Pathology (April 2019) **51(3)**, pp. 246–252

ANATOMICAL PATHOLOGY

Laboratory validation studies in Ki-67 digital image analysis of breast carcinoma: a pathway to routine quality assurance

Morgan Wang[1], Sally McLaren[1], Roopaa Jeyathevan[1], Benjamin Michael Allanson[1], Amanda Ireland[1], Alexandra Kang[1], Katie Meehan[1,2], Carla Thomas[1,2], Cleo Robinson[1,2], Marais Combrinck[1,3], Jennet Harvey[1,2,3], Greg Sterrett[1,3], Benjamin Dessauvagie[1,2,3]

[1]*Breast Subspecialty Unit, PathWest Laboratory Medicine, QEII Medical Centre and Fiona Stanley Hospital Sites, Perth, WA, Australia;* [2]*School of Biomedical Science, UWA, Crawley, WA, Australia;* [3]*Division of Pathology and Laboratory Medicine, Medical School, UWA, Crawley, WA, Australia*

Summary

Ki-67 proliferative index (PI) has prognostic and predictive value in invasive breast carcinoma (IBC), but clinical uptake has been hampered by suboptimal accuracy, reproducibility and standardisation. Published guidelines have addressed pre-analytical and analytical factors to improve Ki-67 PI utility; however, practicalities of ongoing monitoring of Ki-67 PI quality in IBC reporting have not been established. We aimed to evaluate the internal and external quality of our established digital Ki-67 PI IBC reporting practice at a tertiary institution.

In the 5 years since initial validation work, we've completed a series of internal and external quality assurance (QA) projects: (1) an interobserver agreement study, (2) a two site interlaboratory agreement study, (3) determination of the error of our Ki-67 results, (4) an audit of the year-to-year Ki-67 values, (5) an audit of Ki-67 in neoadjuvant chemotherapy (NAC) treated cases, and (6) comparison of our Ki-67 datasets with similar published datasets.

There was excellent concordance (intra-class correlation = 0.98) and good agreement [kappa (κ) = 0.76–0.96] between pathologists, excellent concordance [Pearson correlation (R) = 0.94] and very good agreement (κ = 0.80) between laboratories and excellent concordance (R = 0.92–0.95) and good agreement (κ = 0.67–1.0) over time for our Ki-67 results. No significant difference was observed in Ki-67 data from year-to-year. Expected associations with clinico-pathological prognosticators, pathological complete response following NAC and mitotic index were evident. The median Ki-67 values from the overall and NAC treated datasets were within the range reported in other studies, and our data could be separated into similarly proportioned 'high' and 'low' Ki-67 PI groups when dichotomised as per protocols in other studies.

Collectively, our work provides evidence of adequate internal and external quality control for our digital Ki-67 PI IBC reporting protocols. Given the paucity of formal Ki-67 QA programs, our approach could be emulated, and results compared between laboratories as a framework for internal and external Ki-67 QA.

Key words: Ki-67; digital image analysis; breast cancer.

Received 26 August, revised 15 November, accepted 2 December 2018
Available online 6 March 2019

INTRODUCTION

Ki-67 is a nuclear antigen expressed in all phases of cell proliferation and detectable by a readily available, rapid and affordable immunohistochemical assay. There is strong evidence that 'high' levels of the Ki-67 proliferative index (PI) are associated with poor disease-free and overall survival in invasive breast cancer (IBC)[1–5] and have been shown to indicate the potential value of (neo)adjuvant chemotherapy in certain subsets of patients.[4,6,7] In the era of personalised medicine, accurate and precise evaluation of tissue-based diagnostics and biomarker expression is imperative. The widespread clinical uptake of Ki-67 assessment has been hampered by issues of accuracy, reproducibility and standardisation.[8,9] Specifically, there are issues with Ki-67 PI measurement [including visual assessment vs digital image analysis (DIA)], interpretation (lack of consensus regarding the cut-off value between 'low' and 'high' Ki-67 PI) and selection of the Ki-67 antibody used (different Ki-67 antibody clones produce subtly different Ki-67 PI values).[8,10,11]

An International Ki-67 in Breast Cancer Working Group recently devised a set of recommendations addressing pre-analytical, analytical and interpretation factors in the hope of standardising practice, improving accuracy and reducing variability.[8] However, issues of ongoing quality assurance (QA) for laboratories engaged in Ki-67 reporting were not addressed by the working group and are not well established.

This is in contrast to the routinely reported receptor biomarkers in IBC: oestrogen receptor (ER), progesterone receptors (PR) and human epidermal growth factor receptor-2 (HER2). Clear guidelines regarding QA for these biomarkers and established external QA programs (QAP) and accreditation programs are available. QA is imperative for the reliability, accuracy and reproducibility of results and to ensure optimal patient management and safety, particularly if

Print ISSN 0031-3025/Online ISSN 1465-3931 Crown Copyright © 2019 Published by Elsevier B.V. on behalf of Royal College of Pathologists of Australasia. All rights reserved.
DOI: https://doi.org/10.1016/j.pathol.2018.12.416

therapeutic decisions are to be guided by the biomarker result. Clearly, guidelines for Ki-67 QA are required.

In a pilot study of 86 cases, the Anatomical Pathology (AP) department at PathWest QEII Medical Centre (QEII) previously demonstrated that digital image analysis (DIA) is an accurate, reproducible and efficient method for routine evaluation of Ki-67 PI in IBC in our laboratory.[9] The study also served in training of pathologists and scientists in Ki-67 digital image analysis (DIA) and development of laboratory work-flow for Ki-67 analyses.[9] Given the success of this project, we subsequently commenced reporting digitally assessed Ki-67 PI as a continuous variable in all IBC excision specimens in 2013.

In this context we aimed to evaluate the quality of the established digital Ki-67 PI IBC reporting practice at our busy tertiary institution over a 5 year period and to present methods of Ki-67 QA evaluation which could be emulated in other pathology institutions as part of a framework for internal and external Ki-67 QA.

MATERIALS AND METHODS

Local institutional approvals were obtained for all components of this study.

Ki-67 staining

Ki-67 staining was performed at two PathWest laboratories, using different Ki-67 antibody clones, and slightly different protocols. At PathWest QEII, 4 µm sections were heat retrieved in Ventana CCI buffer at 95°C for 64 min and incubated with Dako M7240 Ki-67 antibody clone for 60 min at 1/400 dilution on a Ventana Ultra automated platform. At PathWest Fiona Stanley Hospital (FSH), 4 µm sections were heat retrieved in Ventana CCI buffer at 95°C for 36 min and incubated with un-dilute Ventana 30-9 Ki-67 antibody clone for 8 min also on the Ventana Ultra platform. At both sites, antibody binding was detected by DAB chromogen and visualised by the Ultraview DAB detection kit with haematoxylin counterstain. Control lymph node tissue was included in each run.

Ki-67 digital image analysis

Ki-67 slides were scanned at ×200 magnification using Aperio ScanScope AT and viewed using Aperio ImageScope software (Leica, Germany). A minimum of 500 cells in the area of highest proliferation (the 'hot spot') was selected for analysis by the Aperio IHC nuclear algorithm (version 1, at factory pre-set values). The region for evaluation and the digital mark-up were always viewed onscreen by a pathologist before accepting Ki-67 PI results. The algorithm is designed to exclude populations of small cells (such as lymphocytes) and spindled cells, and to quantify positive nuclear staining as a percentage of all tumour nuclei (Fig. 1).

Interobserver agreement

Ki-67 PI was performed independently by four pathologists in a series of 20 consecutive IBC excisions from March 2015 at QEII. Pearson correlation coefficient (R) and Kappa agreement (κ) ('high'/'low' at ≥10%) was calculated between the Ki-67 PI values between each pair of pathologists and between each pathologist and the original reported Ki-67 PI and a four way intra-class correlation was determined overall between the four pathologists. A p value of <0.05 was considered significant in this and all subsequent statistical analyses.

Interlaboratory agreement

Ki-67 PI was performed independently at two separate laboratory sites within PathWest in 40 IBC excisions from January 2017; 20 consecutive cases at QEII and 20 consecutive cases at FSH. Two sections were cut from each case; one stained for Ki-67 locally, and the other stained for Ki-67 at the alternate laboratory. Pearson correlation coefficient and Kappa agreement ('high'/'low' at ≥10%) was calculated between the Ki-67 PI values determined at each site.

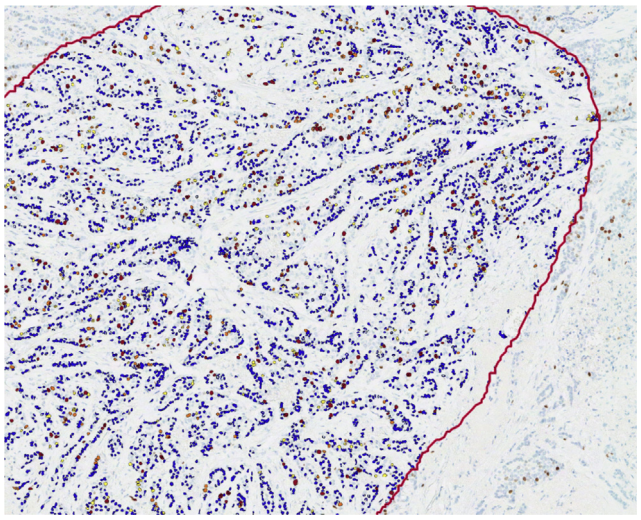

Fig. 1 High power magnification of a Ki-67 stained invasive breast carcinoma demonstrating the digital image analysis mark-up; positive tumour cells mark as yellow, orange or brown depending on intensity; negative tumour cells as blue. Quantification of Ki-67 as a percentage of tumour cells occurs within seconds and is displayed on screen.

Measurement of error

Uncertainty of measurement was determined separately for the inter-observer and the inter-laboratory agreement data, restricted to cases with a Ki-67 score ≤20% (we specifically examined error below 20% as most proposed cut-off values for Ki-67 prognostication are below 20%[1]). Measurement of error was determined by multiplying the standard error in each case by the z-score for a confidence level of 95%.

Audit of Ki-67 PI results

The formal diagnostic reports from all IBC excisions between 2013 and 2016 at PathWest QEII were retrieved via interrogation of the laboratory information system. The Ki-67 PI and standard clinico-pathological parameters were retrospectively extracted. Standard descriptive statistics [mean, median and confidence interval (CI)] were calculated for all continuous variables. Correlations with clinico-pathological variables were also performed after classification of tumours into 'low' and 'high' Ki-67 groups at a cut-off of ≥10%. Mean Ki-67 scores were compared using t-tests for two variables or one-way analysis of variance (Classic or Welch's if unequal variance) for three or more variables. Correlation between continuous variables was determined with Pearson correlation coefficient. Binary outcome proportions were compared using Pearson's Chi-square test.

Audit of pre-treatment Ki-67 PI in neoadjuvant chemotherapy treated patients

Ki-67 PI was retrospectively performed on all core biopsies in the PathWest FSH database between 2013 and 2016 of IBCs subsequently treated with neoadjuvant chemotherapy (NAC) with a matched excision sample in file. The degree of treatment response in the excision specimen was extracted from pathology reports. A Student's t-test was performed to compare Ki-67 values between pathological complete response (pCR) and non-pCR groups.

Comparison of our Ki-67 PI results to external data

Descriptive statistics derived from the QEII Ki-67 PI audit results were compared to similar Ki-67 datasets published in the literature. Studies for comparison were included if they reported at least one of Ki-67 PI median, mean or standard deviation, involved at least 200 cases and Ki-67 assessment was performed on whole-face sections from excisional biopsies of IBC.

Descriptive statistics derived from the evaluation of pre-treatment Ki-67 PI in NAC treated IBCs were compared to similar Ki-67 datasets published in the literature. Studies for comparison were included if they investigated the relationship between Ki-67 PI and response to NAC in IBC and reported at least one of Ki-67 PI median, mean or standard deviation.

Table 1 Correlations of Ki-67 proliferative index between four pathologists and as well as the originally reported Ki-67 proliferative index

	Pathologist 1	Pathologist 2	Pathologist 3	Pathologist 4	Original report
Pathologist 1	1				
Pathologist 2	0.91	1			
Pathologist 3	0.92	0.97	1		
Pathologist 4	0.97	0.95	0.96	1	
Original report	0.95	0.92	0.93	0.93	1

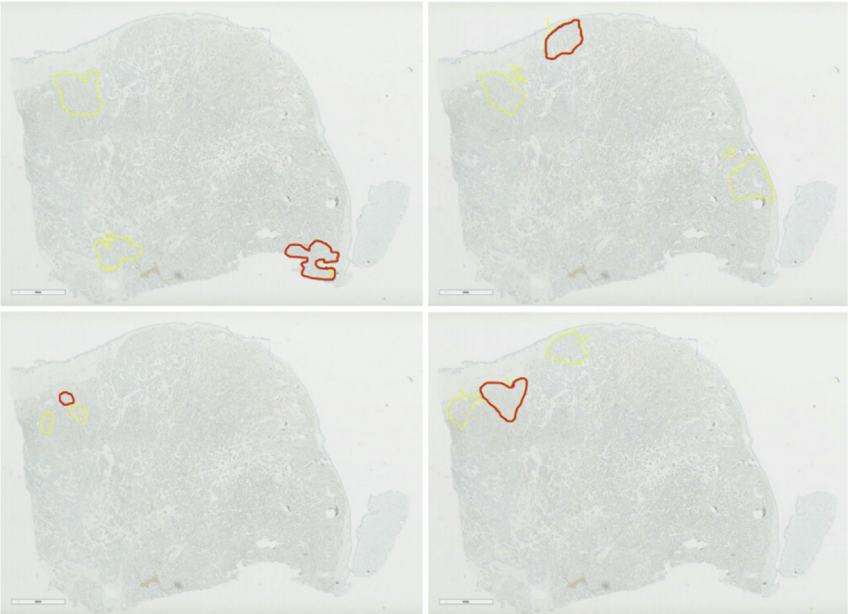

Fig. 2 Scanning magnification of a Ki-67 stained invasive breast carcinoma demonstrating the spatial variation in the location of the hotspots for evaluation (outlined in red) by the four different pathologists.

RESULTS

Reproducibility, agreement and error

We demonstrated excellent concordance (ICC = 0.98) and good agreement (κ = 0.76–0.96) between all pathologists and excellent concordance (R = 0.92–0.95) and good agreement (κ = 0.67–1.0) across time between repeated Ki-67 PI value and the originally reported value (Table 1). These results were achieved despite the area of analysis selected by the different pathologists being spatially separate from one-another and/or different in size in some cases (Fig. 2). The range in error of measurement for the interobserver data was 0.3–3.5% with a mean of 1.8%.

There was excellent concordance (R = 0.94) and very good agreement (κ = 0.80) of Ki-67 PI values between the two laboratory sites (Fig. 3 and 4), again despite the blinded selection of regions for analysis. The range in error of measurement for the inter-laboratory data was 0.98–3.9% with a mean of 1.8%.

Audit of Ki-67 PI results

A total of 1395 IBCs were included in the audit. Size and ER/PR status was reported in all tumours, with mitotic count reported in 1374 tumours, grade in 1391 and HER2 status in 1380. Accompanying lymph nodes were received for 1313 tumours.

The distribution of Ki-67 PI values was stable between 2013 and 2016, with no significant difference between each year ($p > 0.05$) (Table 2). Ki-67 PI correlated highly with mitotic index (MI) ($r = 0.7$, $p < 0.001$) (Fig. 5) and showed significant positive correlations with tumour grade and nodal involvement. Elevated Ki-67 was also associated with HER2 positivity and loss of hormone receptor expression. The 'triple negative' immunophenotype (ER/PR/HER2 negative) was associated with the highest Ki-67 PI (Table 3).

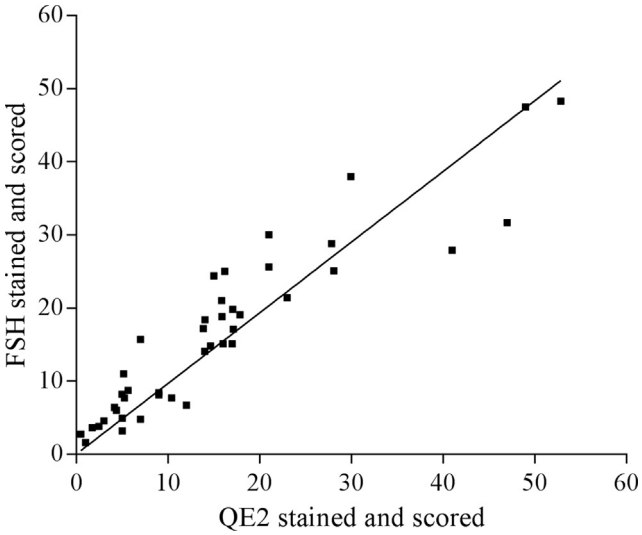

Fig. 3 Cross site correlation of Ki-67 proliferative index values obtained by local staining and reporting protocols.

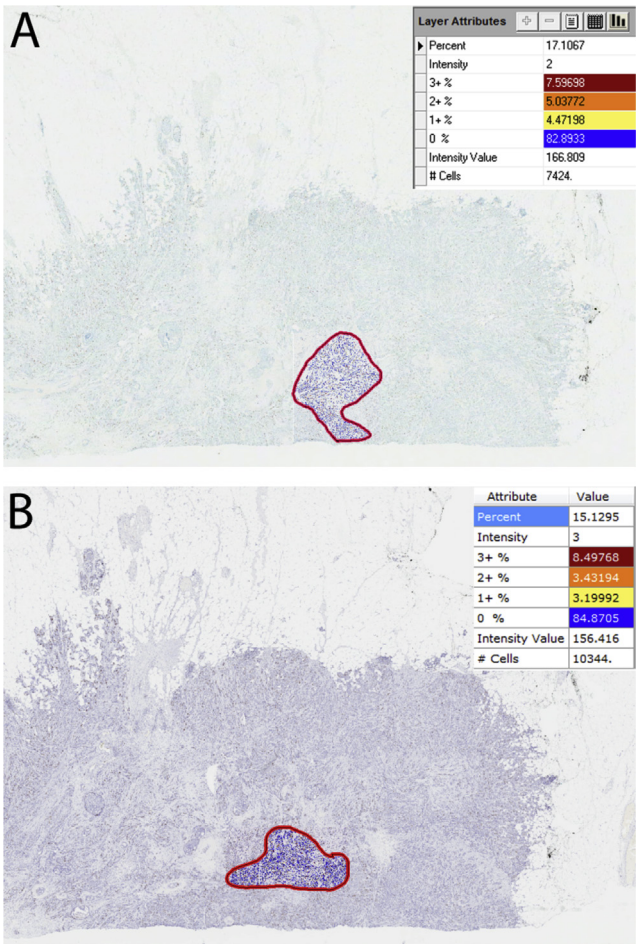

Fig. 4 Cross site Ki-67 proliferative index (PI) digital image analysis. (A,B) Low power images of the Ki-67 stained slides from same case prepared and scored at QEII site (A) and FSH site (B); note that a similar area of the slide was selected for analysis and similar Ki-67 PI values were obtained (insets).

Audit of pre-treatment Ki-67 PI in neoadjuvant chemotherapy treated patients

Ninety-five core biopsies were identified according to our inclusion and exclusion criteria, and of these, 19 cases achieved a pCR (20%). The pre-treatment Ki-67 PI ranged from 2 to 81% across the cohort.

The mean and median pre-treatment Ki-67 PI in IBCs with pCR was 36.4% and 38.0% respectively, compared to 29.4% and 28.5% in incomplete responders (Table 4), identifying a non-significant trend toward a higher Ki-67 PI in the pCR versus non-pCR group ($p = 0.11$) (Fig. 6).

Comparison of Ki-67 results to published data

Eleven studies from the literature met inclusion criteria for comparison (Table 5).[2,12–21] Our median and standard deviation (SD) were towards the lower end of the range reported in the literature, while our mean fell below the reported range. Using a cut-off of ≥10% there was a split of the Nottingham Grade 2 tumours in our cohort into two approximately equal size groups (53% 'high' vs 47% 'low'), similar to the Grade 2 split (49.8% 'high' vs 50.2% 'low'; Table 5) observed in a study by Aleskandarany *et al.*[13] using the same cut-off value.

Four studies from the literature met inclusion criteria for comparison of Ki-67 PI in NAC (Table 6).[22–25] The median pre-treatment Ki-67 PI values of complete responders versus non-complete responders (38.0% vs 28.5%) was similar to median Ki-67 PI values reported for pCR and non-pCR literature (40% vs 20%), while the mean values fell just below the reported range. Employing a cut-off of ≥35%, we demonstrated 31% (12/39) of IBC with pCR in IBCs in the 'high' pre-treatment Ki-67 PI group compared to only 12.5% (7/56) in the 'low' Ki-67 PI group, similar to the 29% vs 17% split observed in a recent study by Denkert *et al.*[26] using the same cut-off value.

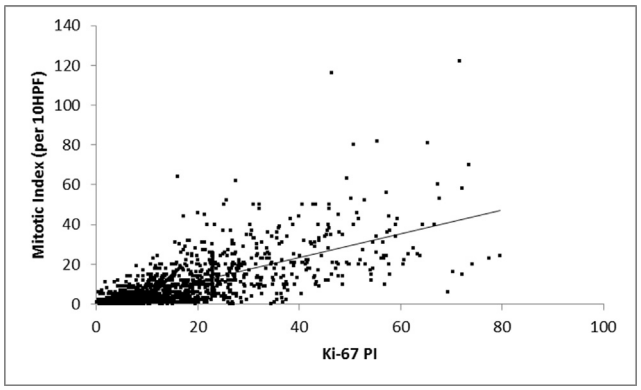

Fig. 5 Correlation between Ki-67 proliferative index and mitotic index.

DISCUSSION

Assessment of Ki-67 PI is a readily available, rapid and affordable ancillary test, with demonstrable prognostic and predictive power. The potential for Ki-67 PI to influence management of IBC is exciting, but its assessment requires robust solutions to the ongoing problems of standardisation, accuracy, reproducibility and quality control. Our practice is to adhere to guidelines set out by the International Ki-67 in Breast Cancer Working Group,[8] whereby specimen handling and slide preparation abide by basic pathology QA processes (such as fixation, specimen storage and immunohistochemistry protocols) and by their more specific recommendations relating to Ki-67 (such as choice of Ki-67 clone and interpretation of positive staining). We also participate in 'NordiQC' one of the few externally assessed

Table 2 Ki-67 proliferative index descriptive statistics by year

	Overall	2013	2014	2015	2016	*p* value
N	1395	305	398	277	415	0.5317
Mean Ki-67 (95% CI)	16.5 (15.7,17.2)	15.7 (14.1,17.3)	16.1 (14.7,17.5)	17.2 (15.5,18.9)	16.8 (15.5,18.1)	
Median Ki-67	12.3	11.1	12.0	12.6	12.9	
Standard deviation	13.9	13.7	13.9	14.3	13.8	

Table 3 Ki-67 proliferative index and clinicopathological parameters

	Mean Ki-67 (95%CI)	*p* value	Ki-67 <10%	Ki-67 ≥10%	*p* value
Number	1395		569 (41%)	826 (59%)	
Grade		<0.0001			<0.0001
1	7.0 (6.4, 7.6)		215 (82%)	48 (18%)	
2	12.0 (11.4, 12.6)		338 (47%)	377 (53%)	
3	30.3 (28.8, 31.8)		14 (3%)	399 (97%)	
ER status		<0.0001			<0.0001
Positive	14.0 (13.3, 14.6)		562 (46%)	661 (54%)	
Negative	34.1 (31.5, 36.7)		7 (4%)	165 (96%)	
HER2 status		<0.0001			<0.0001
Positive	22.5 (20.1, 24.9)		21 (16%)	112 (84%)	
Negative	15.7 (15.0,16.5)		543 (44%)	704 (56%)	
Triple negative status		<0.0001			<0.0001
Yes	37.1 (33.8,40.4)		5 (4%)	109 (96%)	
No	16.1 (15.4, 16.8)		564 (41%)	820 (59%)	
Nodal status		<0.0001			<0.0001
Positive	19.2 (17.9, 20.6)		123 (28%)	309 (72%)	
Negative	15.3 (14.4, 16.3)		405 (46%)	477 (54%)	

Table 4 Total number of cases stratified according to pCR and non-pCR with descriptive statistics

Group	N	Mean Ki-67	Median Ki-67	SD
pCR	19	36.42	38	16.3
non-pCR	76	29.49	28.5	16.9

pCR, pathological complete response.

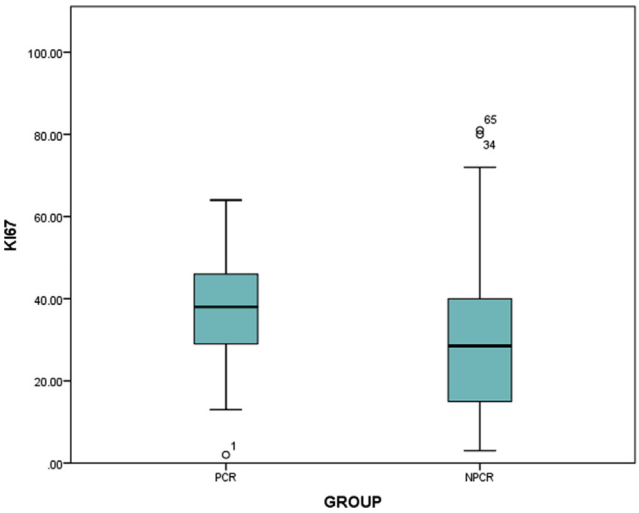

Fig. 6 Box and whisker plot showing a trend toward a higher Ki-67 proliferative index in pCR (left) versus non-pCR (right) groups.

immunohistochemistry proficiency testing programs which includes a breast cancer Ki-67 module. However, unaddressed issues of internal and external quality control abound, and in contrast to the routine receptor biomarkers in breast cancer (ER, PR and HER2), little has been written to address issues of ongoing QA for laboratories engaged in Ki-67 reporting. Over a series of projects documented herein, we devised methods to assess both internal and external QA and in doing so doing have demonstrated that our digital Ki-67 PI in IBC reporting practice meets excellent internal and external QA benchmarks.

We have demonstrated excellent reproducibility and low measurement of error in the digital quantification of Ki-67 PI between pathologists at one site, between pathologists across sites and between pathologists across two time points. This was achieved despite the independent, subjective selection of the region for analysis and routine handling and staining protocols at the two sites (which are subtly different). These results are possible due to (1) adherence at local sites to basic pathology QA processes, such as fixation protocols and use of positive/negative controls; (2) participation in externally assessed immunohistochemistry proficiency testing (RCPA-QAP, NordiQC) which lead to optimal Ki-67 stained slides for analysis; and (3) adherence to basic tenets established in our prior Ki-67 work including assessing areas of highest proliferation (the 'hot spot'), excluding degenerate tumour, benign elements and areas of abundant stroma, and counting between 1000 and 8000 tumour cells.[9,27]

Determining our measurement of error not only supported our assertion that the Ki-67 PI method is robust, accurate and internally valid, but when conveyed to clinicians, it serves as a reminder that Ki-67 values, like any quantifiable laboratory value, are not perfect. Our result indicates that the Ki-67 PI value determined in our laboratory may be out by an average of 1.8%. By providing this error information in the report we remind our clinicians to exercise caution when interpreting Ki-67 PI for decision making purposes at values close to any given cut-off value.

Internal audit revealed the distribution of Ki-67 PI values was stable from 2013 to 2016, with no significant difference in Ki-67 PI values between each year. Elevated Ki-67 PI showed a highly significant positive association with standard, validated, prognostic clinico-pathological variables of mitotic count, grade and nodal status. Elevated Ki-67 was also associated with HER2 positivity, the 'triple negative' phenotype and pCR. Internal reproducibility, small error of measurement, stability over time and expected correlations with clinico-pathological features are components of internal QA and help to confirm the internal consistency in our Ki-67 PI reporting practice.

There is a paucity of formal external QA programs for Ki-67 PI assessment in IBC. In their absence we sought to compare our Ki-67 PI results to datasets from cohorts

Table 5 Comparison of descriptive statistics for Ki-67 PI in invasive breast carcinoma published in the literature

Study	N	Median Ki-67	Mean Ki-67	SD
Inwald *et al.*[2]	3658	15	20.3	18.1
Alco *et al.*[12]	462	20		
Aleskandarany *et al.*[13]	1550	15		
Chen *et al.*[14]	276	25	29.1	22
Fasanella *et al.*[15]	315		22[a]	18[a]
			31[b]	19[b]
Jalava *et al.*[16]	265	17	19.4	12.5
Kim *et al.*[17]	310	10	21.2	21.5
Klintman *et al.*[18]	200	10		
Knutsvik *et al.*[19]	534	18	24	
Tashima *et al.*[20]	4328	20.5	26.2	
Vörös *et al.*[21]	1709	17	23.4	21
Current study	1395	12.3	16.5	13.9

[a] Using SP1 Ki-67 antibody.
[b] Using MM1 Ki-67 antibody.

published in the literature. Our overall Ki-67 PI median and SD are within the range reported in 11 comparable studies in the literature, while the mean lies just below the range. The median pre-treatment Ki-67 PI values in core biopsies of NAC complete responders versus non-complete responders and the mean in non-complete responders compare favourably to similar datasets in the few published studies, while the mean in the pCR group lies just outside the range.

We postulate several reasons as to why our Ki-67 PI mean and median are toward the lower end of the values reported in the literature. Firstly, our overall Ki-67 cohort is unselected, but includes a high proportion of screen-detected cases likely to include smaller, lower grade IBCs, with intrinsic bias towards lower Ki-67 PI, compared to cohorts in some published studies. Additionally, all but one study[15] for comparison utilised visual/manual methods of Ki-67 PI evaluation, which may give higher figures for Ki-67,[8,28,29] in part because of the lower numbers of cells evaluated, typically 500–1000 cells, as opposed to DIA, where the mean number of cells counted can be in excess of 30,000.[9] We have previously demonstrated that a progressive increase in the number of cells for Ki-67 evaluation results in a progressive decrease in the Ki-67 PI of a particular case.[27] We also accept that DIA can incorrectly label negatively stained lymphocytes and stromal cells as tumour cells, artificially lowering the tumour Ki-67 PI. To mitigate this risk, we deliberately avoid selection of stroma and lymphocyte rich areas for our Ki-67 PI DIA. Collectively these factors highlight a need for Ki-67 PI summary statistics from large, screening based cohorts utilising DIA, to enable institutional comparisons for external QA purposes. It also re-iterates the view of the International Ki-67 Working Group that local Ki-67 evaluation methods need to be validated against outcome data in local populations.[8] Such a study is currently underway at our institution.

We undertook further comparisons with external studies to compare proportions of Ki-67 'high' and 'low' cases. A Ki-67 PI cut-off of ≥10% divided the Grade 2 tumours in our cohort into two approximately equal size groups, similar to the Grade 2 split observed in a study by Aleskandarany *et al.*[13] using the same cut-off. In their study a Ki-67 PI of ≥10% selected a 'low risk' group of Grade 2 tumours with biological behaviour akin to Grade 1 tumours. Although of smaller numbers, when our NAC treated cases were separated into 'high' and 'low' Ki-67 PI based on a cut-off of ≥35%, the subsequent pCR rates were similar to those observed by Denkert *et al.*[26] when using the same cut-off value.

The results reported here compare favourably with the descriptive statistics and the proportions of Ki-67 PI 'high' and 'low' cases, particularly in our overall Ki-67 PI datasets. They serve as a measure of external QA and indicate our reporting practice is comparable to external sources.

QA in the reporting of any clinically significant biomarker is required to ensure good practice and patient safety; Ki-67 PI is no different. The investigations described herein arose as solutions to the pitfalls and problems we have encountered in our routine Ki-67 PI reporting practice. These investigations recapitulate components of formal biomarker QA programs, such as at RCPAQAP. In these programs, systematic laboratory errors can be identified by the mandated reporting of a laboratory's data, such as proportions of hormone and HER2 positive/negative breast cases, against expected targets. Outlier laboratories can be identified and root cause analysis instigated. We believe the comparison of our results internally against known clinicopathological associations, and externally against the metrics of published data sets, has recapitulated this QA objective. We anticipate that these processes will be accelerated by other laboratories adopting similar approaches. Interlaboratory comparisons and/or formalised third party external Ki-67 QA programs will lead to robust QA, detection of outlier results and further contribute to agreed/common standards.

In summary, our interobserver reproducibility between pathologists, across laboratory sites and across time for Ki-67 PI is excellent, our Ki-67 PI error measurement is low and our yearly Ki-67 PI values have remained stable over a 4 year period, with the expected positive associations with clinicopathological parameters, collectively reflecting excellent internal QA. Our median and SD for Ki-67 PI and proportions

Table 6 Comparison of descriptive statistics for pre-treatment Ki-67 PI in neoadjuvant chemotherapy in invasive breast carcinoma published in the literature

Study	N	Median Ki-67		Mean Ki-67	
		pCR	Non pCR	pCR	Non pCR
Fasching *et al.*[22]	552			50.6	26.7
Pohl *et al.*[23]	65	40	20		
Sinn *et al.*[24]	54			45	30
Yoshioka *et al.*[25]	64			50.7	33.3
Current study	95	38.0	28.5	36.4	29.4

pCR, pathological complete response.

of Ki-67 'high' and 'low' cases are similar to published datasets. Favourable comparisons to other studies are a measure of good external QA. Ideally, results require validation against local outcome data and in formal external QA programs. Our studies provide evidence of adequate internal and external QA for our digital Ki-67 PI IBC reporting service. Given the paucity of formal Ki-67 QA programs, our approach could be emulated and results compared between laboratories as a framework for local internal and external Ki-67 QA.

Acknowledgements: The authors acknowledge the facilities and the scientific and technical assistance of the Australian Microscopy & Microanalysis Research Facility at the Centre for Microscopy, Characterisation & Analysis, UWA.

Conflicts of interest and sources of funding: The Australian Microscopy & Microanalysis Research Facility at the Centre for Microscopy, Characterisation & Analysis, UWA, is funded by the University, State and Commonwealth Governments. Components of this research were carried out during BD's tenure of a Suzanne Cavanagh Early Career Investigator Grant from Cancer Council Western Australia. The authors state that there are no conflicts of interest to disclose.

Address for correspondence: Ben Dessauvagie, Division of Pathology and Laboratory Medicine, Medical School, UWA, Crawley, WA 6009, Australia. E-mail: ben.dessauvagie@health.wa.gov.au

References

1. de Azambuja E, Cardoso F, de Castro Jr G, *et al*. Ki-67 as prognostic marker in early breast cancer: a meta-analysis of published studies involving 12,155 patients. *Br J Cancer* 2007; 96: 1504–13.
2. Inwald EC, Klinkhammer-Schalke M, Hofstadter F, *et al*. Ki-67 is a prognostic parameter in breast cancer patients: results of a large population-based cohort of a cancer registry. *Breast Cancer Res Treat* 2013; 139: 539–52.
3. Kontzoglou K, Palla V, Karaolanis G, *et al*. Correlation between Ki67 and breast cancer prognosis. *Oncology* 2013; 84: 219–25.
4. Luporsi E, Andre F, Spyratos F, *et al*. Ki-67: level of evidence and methodological considerations for its role in the clinical management of breast cancer: analytical and critical review. *Breast Cancer Res Treat* 2012; 132: 895–915.
5. Stuart-Harris R, Caldas C, Pinder SE, *et al*. Proliferation markers and survival in early breast cancer: a systematic review and meta-analysis of 85 studies in 32,825 patients. *Breast* 2008; 17: 323–34.
6. Criscitiello C, Disalvatore D, De Laurentiis M, *et al*. High Ki-67 score is indicative of a greater benefit from adjuvant chemotherapy when added to endocrine therapy in luminal B HER2 negative and node-positive breast cancer. *Breast* 2014; 23: 69–75.
7. Chen X, He C, Han D, *et al*. The predictive value of Ki-67 before neoadjuvant chemotherapy for breast cancer: a systematic review and meta-analysis. *Future Oncol* 2017; 13: 843–57.
8. Dowsett M, Nielsen TO, A'Hern R, *et al*. Assessment of Ki67 in breast cancer: recommendations from the International Ki67 in breast cancer working group. *J Natl Cancer Inst* 2011; 103: 1656–64.
9. Harvey J, Thomas C, Wood B, *et al*. Practical issues concerning the implementation of Ki-67 proliferative index measurement in breast cancer reporting. *Pathology* 2015; 47: 13–20.
10. Jing N, Fang C, Williams DS. Validity and reliability of Ki-67 assessment in oestrogen receptor positive breast cancer. *Pathology* 2017; 49: 371–8.
11. Gudlaugsson E, Skaland I, Janssen EA, *et al*. Comparison of the effect of different techniques for measurement of Ki67 proliferation on reproducibility and prognosis prediction accuracy in breast cancer. *Histopathology* 2012; 61: 1134–44.
12. Alco G, Bozdogan A, Selamoglu D, *et al*. Clinical and histopathological factors associated with Ki-67 expression in breast cancer patients. *Oncol Lett* 2015; 9: 1046–54.
13. Aleskandarany MA, Rakha EA, Macmillan RD, *et al*. MIB1/Ki-67 labelling index can classify grade 2 breast cancer into two clinically distinct subgroups. *Breast Cancer Res Treat* 2011; 127: 591–9.
14. Chen X, Zhu S, Fei X, *et al*. Surgery time interval and molecular subtype may influence Ki67 change after core needle biopsy in breast cancer patients. *BMC Cancer* 2015; 15: 822.
15. Fasanella S, Leonardi E, Cantaloni C, *et al*. Proliferative activity in human breast cancer: ki-67 automated evaluation and the influence of different Ki-67 equivalent antibodies. *Diagn Pathol* 2011; 6 (Suppl 1): S7.
16. Jalava P, Kuopio T, Juntti-Patinen L, *et al*. Ki67 immunohistochemistry: a valuable marker in prognostication but with a risk of misclassification: proliferation subgroups formed based on Ki67 immunoreactivity and standardized mitotic index. *Histopathology* 2006; 48: 674–82.
17. Kim HS, Park S, Koo JS, *et al*. Risk factors associated with discordant Ki-67 levels between preoperative biopsy and postoperative surgical specimens in breast cancers. *PLoS One* 2016; 11: e0151054.
18. Klintman M, Bendahl PO, Grabau D, *et al*. The prognostic value of Ki67 is dependent on estrogen receptor status and histological grade in premenopausal patients with node-negative breast cancer. *Mod Pathol* 2010; 23: 251–9.
19. Knutsvik G, Stefansson IM, Aziz S, *et al*. Evaluation of Ki67 expression across distinct categories of breast cancer specimens: a population-based study of matched surgical specimens, core needle biopsies and tissue microarrays. *PLoS One* 2014; 9: e112121.
20. Tashima R, Nishimura R, Osako T, *et al*. Evaluation of an optimal cut-off point for the Ki-67 index as a prognostic factor in primary breast cancer: a retrospective study. *PLoS One* 2015; 10: e0119565.
21. Voros A, Csorgo E, Kovari B, *et al*. The use of digital images improves reproducibility of the ki-67 labeling index as a proliferation marker in breast cancer. *Pathol Oncol Res* 2014; 20: 391–7.
22. Fasching PA, Heusinger K, Haeberle L, *et al*. Ki67, chemotherapy response, and prognosis in breast cancer patients receiving neoadjuvant treatment. *BMC Cancer* 2011; 11: 486.
23. Pohl G, Rudas M, Taucher S, *et al*. Expression of cell cycle regulatory proteins in breast carcinomas before and after preoperative chemotherapy. *Breast Cancer Res Treat* 2003; 78: 97–103.
24. Sinn HP, Schneeweiss A, Keller M, *et al*. Comparison of immunohistochemistry with PCR for assessment of ER, PR, and Ki-67 and prediction of pathological complete response in breast cancer. *BMC Cancer* 2017; 17: 124.
25. Yoshioka T, Hosoda M, Yamamoto M, *et al*. Prognostic significance of pathologic complete response and Ki67 expression after neoadjuvant chemotherapy in breast cancer. *Breast Cancer* 2015; 22: 185–91.
26. Denkert C, Loibl S, Muller BM, *et al*. Ki67 levels as predictive and prognostic parameters in pretherapeutic breast cancer core biopsies: a translational investigation in the neoadjuvant GeparTrio trial. *Ann Oncol* 2013; 24: 2786–93.
27. Dessauvagie BF, Meehan K, Thomas C, *et al*. Digital evaluation of proliferative 'hotspots' of more than 8000 cells negatively affects Ki-67 accuracy in breast carcinoma. *Pathology* 2018; 50: S69.
28. Cuzick J, Dowsett M, Pineda S, *et al*. Prognostic value of a combined estrogen receptor, progesterone receptor, Ki-67, and human epidermal growth factor receptor 2 immunohistochemical score and comparison with the Genomic Health recurrence score in early breast cancer. *J Clin Oncol* 2011; 29: 4273–8.
29. Healey MA, Hirko KA, Beck AH, *et al*. Assessment of Ki67 expression for breast cancer subtype classification and prognosis in the Nurses' Health Study. *Breast Cancer Res Treat* 2017; 166: 613–22.