

RESEARCH

Open Access



Grading variation in 2,934 patients with ductal carcinoma in situ of the breast: the effect of laboratory- and pathologist-specific feedback reports

Carmen van Dooijeweert¹, Paul J. van Diest^{1*} , Inge O. Baas², Elsken van der Wall² and Ivette A. G. Deckers³

Abstract

Background: Histologic grade of ductal carcinoma in situ of the breast (DCIS) may become the single biomarker that decides whether patients will be treated. Yet, evidence shows that grading variation in daily practice is substantial. To facilitate quality improvement, feedback reports, in which laboratory-specific case-mix adjusted proportions per grade were benchmarked against other laboratories, were sent to the individual laboratories by March 1, 2018. One year later, the effect of these feedback reports on inter-laboratory variation was studied.

Methods: Synoptic pathology reports of all pure DCIS resection specimens between March 1, 2017 and March 1, 2019 were retrieved from PALGA (the nationwide Dutch pathology registry). Laboratory-specific proportions per grade were compared to the overall proportion in the year before and after feedback. The absolute deviation for all three grades at once, represented by the overall deviation score (ODS), was calculated as the sum of deviations from the grade-specific overall proportions. Case-mix adjusted, laboratory-specific odds ratios (ORs) for high- (grade III) versus low-grade (grade I-II) DCIS were obtained by multivariable logistic regression.

Results: Overall, 2954 DCIS reports from 31 laboratories were included. After feedback, the range between laboratories decreased by 22 and 6.5% for grades II and III, while an increase of 6.2% was observed for grade I. Both the mean ODS (27.2 to 24.1%) and maximum ODS (87.7 to 59.6%) decreased considerably. However, the range of case-mix adjusted ORs remained fairly stable and substantial (0.39 (95% CI: 0.18–0.86) to 3.69 (95% CI: 1.30–10.51)).

Conclusion: A promising decrease in grading variation was observed after laboratory-specific feedback for DCIS grades II-III, while this was not observed for DCIS grade I. Overall, grading variation remained substantial which needs to be addressed considering its clinical implications. Nationwide consensus on a classification, and training of (expert breast) pathologists, for example by e-learning, may help to further improve grading standardization.

Keywords: DCIS, ductal carcinoma in situ, histologic grade, patient management, laboratory-specific feedback, daily pathology practice

* Correspondence: p.j.vandiest@umcutrecht.nl

¹Department of Pathology, University Medical Center Utrecht, PO Box 85500, 3508, GA, Utrecht, the Netherlands

Full list of author information is available at the end of the article



© The Author(s). 2020 **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>. The Creative Commons Public Domain Dedication waiver (<http://creativecommons.org/publicdomain/zero/1.0/>) applies to the data made available in this article, unless otherwise stated in a credit line to the data.

Introduction

Treatment of ductal carcinoma in situ of the breast (DCIS) currently consists of surgery [1, 2], radiotherapy [2–6] and sometimes even (low-dose) tamoxifen [2–4, 7–9]. However, it is believed that an unknown number of DCIS patients are treated for lesions that may never progress into invasive breast cancer [10–13]. Therefore, four randomized controlled clinical trials aim to identify a group of low-risk DCIS patients that, under active surveillance, may safely forgo surgical treatment [11, 12, 14–16].

For all trials, the main biomarker which identifies DCIS as being low-risk, is histologic (nuclear) grade, although different classifications are used by the LORD-, LORIS-, COMET-, and LARRIKIN trials [11, 12, 14–16]. The general hypothesis of these trials is that progression risk, or at least speed of progression is higher for high-grade lesions [17, 18], and if a low-grade DCIS does become invasive, it will be a low-grade invasive carcinoma with favorable characteristics and excellent survival rates after treatment [11, 19].

Besides the fact that grade may become the single biomarker that is used to decide whether patients are treated for their DCIS, grade already plays an important role in clinical patient management. For example, grade influences radiotherapy decisions (omitting a boost, considering partial breast irradiation) [1, 6] and indicates (on biopsy) whether a sentinel lymph node procedure is required [1]. Thus, accurate, consistent, and reproducible grading is of key importance. However, we previously showed that variation in grading, between pathology laboratories and between pathologists within laboratories, is substantial in daily clinical practice on a nationwide scale in the Netherlands [20]. Furthermore, studies in which a set of DCIS was assessed by several pathologists showed significant inter-observer variation, regardless of the used classification, as well [21–23].

As studies have shown that quality of breast cancer care can be improved by auditing and benchmarking [24–29], the results of our previous study were sent to all participating Dutch pathology laboratories as feedback reports, in which their proportions per grade were benchmarked against other laboratories. This enabled pathologists to discuss and reflect upon their grading practices. The present study was conducted to investigate the effect of these feedback reports on grading variation between laboratories on a nationwide scale.

Methods

Data source

All data were retrieved from PALGA, the Dutch nationwide network and registry of histo- and cytopathology, which contains excerpts of all pathology reports from laboratories in the Netherlands [30]. Data within this database are pseudonymized in the laboratories themselves and by a trusted third party (ZorgTTP, Houten, the

Netherlands). In addition, as data on the reporting pathologist was not available in PALGA, laboratories provided the pathologist information directly to the PALGA data-analyst. In a final step, the PALGA data-analyst anonymized all laboratories and pathologists to the researchers. This study was approved by the scientific and privacy committee of PALGA and data were retrieved and handled in compliance with the General Data Protection Regulation Act (GDPR).

Study population

We retrieved all synoptic pathology reports of patients with pure DCIS resection specimens (i.e. without a report of a known adjacent invasive breast cancer) in the Netherlands between March 1, 2017 and March 1, 2019 from PALGA ($n = 3336$). Per pathology report, patient- and tumor characteristics were extracted (sex, age, type of surgery, histologic grade and tumor size). Reports with missing data on histologic grade and/or tumor size were excluded (Supplementary Fig. 1). Pathology reports of residual in situ lesions after neoadjuvant treatment of primary invasive breast carcinomas were excluded. Furthermore, ipsilateral DCIS reports within six months of the previous DCIS resection specimen report were considered paired measurements of which we only included the first report.

In total, 38 out of 42 pathology laboratories in the Netherlands reported resection specimens via the synoptic (PALGA) protocol from March 1, 2017 and onwards. Of these 38 laboratories, we included those that annually reported a minimum of 15 DCIS resection specimens within the protocol.

Feedback reports

Laboratory-specific feedback reports, with proportions per histologic grade of individual laboratories benchmarked against other anonymized laboratories [20], were sent out by March 1, 2018. The general feedback report is available on the PALGA website (in Dutch only) [31], while laboratory-specific reports are only available to the individual laboratories themselves.

Laboratory-specific reports consisted of funnel plots, in which absolute differences in proportion of histologic grade, are plotted against the number of included IBC per laboratory. Within these funnel plots, the national proportions per grade with their 95% confidence intervals as limits were set as targets.

In addition, all laboratories were asked to provide coded information of the reporting pathologist per pathology report, to compare grading practices of the different pathologists within their laboratory. These data on pathologist' level were provided by ten laboratories, which as a result received feedback on both laboratory- and pathologist' level. Thus, pathologists within these

ten laboratories were enabled to discuss and reflect upon both their personal- and overall laboratory grading practices.

As feedback reports were sent to the laboratories by March 1, 2018, we considered the period from March 1, 2017 up to and including March 1, 2018 as pre-feedback period, while the period from March 2, 2018 up to and including March 1, 2019 is considered the post-feedback period.

DCIS grading classification

Histologic grade was defined in the pathology reports as grade I, II or III, without a specification of which classification was used. The Dutch guideline [1] recommends to use the classification of Holland [32]. However, we know from our previous research that numerous different guidelines are used by Dutch pathologists in daily practice [20].

Statistical analysis

Patient and tumor characteristics from the pathology reports were summarized in Table 1. Differences of these characteristics between pre- and post-feedback pathology reports were tested by means of a χ^2 -test for categorical variables and by means of a Mann-Whitney U test for continuous variables.

Overall proportions for DCIS grades I-III were determined pre- and post-feedback. Absolute differences from the overall proportions per laboratory are presented in bar charts per feedback period for grades I-III. We calculated an overall deviation score (ODS), consisting of the sum of absolute deviations from the grade-specific overall proportions per period, to compare the absolute deviation for all grades at once. Differences in ODS of

the individual laboratories between the pre- and post-feedback period were compared by means of a Wilcoxon signed-rank test.

We used several, arbitrary, definitions of change as a possible way to interpret the type of change in laboratories. Laboratories who showed an absolute change of $\leq 2\%$ were defined as 'not shifting'. Among laboratories that showed an absolute change of $> 2\%$, laboratories moving closer to the overall mean after feedback were defined as 'less deviant', while laboratories moving further from the overall mean after feedback were defined as 'more deviant'.

A logistic regression analyses, providing odds ratios (ORs) and 95% confidence intervals (CI) per laboratory, was used to compare relative differences among laboratories. For the logistic regression model, grade was dichotomized into low-grade DCIS (grade I-II) and high-grade DCIS (grade III), as this is the same definition that is used by the majority of the clinical trials [11, 12, 14, 16, 20]. The laboratory best resembling the national distribution on low- (grade I-II) and high-grade (grade III) was chosen as reference laboratory. We performed a multivariate logistic regression analysis to correct for differences in case-mix. Variables were selected based on our previous research [20] and consisted of tumor size and type of surgery.

To compare differences in case-mix adjusted ORs of the individual laboratories, the positive OR difference, i.e. the difference of the laboratory OR to the reference OR of 1.00, was calculated. These positive OR differences of the individual laboratories were compared pre- and post-feedback by Wilcoxon signed rank test.

All analyses were performed by using IBM SPSS Statistics version 15.0.0.2.

Table 1 Characteristics of the 2954 included ductal carcinoma in situ (DCIS) lesions from the PALGA database between March 1, 2017 and March 1, 2018 (PRE feedback), and March 1, 2018 to March 1, 2019 (POST feedback)

| | Total (n = 2954) | PRE (n = 1493) | POST (n = 1461) | p-value ^b |
|------------------------------|------------------|----------------|-----------------|----------------------|
| Histologic grade, n (%) | | | | |
| Grade 1 | 383 (13.0%) | 169 (11.3%) | 214 (14.6%) | 0.016 |
| Grade 2 | 1171 (39.6%) | 590 (39.5%) | 581 (39.8%) | |
| Grade 3 | 1400 (47.4%) | 734 (49.2%) | 666 (45.6%) | |
| Age (y) ^a | 59.8 (10.3) | 59.8 (10.5) | 59.8 (10.1) | 0.922 |
| Sex, n (%) | | | | 0.790 |
| Female | 2943 (99.6%) | 1487 (99.6%) | 1456 (99.7%) | |
| Male | 11 (0.4%) | 6 (0.4%) | 5 (0.3%) | |
| Tumor size (cm) ^a | 2.5 (2.3) | 2.5 (2.3) | 2.5 (2.3) | 0.321 |
| Type of surgery, n (%) | | | | 0.096 |
| Mastectomy | 944 (32.0%) | 456 (30.5%) | 488 (33.4%) | |
| Breast conserving | 2010 (68.0%) | 1037 (69.5%) | 973 (66.6%) | |

^a Mean (SD)

^b PRE vs POST categorical variables by Chi-square test, continuous variables by Mann Whitney U Test

Results

Patient-, tumor- and laboratory characteristics

In total, 2954 DCIS resection specimen reports from 2934 patients were included. For 20 patients, two pathology reports were included, as this either concerned a bilateral tumor ($n = 18$) or an ipsilateral tumor > 6 months after the first diagnosis ($n = 2$). Of the included reports, 1493 were defined as pre-feedback, while 1461 were defined as post-feedback.

Characteristics of the included DCIS resection specimen reports are summarized in Table 1. All included pathology reports originate from 31/42 pathology laboratories. Four laboratories did not implement synoptic reporting between March 1, 2017 and March 1, 2019. The remaining seven laboratories were excluded as they synoptically graded less than 15 DCIS lesions within the PALGA protocol pre- and/or post-feedback. The overall number of DCIS reports before feedback ranged from 19 to 94 (median 48), whereas after feedback the number of synoptic pathology reports ranged from 19 to 86 (median 39).

Mean age (SD) at diagnosis was 59.8 (10.3) years and mean tumor size (SD) was 2.5 (2.3) cm. Overall, only 11 men were included (0.4%). Breast conserving surgery was performed in almost 70% of DCIS patients, although a small decrease was observed after feedback (69.5% versus 66.6%, $p = 0.096$). A significant change in distribution of histologic grade was observed after feedback ($p = 0.016$), as the proportion of grade I increased (from 11.3 to 14.6%), whereas the proportion of grade III decreased (from 49.2 to 45.6%), while grade II remained fairly stable (39.5% versus 39.8%) (Table 1).

Inter-laboratory differences in histologic grading

For grade II, the total range between laboratories decreased markedly by 22% (from 9.1–83.3% to 10.0–62.2%) and, for grade III, to a lesser extent by 6.5% (from 16.7–75.8% to 21.3–73.9%) (Fig. 1). In contrast, the range between laboratories increased by 6.2% for grade I (from 0.0–21.1% to 2.7–30.0%) (Fig. 1).

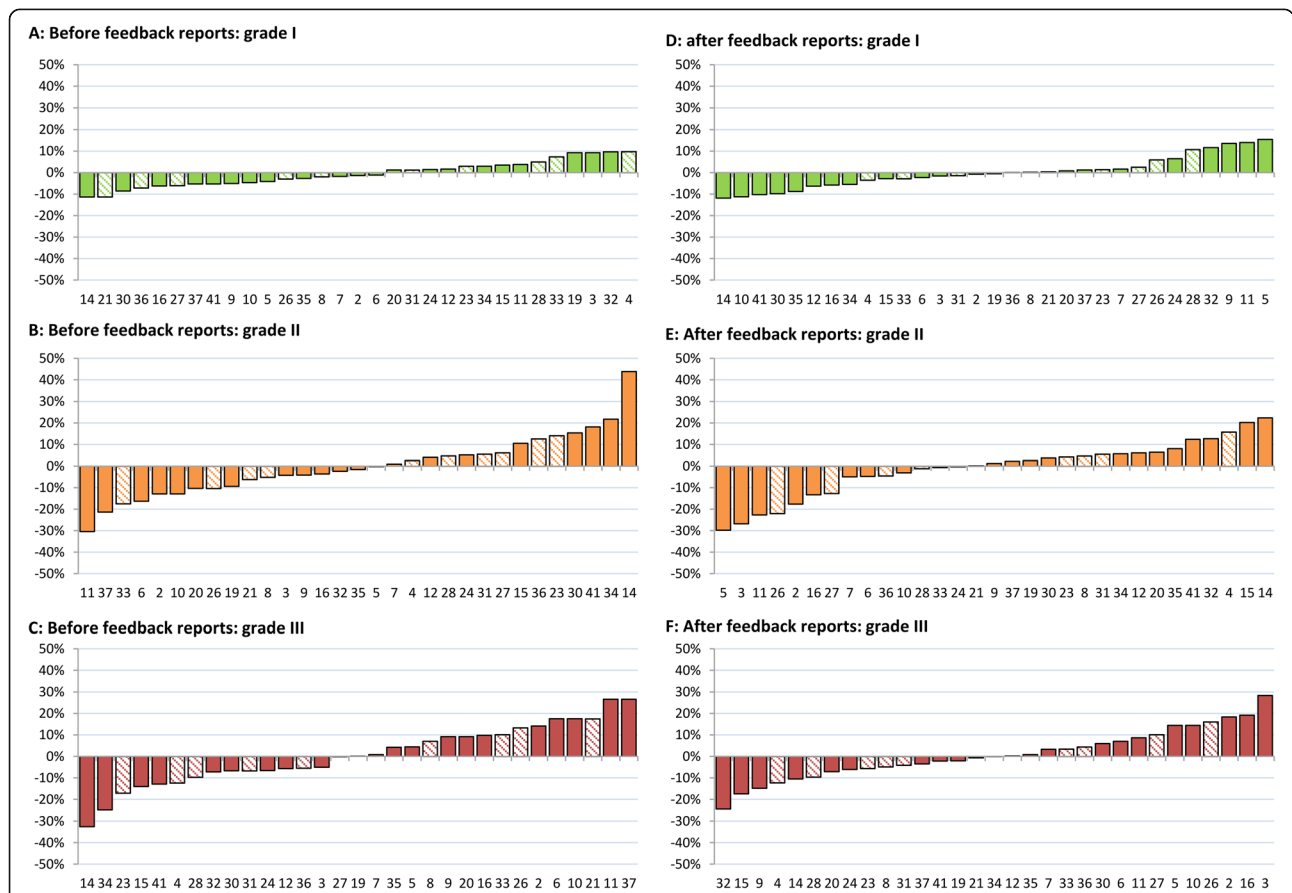


Fig. 1 Inter-laboratory ($n = 31$) variation in histologic grading of ductal carcinoma in situ (DCIS) of the breast before (a-c) ($n = 1493$) and after feedback reports (d-f) ($n = 1461$). Percentages per laboratory show the absolute deviation from the national proportion per histologic grade for grade I (A + D), grade II (B + E) and grade III (C + F). Laboratory numbers for all sub-figures (a-f) correspond. All laboratories are ranged from lower (negative values) to higher proportions (positive values). A decrease in absolute range was observed for grades II and III; grade II (-22.0%), grade III (-6.2%). An increase was observed for grade I (+6.2%). Striped bars indicate laboratories who received feedback on pathologist' level ($n = 10$)

After feedback, the maximum ODS decreased considerably from 87.7 to 59.6%, while the mean ODS also decreased from 27.2 to 24.1% (Fig. 2). The ODS of the individual laboratories did not differ significantly (Wilcoxon signed rank test, $p = 0.456$).

The majority of laboratories became less deviant after feedback for grades II and III (18 laboratories (58.1%) and 17 laboratories (54.8%) respectively), while 11 laboratories (35.5%) became more deviant for both grades. For grade I, the number of laboratories that became less deviant was similar to the number of laboratories that became more deviant; 11 (35.5%) versus 12 (38.8%) laboratories, while 8 (25.8%) laboratories showed a shift of $\leq 2\%$ (Table 2).

Laboratory 19 was chosen as reference laboratory in the multivariate logistic regression model, as it best resembled the national distribution for grade III versus grade I-II (mean deviation for grade III before and after feedback 1.1%). Laboratory-specific ORs ranged from

0.20 (95% CI: 0.06–0.65) to 3.39 (95% CI: 1.34–8.57) before feedback, resulting in an overall OR range of 3.19. After feedback, laboratory-specific ORs ranged from 0.39 (95% CI: 0.18–0.86) to 3.69 (95% CI: 1.30–10.51), with a corresponding OR range of 3.30 (Fig. 3). Subsequently, the overall OR range increased by 3.5% after feedback. Positive OR-differences within the laboratories did not significantly differ before and after feedback (Wilcoxon signed-rank test, $p = 0.886$).

Feedback on pathologist' level

Compared to the slight decrease in mean ODS (from 28.8 to 27.5%) of laboratories who only received feedback on laboratory-level, the mean ODS of the ten laboratories who did also receive feedback on pathologist' level, showed a larger decrease from 24.0 to 17.1%. As for the type of change in laboratories, it seems that laboratories who also received feedback on pathologist'

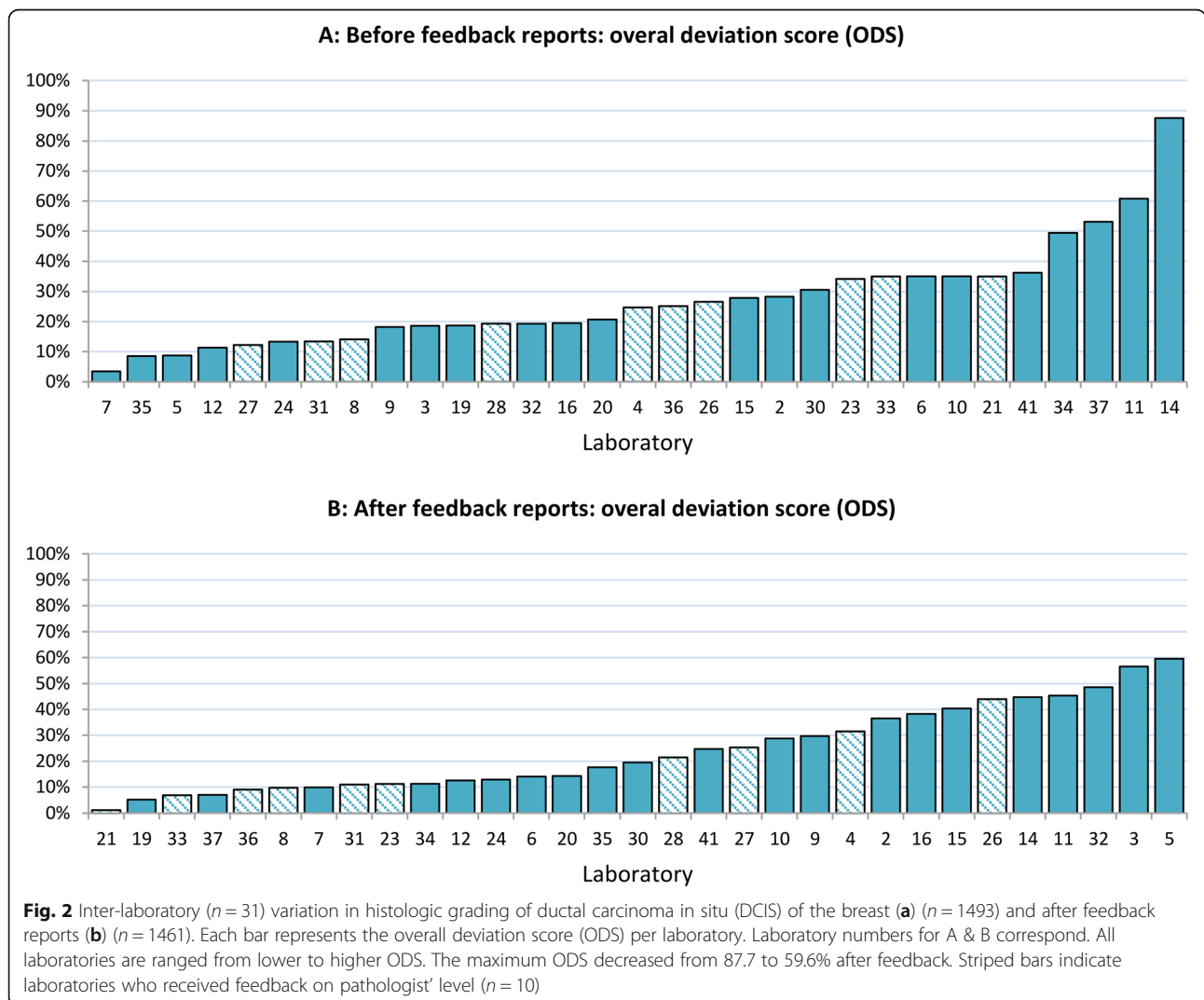


Table 2 Type of change in the laboratories after feedback per histologic grade

| Type of change | Total number of laboratories (n = 31) | Laboratories with feedback on pathologist' level (n = 10) | Laboratories without feedback on pathologist' level (n = 21) | p-value |
|---|---------------------------------------|---|--|---------|
| Grade I | | | | |
| No shift ($\leq 2\%$) | 8 (25.8%) | 1 (10.0%) | 7 (33.3%) | 0.165* |
| Shift (> 2%) | | | | |
| Less deviant | 11 (35.5%) | 6 (60.0%) | 5 (23.8%) | 0.147** |
| More deviant | 12 (38.7%) | 3 (30.0%) | 9 (42.9%) | |
| Grade II | | | | |
| No shift ($\leq 2\%$) | 2 (6.5%) | 1 (10.0%) | 1 (4.8%) | 0.579* |
| Shift (> 2%) | | | | |
| Less deviant | 18 (58.1%) | 6 (60.0%) | 12 (57.1%) | 0.732** |
| More deviant | 11 (35.5%) | 3 (30.0%) | 8 (38.1%) | |
| Grade III | | | | |
| No shift ($\leq 2\%$) | 3 (9.7%) | 2 (20.0%) | 1 (4.8%) | 0.180* |
| Shift (> 2%) | | | | |
| Less deviant | 17 (54.8%) | 6 (60.0%) | 11 (52.4%) | 0.328** |
| More deviant | 11 (35.5%) | 2 (20.0%) | 9 (42.9%) | |

* p-value for shift vs. no shift between laboratories with and without feedback on pathologist' level

** p-value for type of change when laboratories shifted > 2% after feedback between laboratories with and without feedback on pathologist' level

level showed more improvement, however, these differences were not statistically significant (Table 2).

Discussion

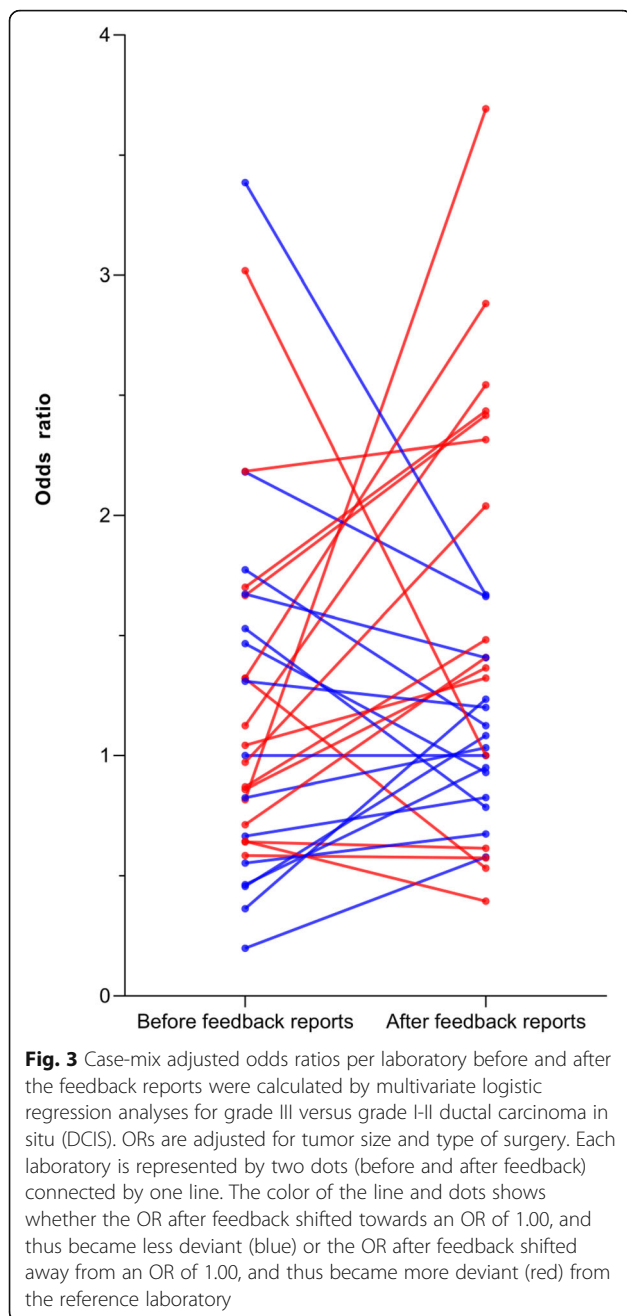
Using data from structured pathology reports, we investigated the effect of feedback reports on nationwide inter-laboratory grading variation of DCIS. The decrease in absolute variation for grades II (6.5%) and III (22.0%), as well as the overall majority of laboratories becoming less deviant after feedback, and the decrease of both the mean and maximum ODS seem to indicate a promising decrease in DCIS grading variation. However, absolute variation increased for grade I (6.2%), and the range of case-mix adjusted ORs remained fairly stable and large after feedback. Furthermore, the absolute range between laboratories remains substantial, and maybe even clinically unacceptable, for all grades.

We hypothesize that the lack of consensus on a grading classification [13, 20], which is also reflected by the use of different classifications by the trials [11, 12, 14–16], may be the explanation for these mixed results as we believe that uniform grading starts with the use of single grading classification by all pathologists. Furthermore, in comparison to grading of invasive breast cancer, which is performed according to the modified Bloom and Richardson grading classification (Elston-Ellis modification) [33, 34], with scoring of the three subcategories (tubular differentiation, nuclear polymorphism and mitotic count) grading of DCIS is less standardized. Recently, dichotomous histopathological assessment of ductal carcinoma was studied as an alternative with (better) acceptable degrees of

interobserver variability, however, interobserver variation remained considerable and at best acceptable [35]. In addition, other recent data showed that even among 35 expert breast pathologists, the threshold for comedonecrosis is highly variable [36]. This highlights the complexity of histologic grading of DCIS and the need for clear and uniform guidelines [36]. This not only important to the possible implementation of trials results into daily clinical practice, but it may also benefit the quality of the trials itself as central pathology review is not always carried out [37].

Interestingly, 7 out of 38 laboratories that used the synoptic PALGA protocol, were excluded for grading less than 15 DCIS resection specimens (via the protocol), which practically means they grade little over one (pure) DCIS specimen per month. For two laboratories this was because they likely only started using the protocol while another laboratory, stopped using the protocol for unknown reasons. Yet, for the remaining four laboratories, numbers per year (pre- or post-feedback) were fairly stable and low. We would like to emphasize, however, that pathologists within these laboratories may still report DCIS resection specimens outside the protocol, and therefore, they may grade more DCIS than our data would suggest. Nevertheless, if these are the actual numbers of DCIS' that are assessed in specific pathology laboratories per year, it may be questioned whether this is desirable, especially with regard to the complexity of histologic grading.

Easy data extraction on a large (nationwide) scale and an increased overall completeness of reports [38] were



the main reasons to only include synoptic pathology reports. Furthermore, over 80% of breast resection specimens is reported via the synoptic protocol [39].

The mean numbers per histologic grade in this study (13.0% for grade I, 39.6% for grade II and 47.4% for grade III) are in line with other studies [20, 40, 41]. Interestingly, we observed a significant change in grade distribution after feedback as the proportion of grade I tumors increased by over 3%, while the proportion of grade III

tumors decreased by a similar percentage. Whether this is initiated by the feedback reports or whether it reflects a true change in the population of DCIS patients remains unknown. Nonetheless, it did make it more difficult to interpret the results regarding deviations from the mean.

Overall, after analyzing the data in an absolute and relative manner, we did observe promising and positive changes, reflected by the decrease in absolute variation for DCIS grade II and III and the decrease of both the mean and maximum ODS. Furthermore, the majority of laboratories became less deviant after feedback for grades II and III. Hence, for grades II and III most deviant laboratories became less deviant, indicating that there is a decrease of the extremes, while changes of individual laboratories (ODS, positive OR-differences) were not significant. In contrast, the results for DCIS grade I showed an increase in the absolute range between laboratories, while the overall range of ORs remained stable and substantially large, ranging from 0.39 (95% CI: 0.18–0.86) to 3.69 (95% CI: 1.30–10.51). This shows that variation in histologic grading is still far from clinically acceptable levels.

Our results confirmed that feedback on the individual level (i.e. the pathologist) may be more effective than feedback on the general level (i.e. laboratories) [42–45]. We observed a larger absolute decrease of the mean ODS of laboratories who also received individual feedback. Furthermore, these laboratories also showed more improvement as compared to laboratories who only received feedback on laboratory-level. Due to the low number of laboratories who received pathologist-specific feedback reports, differences were not statistically significant.

The observed effects on grading variation may not exclusively be attributed to the feedback reports. However, we would like to emphasize that between 2016 and March 1, 2019, besides the feedback reports, no other interventions or guideline changes took place. Furthermore, our previous paper [20] was only first published after feedback reports were sent to the individual pathology laboratories. We believe that these feedback reports may be a useful tool to, at least, monitor grading variation in daily clinical practice.

Conclusion

We observed a promising decrease in grading variation for DCIS grades II and III, while this was not observed for DCIS grade I. The overall variation for all grades remains substantial, and it seems that histologic grade is far from being a clinically acceptable biomarker for DCIS, let alone be the single biomarker that decides whether patients may be treated. In light of the current ongoing trials, improvement and standardization of DCIS grading is adamant. Continuing with feedback

reports, especially on pathologist' level, helps to create awareness and may open the discussion about nationwide consensus on a single grading classification. In addition, adequate training of (expert breast) pathologists, according to a single classification system, for example by e-learning, may help to establish uniform grading in clinical practice over time.

Supplementary information

Supplementary information accompanies this paper at <https://doi.org/10.1186/s13000-020-00970-8>.

Additional file 1 Supplementary Fig. 1. Flowchart of included pathology reports of ductal carcinoma in situ of the breast (DCIS) resection specimens to assess the effect of feedback reports on variation in histologic grading of DCIS.

Abbreviations

DCIS: Ductal carcinoma in situ of the breast (DCIS); PALGA: The nationwide Dutch pathology registry; GDPR: General Data Protection Regulation Act; SKMS: Quality Foundation of the Dutch Association of Medical Specialists

Authors' contributions

All authors (CD, PD, IB, EW, ID) were involved in the design of the study. Data were collected by CD and ID, and analysed by CD. Data interpretation was done by CD, PD, EW and ID. CD wrote the paper. All authors (CD, PD, IB, EW, ID) read and revised the paper, and agreed with the final version of the paper. Funding was obtained by PD.

Funding

This research was funded by the Quality Foundation of the Dutch Association of Medical Specialists (SKMS). 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 manuscript and the decision to submit the manuscript for publication.

Availability of data and materials

All data are available upon request from PALGA (the nationwide registry of histo- and cytopathology in the Netherlands).

Ethics approval and consent to participate

This study was approved by the scientific and privacy committee of PALGA (the nationwide registry of histo- and cytopathology in the Netherlands). As this study was performed retrospectively on anonymised data, and did not involve subjecting patients to anything, consent from the ethics committee was not required according to the Dutch law. All data from the PALGA database are pseudonymized by a trusted third party (ZorgTTP, Houten, The Netherlands). We only ever received fully anonymised data. Data of patients who object against the scientific use of their (anonymised) data are not included in the PALGA database. Therefore waiver of informed consent is applicable. All data were retrieved and handled in compliance with the General Data Protection Regulation (GDPR).

Consent for publication

Not applicable (see above).

Competing interests

All authors declare that they have no conflict of interest.

Author details

¹Department of Pathology, University Medical Center Utrecht, PO Box 85500, 3508, GA, Utrecht, the Netherlands. ²Department of Medical Oncology, University Medical Center Utrecht, Utrecht, the Netherlands. ³Foundation PALGA (the nationwide network and registry of histo- and cytopathology in the Netherlands), Houten, the Netherlands.

Received: 25 February 2020 Accepted: 4 May 2020

Published online: 11 May 2020

References

1. The Netherlands Comprehensive Cancer Organisation (IKNL) (2017). Oncoline: breast Cancer guideline. <https://www.oncoline.nl/borstkanker>. Accessed December 4 2019.
2. Curigliano G, Burstein HJ, PW E, Gnani M, Dubsy P, Loibl S, 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.
3. Correa C, McGale P, Taylor C, Wang Y, Clarke M, Davies C, et al. Overview of the randomized trials of radiotherapy in ductal carcinoma in situ of the breast. *J Natl Cancer Inst Monogr*. 2010;2010(41):162–77.
4. Gorringer KL, Fox SB. Ductal carcinoma in situ biology, biomarkers, and diagnosis. *Front Oncol*. 2017;7:248.
5. Morrow M, Van Zee KJ, Solin LJ, Houssami N, Chavez-MacGregor M, Harris JR, et al. Society of Surgical Oncology-American Society for Radiation Oncology-American Society of Clinical Oncology consensus guideline on margins for breast-conserving surgery with whole-breast irradiation in ductal carcinoma in situ. *J Clin Oncol*. 2016;34(33):4040–6.
6. Smith BD, Bellon JR, Blitzblau R, Freedman G, Haffty B, Hahn C, et al. Radiation therapy for the whole breast: executive summary of an American Society for Radiation Oncology (ASTRO) evidence-based guideline. *Pract Radiat Oncol*. 2018;8(3):145–52.
7. Staley H, McCallum I, Bruce J. Postoperative tamoxifen for ductal carcinoma in situ: Cochrane systematic review and meta-analysis. *Breast*. 2014;23(5):546–51.
8. Cuzick J, Sestak I, Pinder SE, Ellis IO, Forsyth S, Bundred NJ, et al. Effect of tamoxifen and radiotherapy in women with locally excised ductal carcinoma in situ: long-term results from the UK/ANZ DCIS trial. *Lancet Oncol*. 2011;12(1):21–9.
9. DeCensi A, Puntoni M, Guerrieri-Gonzaga A, Caviglia S, Avino F, Cortesi L, et al. Randomized placebo controlled trial of low-dose tamoxifen to prevent local and contralateral recurrence in breast intraepithelial neoplasia. *J Clin Oncol*. 2019;37(19):1629–37.
10. Benson JR, Jatoi I, Toi M. Treatment of low-risk ductal carcinoma in situ: is nothing better than something? *Lancet Oncol*. 2016;17(10):e442–e51.
11. Elshof LE, Tryfonidis K, Slaets L, van Leeuwen-Stok AE, Skinner VP, Dif N, et al. Feasibility of a prospective, randomised, open-label, international multicentre, phase III, non-inferiority trial to assess the safety of active surveillance for low risk ductal carcinoma in situ - the LORD study. *Eur J Cancer*. 2015;51(12):1497–510.
12. Francis A, Thomas J, Fallowfield L, Wallis M, Bartlett JMS, Brookes C, et al. Addressing overtreatment of screen detected DCIS; the LORIS trial. *Eur J Cancer*. 2015;51(16):2296–303.
13. van Seijen M, Lips EH, Thompson AM, Nik-Zainal S, Futreal A, Hwang ES, et al. Ductal carcinoma in situ: to treat or not to treat, that is the question. *Br J Cancer*. 2019;121(4):285–92.
14. Hwang ES, Hyslop T, Lynch T, Frank E, Pinto D, Basila D, et al. The COMET (comparison of operative versus monitoring and endocrine therapy) trial: a phase III randomised controlled clinical trial for low-risk ductal carcinoma in situ (DCIS). *BMJ Open*. 2019;9(3):e026797.
15. Groen EJ, Elshof LE, Visser LL, Rutgers EJT, Winter-Warnars HAO, Lips EH, et al. Finding the balance between over- and under-treatment of ductal carcinoma in situ (DCIS). *Breast*. 2017;31:274–83.
16. Lippey J, Spillane A, Saunders C. Not all ductal carcinoma in situ is created equal: can we avoid surgery for low-risk ductal carcinoma in situ? *ANZ J Surg*. 2016;86(11):859–60.
17. Sanders ME, Schuyler PA, Dupont WD, Page DL. The natural history of low-grade ductal carcinoma in situ of the breast in women treated by biopsy only revealed over 30 years of long-term follow-up. *Cancer*. 2005;103(12):2481–4.
18. Toss M, Miligy I, Thompson AM, Khout H, Green AR, Ellis IO, et al. Current trials to reduce surgical intervention in ductal carcinoma in situ of the breast: critical review. *Breast*. 2017;35:151–6.
19. Lopez-Garcia MA, Geyer FC, Lacroix-Triki M, Marchió C, Reis-Filho JS. Breast cancer precursors revisited: molecular features and progression pathways. *Histopathology*. 2010;57(2):171–92.
20. van Dooijeweert C, van Diest PJ, Willems SM, Kuijpers C, Overbeek LH, Deckers IAG. 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.
21. Douglas-Jones AG, Morgan JM, Appleton MA, Attanoos RL, Caslin A, Champ CS, et al. Consistency in the observation of features used to classify duct carcinoma in situ (DCIS) of the breast. *J Clin Pathol.* 2000;53(8):596–602.
 22. Elston CW, Sloane JP, Amendoeira I, Apostolikas N, Bellocq JP, Bianchi S, et al. Causes of inconsistency in diagnosing and classifying intraductal proliferations of the breast. European Commission working group on breast screening pathology. *Eur J Cancer.* 2000;36(14):1769–72.
 23. Schuh F, Biazus JV, Resetkova E, Benfica CZ, Edelweiss MI. Reproducibility of three classification systems of ductal carcinoma in situ of the breast using a web-based survey. *Pathol Res Pract.* 2010;206(10):705–11.
 24. Laronga C, Gray JE, Siegel EM, Lee JH, Fulp WJ, Fletcher M, et al. Florida initiative for quality Cancer care: improvements in breast cancer quality indicators during a 3-year interval. *J Am Coll Surg.* 2014;219(4):638–45.e1.
 25. Sacerdote C, Bordon R, Pitarella S, Mano MP, Baldi I, Casella D, et al. Compliance with clinical practice guidelines for breast cancer treatment: a population-based study of quality-of-care indicators in Italy. *BMC Health Serv Res.* 2013;13:28.
 26. Sringley J, Lankshear S, Brierley J, McGowan T, Divaris D, Yurcan M, et al. Closing the quality loop: facilitating improvement in oncology practice through timely access to clinical performance indicators. *J Oncol Pract.* 2013;9(5):e255–61.
 27. van Bommel AC, Spronk PE, Vrancken Peeters MT, Jager A, Lobbes M, Maduro JH, et al. Clinical auditing as an instrument for quality improvement in breast cancer care in the Netherlands: the national NABON breast Cancer audit. *J Surg Oncol.* 2017;115(3):243–9.
 28. van Dam PA, Verkinderen L, Hauspy J, Vermeulen P, Dirix L, Huizing M, et al. Benchmarking and audit of breast units improves quality of care. *Facts Views Vis Obgyn.* 2013;5(1):26–32.
 29. Veerbeek L, van der Geest L, Wouters M, Guicherit O, Does-den Heijer A, Nortier J, et al. Enhancing the quality of care for patients with breast cancer: seven years of experience with a Dutch auditing system. *Eur J Surg Oncol.* 2011;37(8):714–8.
 30. Casparie M, Tiebosch ATMG, Burger G, Blauwgeers H, van de Pol A, van Krieken JHJM, 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.
 31. Deckers IAG. Feedback variation in grading of invasive breast cancer and DCIS (in Dutch). In: PALGA, the nationwide network and registry of histo- and cytopathology in the Netherlands, Houten, the Netherlands; 2018. https://www.palga.nl/assets/uploads/Spiegelinfo/PALGA%20rapport%20-%20Terugkoppeling%20variatie%20in%20gradering%20mammacarcinoom_algemeen.pdf. Accessed December 4 2019.
 32. Holland R, Peterse JL, Millis RR, Eusebi V, Faverly D, van de Vijver MJ, et al. Ductal carcinoma in situ: a proposal for a new classification. *Semin Diagn Pathol.* 1994;11(3):167–80.
 33. 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.
 34. 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.
 35. Dano H, Altinay S, Arnould L, Bletard N, Colpaert C, Dedeurwaerdere F, et al. Interobserver variability in upfront dichotomous histopathological assessment of ductal carcinoma in situ of the breast: the DCISion study. *Mod Pathol.* 2019.
 36. Harrison BT, Hwang ES, Partridge AH, Thompson AM, Schnitt SJ. Variability in diagnostic threshold for comedo necrosis among breast pathologists: implications for patient eligibility for active surveillance trials of ductal carcinoma in situ. *Mod Pathol.* 2019.
 37. Van Bockstal MR, Agahozo MC, Koppert LB, CHM v D. A retrospective alternative for active surveillance trials for ductal carcinoma in situ of the breast. *Int J Cancer.* 2019.
 38. Sluijter CE, van Lonkhuijzen LR, van Slooten HJ, Nagtegaal ID, Overbeek LI. The effects of implementing synoptic pathology reporting in cancer diagnosis: a systematic review. *Virchows Arch.* 2016;468(6):639–49.
 39. PALGA Foundation (2018) Annual report. <https://www.palga.nl/assets/uploads/Jaarverslag/2018%20Palga%20web.pdf>. Accessed December 4 2019.
 40. van Luijt PA, Heijnsdijk EA, Fracheboud J, Overbeek LI, Broeders MJ, Wesseling J, et al. The distribution of ductal carcinoma in situ (DCIS) grade in 4232 women and its impact on overdiagnosis in breast cancer screening. *Breast Cancer Res.* 2016;18(1):47.
 41. Cutuli B, Lemanski C, De Lafontan B, Chauvet MP, De Lara CT, Mege A, et al. Ductal carcinoma in situ (DCIS): a French National Survey. Analysis of 2,125 patients. *Clin Breast Cancer.* 2019.
 42. Brehaut JC, Colquhoun HL, Eva KW, Carroll K, Sales A, Michie S, et al. Practice feedback interventions: 15 suggestions for optimizing effectiveness. *Ann Intern Med.* 2016;164(6):435–41.
 43. Kluger AN, DeNisi A. The effects of feedback interventions on performance: a historical review, a meta-analysis, and a preliminary feedback intervention theory. *Psychol Bull.* 1996;119(2):254–84.
 44. Ilgen DR, Fisher CD, Taylor MS. Consequences of individual feedback on behavior in organizations. *J Appl Psychology.* 1979;64(4):349–71.
 45. Archer-Kath J, Johnson DW, Johnson RT. Individual versus group feedback in cooperative groups. *J Soc Psychol.* 1994;134(5):681–94.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Ready to submit your research? Choose BMC and benefit from:

- fast, convenient online submission
- thorough peer review by experienced researchers in your field
- rapid publication on acceptance
- support for research data, including large and complex data types
- gold Open Access which fosters wider collaboration and increased citations
- maximum visibility for your research: over 100M website views per year

At BMC, research is always in progress.

Learn more biomedcentral.com/submissions

