dutch Pathology Registry; PR: progesterone; SD: standard deviation

Additional Supporting Information may be found in the online version of this article.

Conflict of interests: SW is a board member of the medical advisory board of Roche, Pfizer, AstraZeneca, MSD and Cergentis and receives unrestricted research grants from Roche, Pfizer, AstraZeneca and MSD. For the remaining authors, no conflict of interest was declared.

Grant sponsor: Quality Foundation of the Dutch Association of Medical Specialists (SKMS)
DOI: 10.1002/ijc.32330

History: Received 22 Mar 2019; Accepted 28 Mar 2019; Online 11 Apr 2019.

Correspondence to: Paul J van Diest, MD, PhD, Professor and head, Department of Pathology, University Medical Center Utrecht, PO Box 85500, 3508 GA Utrecht, The Netherlands, Tel.: +31-88-755-6565, Fax: +31-88-756-9593, E-mail: p.j.vandiest@umcutrecht.nl

What's new?

Histologic grade serves a critical prognostic role in invasive breast cancer (IBC) and is used to guide therapeutic decisions. Evidence indicates, however, that IBC grading varies considerably. Here, grading variation in clinical practice was evaluated using real-life data from laboratories in the nationwide Dutch Pathology Registry. Laboratories varied in IBC grade I, II, and III reporting. Among grading components, nuclear polymorphism showed the greatest difference between laboratories. Within laboratories, one-third of pathologists deviated significantly from national proportions for IBC grade I. Despite deployment of uniform guidelines across laboratories, IBC histologic grading is not necessarily performed in a consistent manner.

lymph node negative tumours,[9–11] histologic grade is determinative in patient management in a substantial number of cases, including the use of genetic profiling tests.[8,12–15] Additionally, in contrast to prognostic genetic profiling tests, the evaluation of histologic grade is cheap and can in principle be performed in all cases of breast cancer.[16]

The most widely used grading system for invasive breast cancer (IBC) is the modified Bloom and Richardson guideline (Elston–Ellis modification of Scarff–Bloom–Richardson grading system, also known as the Nottingham grading system),[17,18] which combines the assessment of cell morphology (nuclear polymorphism), measurement of differentiation (tubule formation) and assessment of proliferation (mitotic count), resulting in a total score and derived grade.[18] This system is considered suitable for evaluating IBC in routine clinical setting and is globally incorporated in breast cancer guidelines.[3] Furthermore, Lundin *et al.* concluded that even when assessed by pathologists who have had no special training in breast cancer pathology, histologic grading in breast cancer is of substantial and independent prognostic value.[19]

While grading has systematically been proven to be prognostically very important, accurate, consistent, and reproducible grading by pathologists is of key-importance for identification of individual patients who, based on their prognosis, may or may not benefit from adjuvant treatment. However, current evidence suggests that there is considerable and clinically relevant variation in the grading of IBC. Previous studies, in which a set of IBC was reviewed by several pathologists, mostly showed an overall reproducibility that was no more than moderate.[16,20–22] Yet, these conclusions are derived from smaller studies where grading was performed in study setting, and thus this may not resemble real-life grading in daily clinical practice. Moreover, individual practicing pathologists may not have felt addressed by these data, as it did not provide them insight into their own grading practice.

In this context, nationwide daily clinical practice studies did show that there is considerable variation between laboratories and individual pathologists in the grading of, for example, colorectal adenomas[23] and colorectal adenocarcinomas.[24] In addition, we previously reported substantial nationwide interlaboratory and intralaboratory variation in grading of ductal carcinoma *in situ* of the breast (DCIS).[25] Grading of DCIS is methodologically different from and less standardized than grading of IBC, and DCIS treatment is currently independent of any histopathologic features, whereas treatment of IBC is widely guided by histologic grading. Therefore, it is unclear whether the same conclusions can be drawn for IBC. In light of its current, and important, clinical consequences, it is also particularly important to create insight and awareness in grading variation of IBC.

We studied the laboratory-specific variation in histologic grading of IBC in a nationwide study in the Netherlands. Using the Dutch nationwide pathology registry (PALGA), we assessed the variation in histologic grading of over 33,000 patients with IBC, between Dutch pathology laboratories and between individual pathologists using real-life data from synoptic (structured) pathology reports from daily pathology practice. In addition, we also analyzed the variation of the three components of grading (according to the modified Bloom and Richardson classification) between laboratories. Furthermore, we conducted a questionnaire among pathologists to gain insight into their grading practices. As grade is an important decision-tool in adjuvant treatment, inter- and intra-laboratory variation in grading may lead to under- and over-treatment of a substantial percentage of primary breast cancer patients. Creating insight into these laboratory-specific differences may help to design an intervention to improve standardization among laboratories and pathologists.

Materials and Methods

Data source and study population

Data were extracted from PALGA, the nationwide network and registry of histopathology and cytopathology in the Netherlands, which contains excerpts of all pathology reports from Dutch Pathology laboratories since 1991.[26] All data from the PALGA database are pseudonymized by a trusted third party (ZorgTTP, Houten, The Netherlands). Consent was given by all Dutch laboratories for the storage of their data by PALGA, and for scientific use of these data. Pathology laboratories were initially anonymized and further consent was obtained for additional analysis of inter-pathologist variation within the individual laboratories ($n = 8$ laboratories). The scientific and privacy committee of PALGA approved this study. All data were retrieved and handled in compliance with the General Data Protection Regulation Act (GDPR).

We retrieved all synoptic pathology reports of patients with resection specimens of IBC between January 1, 2013, and December 31, 2016, in the Netherlands (n = 48,667). Synchronous IBC was defined as an ipsilateral lesion within 6 months of the first IBC resection. These lesions were considered paired

measurements of which only the first was included. We solely included patients with primary tumors, thereby excluding resection specimens with complete regression of the tumor, specimens without a tumor after biopsy, and specimens of reexcisions. As neoadjuvant therapy may influence grading,[27–29] pathology reports of patients who received neoadjuvant treatment were excluded (Supporting Information Fig. S1).

In total, 40 out of 46 Dutch laboratories synoptically reported IBC on breast resection specimens. Of these, we included those that synoptically reported ≥250 IBC during the study period. For interpathologist variation within individual laboratories, we only analyzed data from pathologists from the eight participating laboratories who synoptically reported ≥20 IBC during the study period.

For each patient, we extracted sex and age, type of surgery, IBC tumor size, histologic subtype, histologic grade, ER/PR-receptor status (immunohistochemistry [IHC]) and HER2-receptor status (IHC and/or *in situ* hybridization [ISH]). Reports of IBC with any missing data were excluded from further data analysis (Supporting Information Fig. S1).

Analysis of histologic grading

In the PALGA synoptic reporting module, histologic grade was determined according to the modified Bloom and Richardson guideline (Elston–Ellis modification of Scarff–Bloom–Richardson grading system, also known as the Nottingham grading system). According to this guideline, the three components (tubule formation, nuclear polymorphism and mitotic count) are scored from 1 to 3, which results in a total score and derived overall histologic grade (score 3–5 = Grade I, score 6–7 = Grade II, score 8–9 = Grade III).[17,18] The primary outcome measure of this study was the interlaboratory variation in histologic grading of IBC and separate for its three components. Secondary outcome measure was the interpathologist variation in histologic grading within a single laboratory.

Histologic grading in relation to clinical management

To gain insight into the influence of histologic grading on therapeutic patient management, we identified a subgroup of patients who, in view of current national guidelines,[8] were eligible for adjuvant systemic chemotherapy solely on the basis of histologic grade. This concerned patients ≥35 years of age with a negative HER2-receptor status and a tumor size of 1.1–2 cm, or, in patients <35 years, those with a negative HER2-receptor status and a tumor of ≤1 cm, or a positive HER2-receptor status and a tumor of <0.5 cm. In these patients, Grades II–III tumors qualify for adjuvant chemotherapy, whereas this is not recommended for patients with Grade I tumors. In addition, for this group, it was checked whether the total (modified Bloom and Richardson) score was on a switch point of grades, (i.e., scores 5 [Grade I] or 6 [Grade II] and scores 7 [Grade II] or 8 [Grade III]), where the difference of only one point on the total Bloom and Richardson score could already alter the overall histologic grade and thereby chemotherapy indication.

Questionnaire survey among pathologists

A questionnaire survey was sent to all 46 pathology laboratories in the Netherlands to identify how pathologists determine the histologic grade of IBC in daily clinical practice. The survey contained questions on whether pathologists consider themselves specialized breast pathologists, the number of years of experience as a pathologist, how they count mitoses and how they deal with heterogeneity of histologic grade within one specimen (Supporting Information Fig. S2).

Statistical analysis

Patient and tumor characteristics were summarized and differences between histologic grades were tested by means of a χ^2-test for categorical variables and by a nonparametric Kruskal–Wallis test for continuous variables.

The overall proportions of histologic Grades I, II and III were determined and considered the national proportion. Absolute differences in proportion of histologic grades between laboratories are presented in funnel plots per grade, in which the proportions per laboratory are plotted against the number of included IBC per laboratory (Fig. 1). The target of these funnel plots was set at the national proportions with their 95% confidence intervals (CI) as limits.[30] Absolute interlaboratory differences in proportions of the three components of grading were also analyzed.

To compare relative differences among laboratories, odds ratios (OR) and 95% CIs per laboratory were calculated by logistic regression. As there is no clear binary cut-off for low-grade and high-grade IBC in clinical practice, we performed two logistic regression analyses, with different definitions of low- and high-grade IBC. In our first logistic regression analysis, we defined low-grade IBC as Grade I and high-grade IBC as Grades II–III. In our second logistic regression analysis, we defined low-grade IBC as Grades I–II and high-grade IBC as Grade III. Both analyses resulted in ORs and 95% CI for high- *vs.* low-grade IBC.

For the choice of the reference laboratory of the logistic regression models, the sum of absolute deviations from the grade specific national proportions was calculated to compare the absolute deviation for all three grades at once. The laboratory with the lowest sum-score was deemed best resembling the national distribution and was thereupon chosen as reference laboratory.

Two multivariable logistic regression analyses for high- *vs.* low-grade IBC were performed to correct for differences in case-mix. To identify potential confounding factors, we selected clinicopathological variables *a priori* based on literature[18,31–35] and on pathologists' experience. These factors included age, sex, tumor size, type of surgery, histologic subtype, hormone-receptor status and HER2-receptor status. Hormone-receptor status (ER/PR) was considered positive when either or both the ER- or PR-receptor were positive. According to the Dutch guideline, and within the synoptic PALGA protocol module, the ER- and PR-receptor status is considered positive when ≥10% of the tumor cells show ER- and PR-specific staining on IHC.[8] Overall, hormone-receptor status was taken into account as a binary variable (either positive (≥10%) or negative (<10%)) and not as a continuous variable

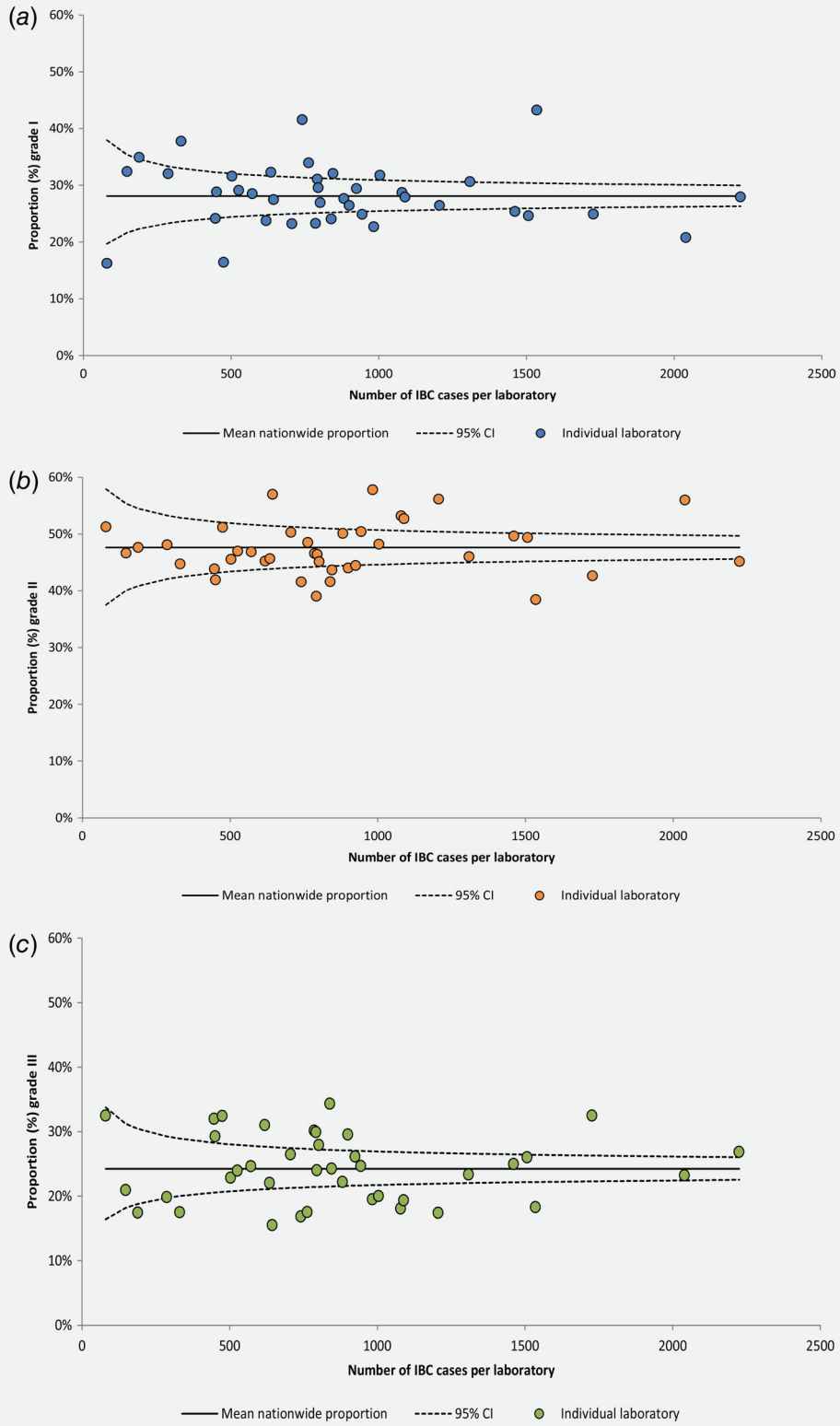

Figure 1. Funnel plots showing the observed proportion per IBC grade per laboratory (dots) relative to the mean national proportion and its 95% confidence intervals for IBC Grades I (*a*), II (*b*) and III (*c*) (2013–2016). [Color figure can be viewed at wileyonlinelibrary.com]

(percentage of stained tumor nuclei). All variables, except for sex, as the number of males was too low, appeared to be significantly associated with grade and were therefore included in both final multivariate models. It was checked whether males clustered in specific laboratories, but this was not the case. The adjusted ORs (95% CI) are presented in a forest plot (Fig. 2).

For analysis of the interpathologist variation within the laboratories, we merely compared the proportions per histologic grade between pathologists by Fisher exact test (Monte Carlo option; Fig. 3, Supporting Information Figs. S3*A* and S3*B*). Results of the questionnaire were summarized by frequencies and percentages.

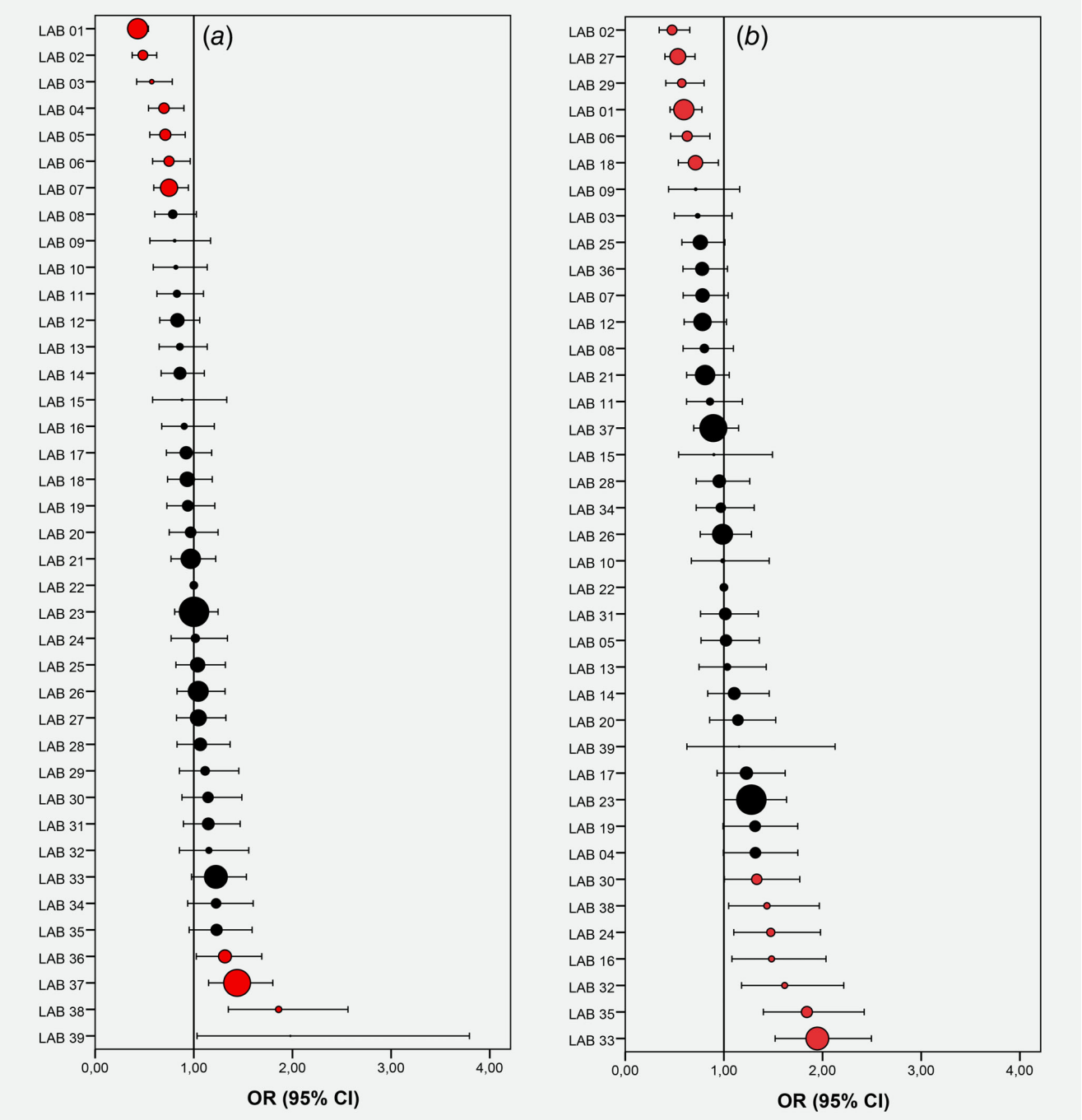

Figure 2. Forest plots showing the adjusted odds ratios (OR) and 95% confidence intervals (CI) of invasive breast cancer (IBC) Grades II–III vs. IBC Grade I (*a*) and of IBC Grade III versus IBC Grades I–II (*b*) in comparison to the reference laboratory (#22). Dot size indicates the total number of analyzed synoptically reported IBC lesions per laboratory. Red dots indicate laboratories with a significantly deviant OR as compared to the reference laboratory. ORs are adjusted for age, tumor size, type of surgery, histologic subtype, hormone receptor status and HER2 receptor status.

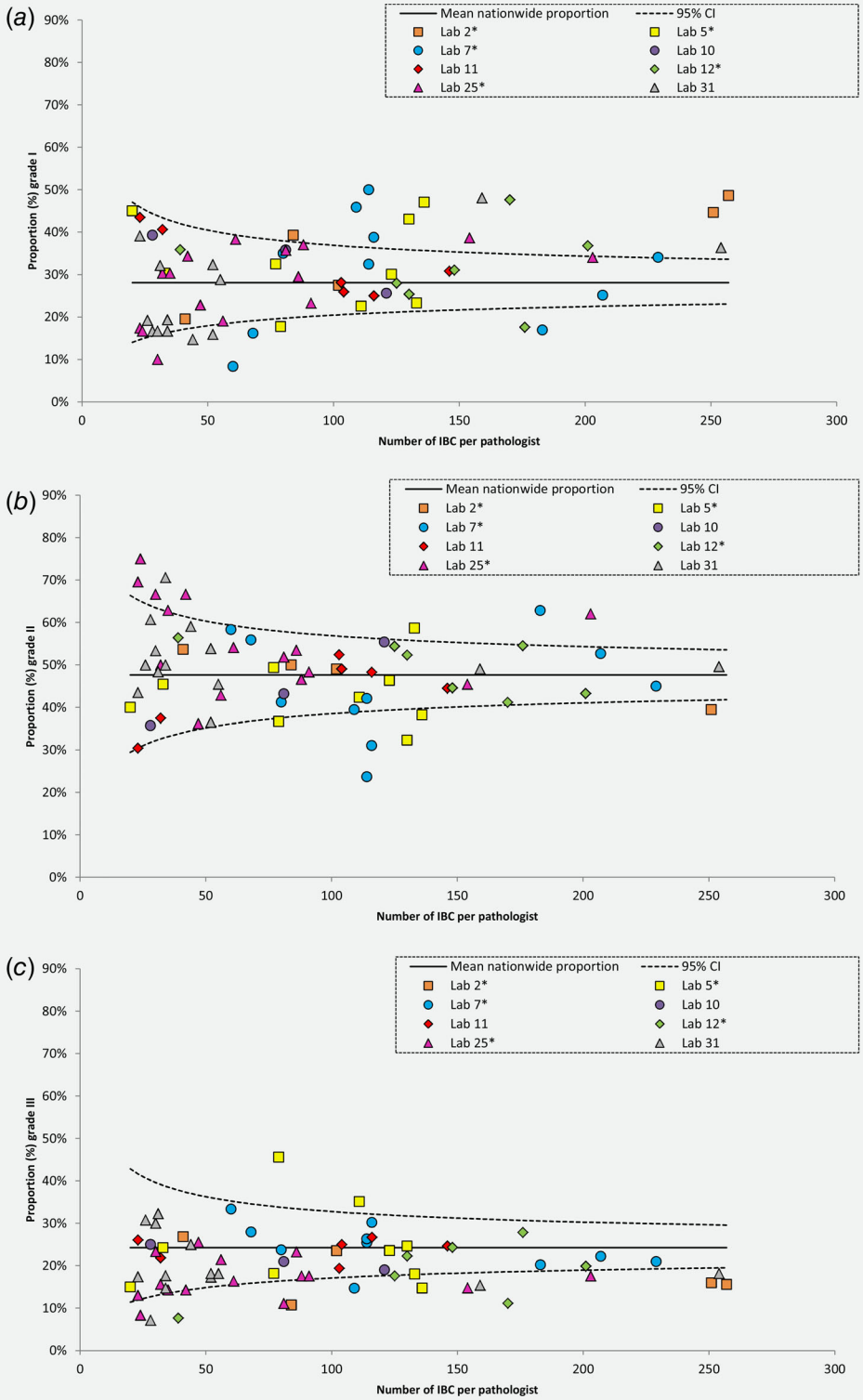

Figure 3. Funnel plots showing the observed proportion of invasive breast cancer (IBC) lesions per grade per pathologist (dots) of eight laboratories relative to the mean national proportion for IBC Grades I (*a*), II (*b*) and III (*c*) (2013–2016). *Indicates that the distribution of Grades I–III significantly differed between pathologists within the individual laboratory (calculated by Fishers Exact test; Monte Carlo option).

Values of $p < 0.05$ were considered statistically significant. All statistical analyses were performed by using IBM SPSS Statistics version 21.

Results

Characteristics of IBC lesions and laboratories

A total of 33,792 IBC lesions from 33,043 patients were included in our data analysis. For some patients, we included more than one pathology report as this concerned either a bilateral tumor or an ipsilateral tumor more than 6 months after the first IBC resection. All patients originate from a total of 39/46 Dutch pathology laboratories as one laboratory graded less than 250 IBC lesions within the synoptic PALGA protocol module and six laboratories had not yet implemented synoptic reporting at the time of the study (Supporting Information Fig. S1). Characteristics of these included patients and corresponding invasive breast tumors are listed in Table 1.

Mean (SD) age at diagnosis was 62.2 (12.1) years and patients were predominantly female (99.2%). Breast-conserving surgery was performed in the majority of patients (63.9%). Higher histologic grade was positively associated with mastectomy rate and tumor size, and with a negative ER/PR-receptor status and a positive HER2-receptor status.

The number of synoptically reported IBC lesions per laboratory ranged from 80 to 2,225 (median 795). Overall national proportions for IBC Grades I, II and III were 28.1, 47.6 and 24.3%.

Interlaboratory differences in histologic grading

Laboratories varied mostly in the reporting of IBC Grade I (16.3–43.3%), followed by IBC Grade II (38.4–57.8%) and IBC Grade III (15.5–34.3%). Overall, more than half of the laboratories (22/39) showed proportions outside the 95% CI for both Grade I and Grade III (56.4%), whereas this was the case for 41.0% of the laboratories for Grade II (Fig. 1).

The sum-score was lowest and only 1.6% for laboratory 22, which was thereupon chosen as reference laboratory. The maximum sum-score, in contrast, was 30.3% (Laboratory 1). Using the first definition of high-grade IBC (Grades II–III), multivariate logistic regression showed that 11 laboratories (28.2%) reported a significantly higher ($n = 4$) or lower ($n = 7$) proportion of high-grade IBC (Grades II–III) than the reference laboratory (Fig. 2*a*). Adjusted ORs of individual laboratories ranged from 0.43 (95% CI: 0.35–0.54) to 1.98 (95% CI: 1.03–3.79).

Using the second, alternative, definition of high-grade IBC (Grade III), multivariate logistic regression analyses showed that 13 laboratories (33.3%) reported a significantly higher ($n = 7$) or lower ($n = 6$) proportion of high-grade IBC (III) than the reference laboratory (Fig. 2*b*). Adjusted ORs of individual laboratories ranged from 0.47 (95% CI: 0.34–0.65) to 1.95 (95% CI: 1.52–2.50).

After correction for case-mix in both analyses, using different definitions of high-grade IBC, 20 laboratories (51.3%) had at least one significantly higher or lower OR than the reference laboratory. Four laboratories (10.3%) had significantly deviant ORs on both analyses (Fig. 2).

Interlaboratory differences in components of histologic grading

Regarding the three components of grading, most variation between laboratories was observed for nuclear polymorphism (broadest range in category 3 [severe] 14.2–55.0%), followed by mitotic count

Table 1. Characteristics of the 33,792 included invasive breast cancer (IBC) lesions from the PALGA database 2013–2016

	Total ($n = 33,792$)	Grade 1 ($n = 9,495$)	Grade 2 ($n = 16,098$)	Grade 3 ($n = 8,199$)	*p*
Age (years)[1]	62.2 (12.1)	62.4 (10.8)	62.8 (11.8)	60.7 (13.8)	0.000
Sex, *n* (%)					
Female	33,537 (99.2%)	9,441 (99.4%)	15,967 (99.2%)	8,129 (99.1%)	0.045
Male	255 (0.8%)	54 (0.6%)	131 (0.8%)	70 (0.9%)	
Tumor size (cm)[1]	1.9 (1.3)	1.4 (0.9)	1.9 (1.4)	2.3 (1.5)	0.000
Type of surgery, *n* (%)					
Mastectomy	12,209 (36.1%)	2,548 (26.8%)	6,172 (38.3%)	3,489 (42.6%)	0.000
Breast conserving	21,583 (63.9%)	6,947 (73.2%)	9,926 (61.7%)	4,710 (57.4%)	
Histologic subtype, *n* (%)					
Ductal	28,547 (84.5%)	8,727 (91.9%)	12,382 (76.9%)	7,438 (90.7%)	0.000
Lobular	4,432 (13.1%)	647 (6.8%)	3,465 (21.5%)	320 (3.9%)	
Other	813 (2.4%)	121 (1.3%)	251 (1.6%)	441 (5.4%)	
ER/PR receptor status, *n* (%)					
Positive	29,576 (87.5%)	9,373 (98.7%)	15,162 (94.2%)	5,041 (61.5%)	0.000
Negative	4,216 (12.5%)	122 (1.3%)	936 (5.8%)	3,158 (38.5%)	
HER2-receptor status, *n* (%)					
Positive	3,340 (9.9%)	212 (2.2%)	1,335 (8.3%)	1,793 (21.9%)	0.000
Negative	30,452 (90.1%)	9,283 (97.8%)	14,763 (91.7%)	6,406 (78.1%)	

[1]Mean (SD).

(broadest range in category 1 (≥13 mitoses per 2 mm^2): 47.2–75.2%) and tubular formation (broadest range for category 3 (<10% of cells with tubular differentiation): 52.7–74.2%).

Overall, the majority of tumors (76.8%) had a total grading score on a switch point of grades, that is, scores 5 or 6 (49.4%) and scores 7 or 8 (27.4%; Table 2), for which the difference of only one point on the total Bloom and Richardson score could alter the overall histologic grade and thereby the indication for chemotherapy.

Intralaboratory differences in histologic grading

Sixty-eight pathologists from the eight participating laboratories synoptically reported ≥20 tumors during the study period. Per laboratory, the number of analyzed pathologists ranged from 3 to 15 (median 8). In addition, the number of analyzed IBC lesions per pathologist ranged from 20 to 257 (median 82.5). Overall, 22 pathologists (32.4%) graded significantly deviant compared to the national proportions for IBC Grade I, while this was the case for 16 pathologists (23.5%) for Grade II and for 14 pathologists (20.6%) for Grade III (Fig. 3).

Most variation between pathologists within the individual laboratories was observed within laboratory 7 for Grade I (range 8.3–50.0%) and Grade II (range 23.7–62.8%), whereas most variation for Grade III (range 14.7%–45.6%) was observed in laboratory 5. For five laboratories (62.5%) the distribution of histologic grade (i.e., the proportions of Grades I–III) significantly differed between pathologists within that laboratory (Supporting Information Figs. S3*A* and S3*B*).

Indication for adjuvant systemic chemotherapy

In 19,461 of the 33,792 IBC lesions (57.6%), the pathology reports held complete information on all relevant variables that in current clinical practice are used to establish the indication for adjuvant chemotherapy in primary breast cancer (i.e., lymph node status, HER2-status, age, tumor size and histologic grade).

Histologic grade determined the indication for adjuvant chemotherapy in 5,821 patients (29.9%; Fig. 4). Of this group, 1,801 tumors (30.9%) were reported as Grade I and thus, according to current guidelines,[8] would not have had an indication for adjuvant chemotherapy. In 4,020 tumors (69.1%), solely based on histologic grade, adjuvant chemotherapy has likely been advised, as they were reported as Grade II or III tumors. In total, of the tumors in which the indication for adjuvant chemotherapy was dependent on histologic grade (n = 5,821), 3,187 (54.8%) even had a total score on the switch point of Grades I and II (i.e., an overall score of five [Grade I] or six [Grade II]).

Table 2. Scores of the three components of the modified Bloom and Richardson classification and overall score for the 33,972 included invasive breast cancer (IBC) lesions from the PALGA database 2013–2016

Characteristics		Total (n = 33,792)	Grade I (n = 9,495)	Grade II (n = 16,098)	Grade III (n = 8,199)
Tubular differentiation, n (%)					
1	>75% of cells	3,895 (11.5%)	3,698 (38.9%)	197 (1.2%)	0 (0.0%)
2	10–75% of cells	8,724 (25.8%)	4,694 (49.4%)	3,371 (20.9%)	659 (8.0%)
3	<10% of cells	21,173 (62.7%)	1,103 (11.6%)	12,530 (77.8%)	7,540 (92.0%)
Nuclear polymorphism, n (%)					
1	Mild 1[1]	2,942 (8.7%)	2,818 (29.7%)	124 (0.8%)	0 (0.0%)
2	Moderate 2[2]	20,741 (61.4%)	6,545 (68.9%)	12,258 (76.1%)	1,938 (23.6%)
3	Severe 3[3]	10,109 (29.9%)	132 (1.4%)	3,716 (23.1%)	6,261 (76.4%)
Mitotic count, n (%)					
1	<7 per 2 mm^2	21,164 (62.6%)	9,273 (97.7%)	11,891 (73.9%)	0 (0.0%)
2	≥8 ≤12 per 2 mm^2	5,163 (15.3%)	213 (2.2%)	3,270 (20.3%)	1,680 (20.5%)
3	≥13 per 2 mm^2	7,465 (22.1%)	9 (0.1%)	937 (5.8%)	6,519 (79.5%)
Total score					
3	Grade I	1,127 (3,3%)	1,127 (11.9%)	–	–
4	Grade I	2,796 (8.3%)	2,796 (29.5%)	–	–
5	Grade I	5,572 (16.5%)	5,572 (58.7%)	–	–
6	Grade II	11,127 (32.9%)	–	11,127 (69.1%)	–
7	Grade II	4,971 (14.7%)	–	4,971 (30.9%)	–
8	Grade III	4,277 (12.7%)	–	–	4,277 (52.2%)
9	Grade III	3,922 (11.6%)	–	–	3,922 (47.8%)

[1]Nuclei are small with little increase in size in comparison with normal breast epithelial cells, regular outlines, uniform nuclear chromatin, little variation in size.
[2]Cells larger than normal with open vesicular nuclei, visible nucleoli, and moderate variability in both size and shape.
[3]Vesicular nuclei, often with prominent nucleoli, exhibiting marked variation in size and shape, occasionally with very large and bizarre forms.

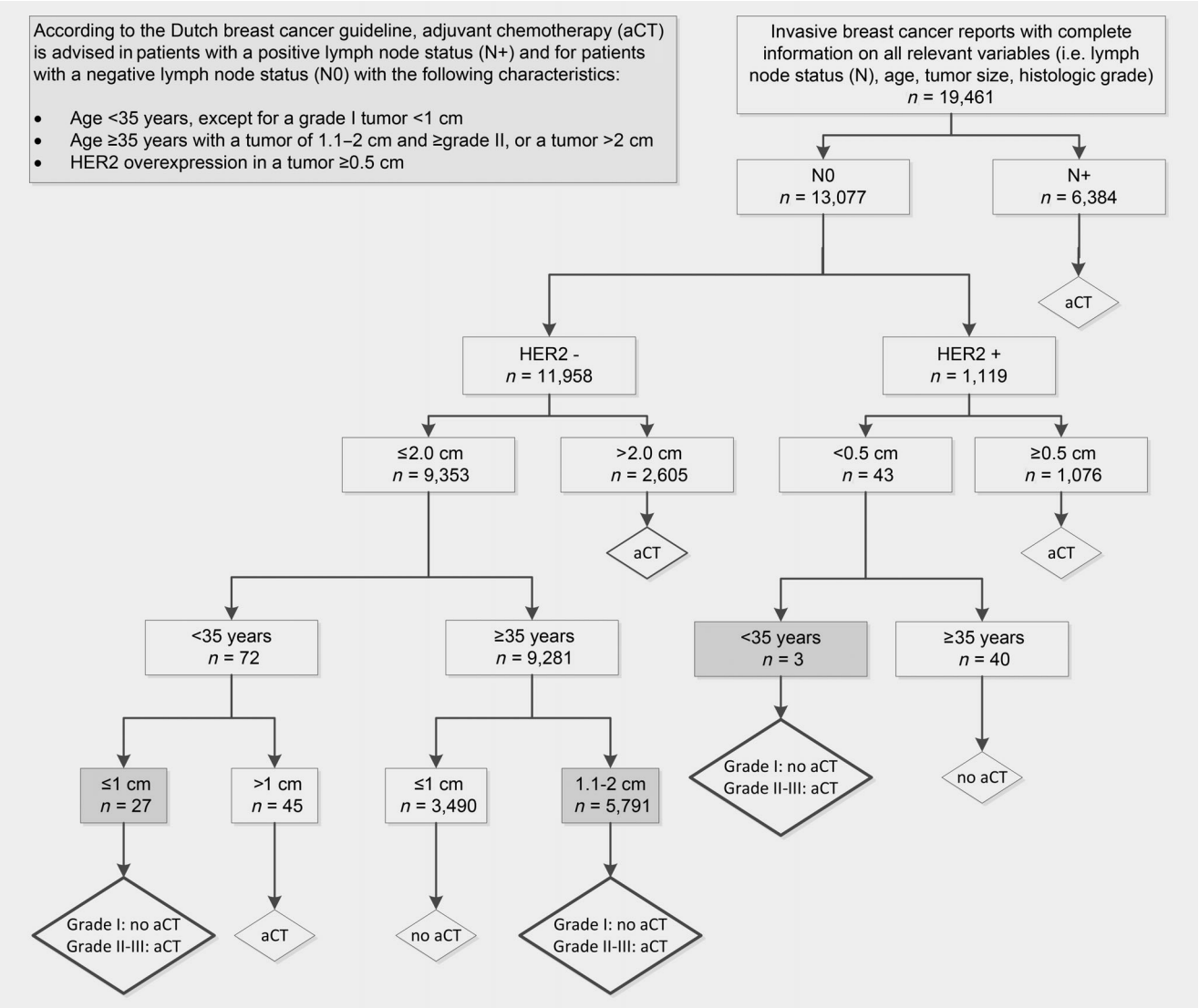

Figure 4. Flowchart showing the decision tree for adjuvant chemotherapy (aCT) in breast cancer patients according to the current Dutch guideline for the 19,461 tumors that held complete information on all relevant variables (i.e., lymph node status [N], age, tumor size and histologic grade). Grey squares indicate tumors of patients in which the aCT indication is dependent on histologic grade (n = 5,821, 29.9%).

Results of questionnaire survey

Seventy-nine pathologists out of the approximately 320 practicing pathologists in the Netherlands (25%),[24] responded to our online questionnaire, of which 19.0% worked in an academic hospital at the time. Thirty-seven (46.8%) pathologists denoted themselves as experts in breast pathology. Grading practice of generalized and specialized breast pathologists did not seem to differ (data not shown). All pathologists reported the modified Bloom and Richardson grading guideline as a reference for histologic grading of IBC, however, 11 pathologists (13.9%) also responded that, in their opinion, specific histologic subtypes per definition have a specific grade. In case of heterogeneity of histologic grade within one specimen, the majority of pathologists (76.0%) report the highest grade as overall histologic grade.

Discussion

In this nationwide cohort of 33,043 invasive breast cancer patients, approximately half of the lesions were reported as Grade II (47.6%), whereas Grade I and Grade III were scored in about a quarter of all lesions (28.1% and 24.3%, respectively). The observed overall proportions per grade are in line with previous cohort studies that showed a similar distribution pattern for IBC Grades I (15–30%), II (41–62%) and III (22–33%), although specific percentages vary.[36–41]

As synoptic reporting, compared to narrative reporting, results in an increased overall completeness of pathology reports,[42] and as it enables easy data extraction because all variables are stored in a standardized manner, data included in this study were solely from synoptically reported IBC lesions. Currently, over 80 percent of (pre)malignant breast lesions is reported via the synoptic

PALGA protocol by pathologists in the Netherlands.[43] As a control, we compared our data with aggregated data from the Netherlands Cancer Registry, which also holds narrative reports, and observed a similar distribution (data not shown), indicating that the distribution of histologic grade in our population, based on synoptic reporting, is likely to be representative for all IBC patients in the Netherlands.

Laboratory-specific data were analyzed in an absolute and relative manner, in which individual laboratories were compared to both the national proportion and a reference laboratory, all indicating that interlaboratory differences in histologic grading of IBC are substantial. This was illustrated by the large range of proportions per grade, by the sum score with variation of up to 30.3% and by the fact that the reported proportions per grade and the adjusted ORs were significantly deviant from the national distribution in approximately half of the laboratories.

In addition to the substantial interlaboratory variation, significant intralaboratory differences were also observed between pathologists within five of eight analyzed laboratories (62.5%). These findings emphasize that, even within the laboratories, histologic grading is not performed in a similar manner among pathologists, although the same guidelines are used in all laboratories and by all pathologists (modified Bloom and Richardson guideline). In the eight participating laboratories, 38 of 68 pathologists (55.9%) synoptically graded <100 IBC in the study period of 4 years. Although there is no external standard or benchmark to indicate whether a pathologist is an expert in IBC grading, this may imply that there are too many pathologists with too little experience in grading. However, in absolute manner, pathologists may grade more tumors than it seems from our results, as approximately 20% of IBC cases are still graded outside the synoptic protocol. Furthermore, our results show that both pathologists who grade few and pathologists who grade many IBC show significantly deviant proportions. Nevertheless, the results of this study raise the question of whether it is desirable that some pathologists may actually only grade a few IBC cases per year. This may be the subject of future research.

In line with previous studies,[16,20–22] most variation between laboratories was observed for nuclear polymorphism, which might be explained by the fact that scoring of this category is least quantitative, when compared to mitosis counting and to scoring the percentage of tubular differentiation. In addition, more than three-quarters of all patients had a tumor with a total score on a switch point of grades (i.e., scores 5 or 6, and scores 7 or 8), which shows that the variation of only one point in the total score of the three components may already change its subsequent histologic grade, and thus may influence patient management.

The results of this study hopefully raise awareness among pathologists and clinical oncologists, emphasizing that treatment decisions depend on histologic grade in a substantial number of patients and that, for individual patients, the difference of only one point on the total score could mean the difference between adjuvant chemotherapy or not. Therefore, accurate, consistent and reproducible grading is of utmost importance. However, this study also shows that histopathologic grading may currently not meet high enough clinical standards for individual patients, which is a crucial first step to improvement. Furthermore, pathologists are enabled to discuss and reflect on their grading practices as these "mirror" data were also sent to the laboratories by PALGA, which may lead to regression to the mean. In addition, these data should not only be discussed by pathologists but also in multidisciplinary meetings with clinical oncologists. In this context, the Dutch Society of Pathology is already considering annual benchmarking of histologic grading of IBC based on "mirror" PALGA data, which may be adopted much broader in the field. In addition, future research might focus on specifically training pathologists in the assessment of histologic grade, which is underlined by Elston and Ellis, who emphasize that grading of IBC should only be undertaken by trained pathologists.[44] Pathologists might, for example, be trained by an e-learning, which could attribute to better synchronization of histologic grading.

Despite the indisputable need to improve histologic grading practices, it should be noted that other variables guiding breast cancer patient management have limitations as well. For example, HER2 and ER scoring, and the assessment of small nodal metastases are also subject to interobserver variation.[3,45] In addition, in contrast to other prognostic parameters, like genetic profiling tests, the evaluation of histologic grade is cheap and can in principle be performed in all breast cancer cases.[12,16] Furthermore, although molecular or genetic measures of prognosis may become increasingly important in the risk stratification of IBC, it is believed by Elston and Ellis that the future clinical application of molecular measures will be in combination with, and analogous to histologic grade, which is underlined by current (international) guidelines.[8,13,14,44] What is more, the decision to apply expression profiling of IBC is to a great extent based on histologic grade of the tumor.[8,12,14,15,46] Thus, the assessment of histologic grade may remain of great clinical importance as one of the best established prognostic factors for patients with breast cancer.

The impact of histologic grading is further underlined by our findings that treatment decisions on adjuvant therapy, according to the current guidelines, are solely dependent on histologic grade in almost one in every three patients, which highlights that histologic grading is of great clinical importance, as it influences treatment decisions and may subsequently influence outcome in a substantial part of patients. More than half of this group of patients (54.8%), for whom the indication for adjuvant chemotherapy was dependent on histologic grade (n = 5,821), even had a score on the switch point of grades (i.e., score 5 or score 6), indicating that a difference of only one point on the total score would already alter their indication for adjuvant chemotherapy. With the observed substantial grading variation in this study, it is very likely that this may have

influenced treatment decisions. Whether this subsequently influenced outcome of these patients should be the subject of future research. Overall, variation in grading may very easily lead to different treatment indications in a substantial part of patients.

Conclusion

In conclusion, the results of this large nationwide study show that there is substantial variation in the histopathologic grading of IBC, both between and within pathology laboratories. Reducing variation in grading is highly clinically relevant, as, for almost one in every three patients, the decision on adjuvant systemic chemotherapy solely depends on histologic grade. Hence, it is very likely that variation in grading influences treatment decisions and subsequently may influence outcome and exposure to unnecessary toxicity of individual patients. Interventions to improve nationwide histologic grading, for example, by e-learning, may especially focus on the assessment of nuclear polymorphism, as most variation was observed in this category.

Acknowledgements

The authors thank all pathology laboratories that participated in this study. We also thank Dr. C.H.M. van Deurzen, Dr. H.J. van Slooten, Dr. J. Wesseling and Dr. P.J. Westenend for their advice and for their involvement in the project-steering-group, and the registration team of the Netherlands Comprehensive Cancer Organization (IKNL) for the collection of data for the Netherlands Cancer Registry as well as IKNL staff for scientific advice. This research was funded by the Quality Foundation of the Dutch Association of Medical Specialists (SKMS). Funding was obtained by PD. The funding source had no role in the design and conduct of this study, nor did it have a role in the review or approval of the article and the decision to submit the article for publication.

Author Contributions

Design of the study: All authors. Data collection: van Dooijeweert C, Overbeek LIH and Deckers IAG; data analysis: van Dooijeweert C, van Diest PJ and Deckers IAG. Data interpretation: van Dooijeweert C, van Diest PJ, van der Wall E and Deckers IAG. Article writing: van Dooijeweert C. Read and revised the paper, and agreed with the final version of the paper: All authors. Funding: van Diest PJ.

References

1. National Institute for Health and Environment (RIVM). *Chance of breast cancer*. National Institute for Health and Environment (RIVM). Available from: https://www.rivm.nl/bevolkingsonderzoek-borstkanker/wat-is-borstkanker/kans-op-borstkanker [date last accessed March 20, 2019].
2. Netherlands Cancer Registry. *Cijfers over Kanker*. Netherlands Cancer Registry supplied by IKNL. Available from: www.cijfersoverkanker.nl [date last accessed March 20, 2019].
3. Rakha EA, Reis-Filho JS, Baehner F, et al. Breast cancer prognostic classification in the molecular era: the role of histological grade. *Breast Cancer Res.* 2010;12(4):207.
4. 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.
5. Sundquist M, Thorstenson S, Brudin L, et al. Applying the Nottingham Prognostic Index to a Swedish breast cancer population. South East Swedish Breast Cancer Study Group. *Breast Cancer Res Treat.* 1999;53(1):1–8.
6. Frkovic-Grazio S, Bracko M. Long term prognostic value of Nottingham histological grade and its components in early (pT1N0M0) breast carcinoma. *J Clin Pathol.* 2002;55(2):88–92.
7. Galea MH, Blamey RW, Elston CE, et al. The Nottingham Prognostic Index in primary breast cancer. *Breast Cancer Res Treat.* 1992;22(3):207–19.
8. The Netherlands Comprehensive Cancer Organisation. *Breast Cancer Guideline 2017 [updated 2012-07-02]*. Available from: https://www.oncoline.nl/borstkanker [date last accessed October 30, 2018].
9. Anttinen J, Kautiainen H, Kuopio T. Role of mammography screening as a predictor of survival in postmenopausal breast cancer patients. *Br J Cancer.* 2006;94(1):147–51.
10. Shen Y, Yang Y, Inoue LY, et al. Role of detection method in predicting breast cancer survival: analysis of randomized screening trials. *J Natl Cancer Inst.* 2005;97(16):1195–203.
11. Hanrahan EO, Valero V, Gonzalez-Angulo AM, et al. Prognosis and management of patients with node-negative invasive breast carcinoma that is 1 cm or smaller in size (stage 1; T1a,bN0M0): a review of the literature. *J Clin Oncol.* 2006;24(13):2113–22.
12. Cardoso F, van't Veer LJ, Bogaerts J, et al. 70-Gene Signature as an Aid to Treatment Decisions in Early-Stage Breast Cancer. *N Engl J Med.* 2016;375(8):717–29.
13. Senkus E, Kyriakides S, Ohno S, et al. Primary breast cancer: ESMO Clinical Practice Guidelines for diagnosis, treatment and follow-up†. *Ann Oncol.* 2015;26(suppl_5):v8–v30.
14. Curigliano G, Burstein HJ, PW E, et al. De-escalating and escalating treatments for early-stage breast cancer: the St. Gallen International Expert Consensus Conference on the Primary Therapy of Early Breast Cancer 2017. *Ann Oncol.* 2017;28(8):1700–12.
15. Krop I, Ismaila N, Andre F, et al. Use of biomarkers to guide decisions on adjuvant systemic therapy for women with early-stage invasive breast cancer: American Society of Clinical Oncology Clinical Practice Guideline Focused Update. *J Clin Oncol.* 2017;35(24):2838–47.
16. Boiesen P, Bendahl PO, Anagnostaki L, et al. Histologic grading in breast cancer—reproducibility between seven pathologic departments. *Acta Oncol.* 2000;39(1):41–5.
17. 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.
18. 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.
19. Lundin J, Lundin M, Holli K, et al. Omission of histologic grading from clinical decision making may result in overuse of adjuvant therapies in breast cancer: results from a nationwide study. *J Clin Oncol.* 2001;19(1):28–36.
20. Frierson HF Jr, 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.
21. Italian Network for Quality Assurance of Tumour Biomarkers (INQAT) Group. Quality control for histological grading in breast cancer: an Italian experience. *Pathologica.* 2005;97(1):1–6.
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. Kuijpers CC, Sluijter CE, von der Thusen JH, et al. Interlaboratory variability in the grading of dysplasia in a nationwide cohort of colorectal adenomas. *Histopathology.* 2016;69(2):187–97.
24. Kuijpers CC, Sluijter CE, von der Thusen JH, et al. Interlaboratory Variability in the Histologic Grading of Colorectal Adenocarcinomas in a Nationwide Cohort. *Am J Surg Pathol.* 2016;40(8):1100–8.
25. van Dooijeweert C, van Diest PJ, Willems SM, et al. Significant inter- and intra-laboratory variation in grading of ductal carcinoma in situ of the breast: a nationwide study of 4901 patients in the Netherlands. *Breast Cancer Res Treat.* 2019;174(2):479–88.
26. Casparie M, Tiebosch AT, Burger G, et al. Pathology databanking and biobanking in The Netherlands, a central role for PALGA, the nationwide

histopathology and cytopathology data network and archive. *Cell Oncol.* 2007;29(1):19–24.
27. Adams AL, Eltoum I, Krontiras H, et al. The effect of neoadjuvant chemotherapy on histologic grade, hormone receptor status, and HER2/neu status in breast carcinoma. *Breast J.* 2008;14(2): 141–6.
28. Zheng S, Zhang BL, Xiao T, et al. Comparison of histopathologic changes and expression of biomarkers in breast carcinoma before and after neoadjuvant chemotherapy. *Zhonghua Bing Li Xue Za Zhi* 2011;40(7):465–70.
29. Yin HF, Wang YH, Qin XQ, et al. Effect of neoadjuvant chemotherapy on histologic grade and expression of biological markers in breast cancer. *Zhonghua Zhong Liu Za Zhi* 2009;31(11):858–62.
30. Spiegelhalter DJ. Funnel plots for comparing institutional performance. *Stat Med.* 2005;24(8): 1185–202.
31. Veneroso C, Siegel R, Levine PH. Early age at first childbirth associated with advanced tumor grade in breast cancer. *Cancer Detect Prev.* 2008;32(3):215–23.
32. Huang HJ, Neven P, Drijkoningen M, et al. Association between tumour characteristics and HER-2/neu by immunohistochemistry in 1362 women with primary operable breast cancer. *J Clin Pathol.* 2005;58(6):611–6.
33. Rosenberg J, Chia YL, Plevritis S. The effect of age, race, tumor size, tumor grade, and disease stage on invasive ductal breast cancer survival in the U.S. SEER database. *Breast Cancer Res Treat.* 2005;89(1):47–54.
34. Sullivan PS, Apple SK. Should histologic type be taken into account when considering neoadjuvant chemotherapy in breast carcinoma? *Breast J.* 2009; 15(2):146–54.
35. Tubiana-Hulin M, Stevens D, Lasry S, et al. Response to neoadjuvant chemotherapy in lobular and ductal breast carcinomas: a retrospective study on 860 patients from one institution. *Ann Oncol.* 2006;17(8):1228–33.
36. Blamey RW, Hornmark-Stenstam B, Ball G, et al. ONCOPOOL—a European database for 16,944 cases of breast cancer. *Eur J Cancer.* 2010;46(1): 56–71.
37. Kiaer HW, Laenkholm AV, Nielsen BB, et al. Classical pathological variables recorded in the Danish Breast Cancer Cooperative Group's register 1978-2006. *Acta Oncol.* 2008;47(4): 778–83.
38. Moller H, Henson K, Luchtenborg M, et al. Short-term breast cancer survival in relation to ethnicity, stage, grade and receptor status: national cohort study in England. *Br J Cancer.* 2016;115(11): 1408–15.
39. Puig-Vives M, Sanchez MJ, Sanchez-Cantalejo J, et al. Distribution and prognosis of molecular breast cancer subtypes defined by immunohistochemical biomarkers in a Spanish population-based study. *Gynecol Oncol.* 2013;130(3):609–14.
40. Seneviratne S, Lawrenson R, Scott N, et al. Breast cancer biology and ethnic disparities in breast cancer mortality in New Zealand: a cohort study. *PLoS One.* 2015;10(4):e0123523.
41. Sun J, Chen C, Wei W, et al. Associations and indications of Ki67 expression with clinicopathological parameters and molecular subtypes in invasive breast cancer: A population-based study. *Oncol Lett.* 2015;10(3):1741–8.
42. Sluijter CE, van Lonkhuijzen LR, van Slooten HJ, et al. The effects of implementing synoptic pathology reporting in cancer diagnosis: a systematic review. *Virchows Archiv.* 2016;468(6):639–49.
43. PALGA Foundation. *Annual Report 2017.* Available from: https://www.palga.nl/assets/uploads/Jaarverslag/Jaarverslag_2017_websiteversie.pdf [date last accessed October 20, 2018].
44. 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* 2002;41(3A):154–61.
45. Orlando L, Viale G, Bria E, et al. Discordance in pathology report after central pathology review: Implications for breast cancer adjuvant treatment. *Breast* 2016;30:151–5.
46. Groenendijk FH, Jager A, Cardoso F, et al. A nationwide registry-based cohort study of the MammaPrint genomic risk classifier in invasive breast cancer. *Breast.* 2018;38:125–31.