Title page

MMED research project: Eunice Joy van den Berg

Student number: 8321954

Submitted for Graduation: 21 November 2019

Title:

KI67 immunohistochemistry quantification in breast carcinoma: A comparison of visual estimation, counting and Immunoratio©.

Prepared in submissible format according to the requirements for the journal: "Applied immunohistochemistry and molecular morphology".

Authors:

1. Eunice Joy van den Berg FCPath(SA)Anat National Health Laboratory Service and University of the Witwatersrand Johannesburg

2. Raquel Duarte PhD Department of Internal Medicine. University of the Witwatersrand.

3. Caroline Dickens PhD Department of Internal Medicine University of the Witwatersrand.

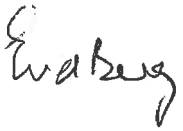
4. Maureen Joffe PhD Chris Hani Baragwanath Academic Hospital Breast Clinic.

5. Reena Mohanlal FCPath(SA)Anat National Health Laboratory Service and University of the Witwatersrand.

CONTENTS

DECLARATION BY CANDIDATE

I, Eunice Joy van den Berg, Student no 8321954, declare that this research project entitled "Ki67 immunohistochemistry quantification in breast carcinoma: a comparison of visual estimation, counting and Immunoratio©" is my own work, and that I have been the main contributor towards the conception and design, literature review, collection and assembly of data, data interpretation and manuscript writing.

Signed:  Date: 21 November 2019

Eunice Joy van den Berg

Acknowledgements:

Dr. H. Cubasch and the breast surgery unit at Chris Hani Baragwanath Hospital and their patients.

Mr. Eric Liebenberg for photography.

Abstract:

Background: Molecular analysis has shown that breast carcinomas can be classified into several intrinsic subtypes, with implications for management and prognosis. In the majority of pathology laboratories molecular analysis of each case is not possible and immunohistochemistry is used for subtyping. This includes analysis of hormone receptors as well as HER-2-neu and Ki67. The methodology for the interpretation of the proliferation index using Ki67 remains an area of uncertainty. We investigated the degree of agreement between different methods of Ki67 interpretation.

Materials and methods: We analysed 204 breast core biopsies diagnostic of breast carcinoma using visual estimation/eyeballing, Immunoratio©, and counting by two pathologists. The correlation between the different methods and the inter-observer agreement between the two pathologists was assessed. Specific analysis was also done with respect to classification of cases into low Ki67 groups (using Ki67 values less than 14% and less than 20%) since this is critical in classifying tumours into Luminal A and Luminal B subtypes.

Results: Correlation between the different methods was best achieved comparing Immunoratio© and counting, and worst comparing counting and eyeballing. Correlation was better when considering inter-observer variability (counting by two pathologists). Comparing the number of cases classified as low Ki67 (less than 14% and less than 20%) the Cohen's Kappa statistic varied from $\kappa = 0.267$ to $\kappa = 0.814$ with different methods. When limiting the analysis to cases with a Ki67 of 10-25% according to any method, there was greater disagreement.

Conclusion: At the higher and lower Ki67 levels, the correlation between the methods of assessment was acceptable, however, at levels close to the cut-off values for Luminal A versus Luminal B, several patients would be differently classified by the different methods and therefore potentially receive suboptimal management.

Key words: Ki67, immunohistochemistry, eyeballing, digital image analysis, Immunoratio©, counting, breast carcinoma.

Abbreviations: IR: Immunoratio©. CP1: Counting by pathologist 1. CP2: Counting by pathologist 2. P1: Pathologist 1. P2: Pathologist 2. EB: Eyeballing. DIA: digital image analysis.

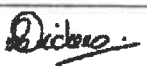
Contribution of Candidate to the submissable article: "Ki67 immunohistochemistry quantification in breast carcinoma: a comparison of visual estimation, counting and Immunoratioⓒ"

This letter confirms that the contribution of E.J. van den Berg to the article includes the following:

1. **Conception and design.**
2. **Literature review.**
3. **Collection and assembly of data.**
4. **Data interpretation.**
5. **Manuscript writing.**

We are satisfied that she may submit this article as her research report for the MMed degree.

Signed by co-authors:

Name	Signature	Date
Caroline Dickens		28/03/2019
Raquel Duarte		28/03/2019
Maureen Joffe		28/03/2019
Reena Mohanlal		28/03/2019

Introduction

Microarray-based molecular analysis of breast carcinomas has shown distinct sub-groups with a spectrum of molecular and biological characteristics. (1) Perou and Sorlie in 2000 used a hierarchical method of clustering of gene expression to divide breast cancers into subtypes that correlate with their morphologic and clinical features. (2) They defined five groups: luminal A, luminal B, normal breast-like, HER2-enriched and basal-like. The "normal breast-like" group may represent an artefact due to insufficient representation of tumour in some of the specimens. The luminal A subtype represents approximately 60% of breast carcinomas, luminal B 10%, HER2-enriched 15% and basal-like 15% (3). Subsequently, other groups have been described. Similar results have been found with genomic analysis (4) and with multiparameter tests including PAM50 (1) and MammaPrint©.(5) This has led to a more personalised approach to therapy.(6)

Employing molecular methods routinely is often not viable in view of the financial cost and time constraints, and the St. Gallen conference of 2013 advised the use of immunohistochemical stains for classification of breast cancers into intrinsic subtypes. (7) This was based on the demonstration by researchers that a panel of immunohistochemical stains could serve as a surrogate for gene expression profiling. One such study was that of Cheang et al. who obtained a 72% sensitivity and 77% specificity using immunohistochemistry compared to gene expression profiling in classification of luminal A and B lesions. (8) Analysis of these stains has been used to serve as a surrogate assessment as follows: (9)

1.Luminal breast cancer

a. Luminal A: ER positive (+) and PR + (PR > 20%), HER2 negative (-) and low Ki67 (<14%).

b. Luminal B: ER+, PR+ (PR>20%), HER2-, Ki67 > 14%

or ER+, HER2-, Ki67 14-19%, and PR < 20%.

2. HER2 Positive

a. Luminal HER2 Subtype: ER+ (but often not as strongly positive as HER2 negative tumours)

b. HER2+-Enriched subtype: ER-, PR-, HER2+

3.Triple negative.

Basal-like subtype: ER-, PR-, HER2-, Cytokeratin5+ and/or EGFR+

Non- classified triple negative subtype: Negative for ER, PR, HER2, Cytokeratin5 and EGFR.

These subtypes are predictive of different responses to chemotherapeutic agents. (10)

During the investigation of specific antigens for diagnosis of lymphoma, the Ki67 antigen was identified as a proliferation marker by Johannes Gerdes. (11) He demonstrated that Ki67 reflects the growth fraction in neoplasms.(12) Prior to this, indirect methods such as ^{3}H thymidine determination and flow cytometry had been used (13), but these could not be applied on a routine basis. After developing the process of antibody cloning MIB1 (Molecular Borstol Index) was produced, which was subsequently used to determine the Ki67 index in formalin fixed paraffin embedded tissue (FFPE). (11) Later the role of the Ki67 protein was further described: to keep the condensed chromosomes dispersed in the cell's cytoplasm during mitosis. (14) The Ki67 antigen is a nuclear non-histone protein expressed in significant quantities throughout the cell cycle, other than in the quiescent phase and early cell cycle and thus quantifies proliferation within a cell population. Anti-Ki67 labelling has been validated as a useful proliferation marker in various neoplasms including haematolymphoid neoplasms, astrocytoma, oligodendroglioma, colon carcinoma and breast carcinoma. (15,16,17)

Many, but not all researchers have demonstrated Ki67's predictive and prognostic value. (16,18) Cheang et al. (8) used a biomarker panel to separate tumour subtypes, with a Ki67 cut off of 14% to classify Luminal A and Luminal B subtypes. On follow-up the latter had a worse prognosis for disease recurrence and for survival. De Azambujo et al. analysed the results of 29 studies. This meta-analysis demonstrated a poorer prognosis for disease-free survival and overall survival in carcinomas with high Ki67 levels, however, a limitation was that univariate analysis only was done. (18). Acs and colleagues demonstrated that the Ki67 index had significance as an independent prognostic factor

for overall survival. (19) A meta-analysis performed by Stuart-Harris et al (20) demonstrated strong correlation between Ki-67 values and overall survival. These researchers did not, however, advise routine use of Ki-67, stating that it was uncertain whether it would add to prognostic indicators already in use.

The Ki-67 index has been shown to be predictive of response to chemotherapy. A study by Cristicitiello et al. demonstrated the heterogeneity of response to adjuvant chemotherapy in oestrogen receptor positive breast carcinomas. They showed that addition of chemotherapy to hormonal therapy was associated with decreased recurrence risk in patients with highly proliferative tumours. (21) Sueta et al. examined luminal-type tumours in the light of their known poorer response to chemotherapy than other subtypes. They confirmed the value of a high proliferation index as a predictor of complete pathological response in Luminal-type tumours with a high Ki-67 index. (22) Patients with luminal A breast cancers have been found to benefit most from hormone therapy without chemotherapy, whereas those with luminal B tumours respond better to chemotherapy. (6) Ellis and colleagues used a multiparameter analysis method including Ki-67, and this proved to be useful in predicting response to neo-adjuvant endocrine therapy. "On-treatment" Ki-67 values demonstrated accuracy in predicting relapse risk. (23)

There is widespread controversy regarding interpretation of Ki67. It has been found to have unacceptable inter-observer and inter-laboratory variability (24) and some suggest that routine reporting of Ki67 should not be done in breast carcinoma in view of its lack of reliability. (25) Many variables impact the interpretation of Ki67 immunostaining. Various methods are used for assessment of percentage of Ki67 positivity, including "eyeballing" (EB) ie. visual scanning and estimating the percentage staining, formally counting the cells, and digital image analysis (DIA) methods. (24) It is uncertain whether the highest staining areas ("hot spots"), average or low staining areas should be assessed, or multiple areas and the mean calculated. Proponents of the

"hot spot" method argue that the "hot spots" are the areas representing the tumour component most likely to metastasize and therefore the most important component prognostically. (26) Assessment of the tumour invasive edge has been recommended. (27) If manual counting of positive cells is done, there is a variable approach to the number of cells counted. The International Ki67 in Breast Cancer Working Group recommended a minimal of 500 cells, and ideally 1000 cells to be counted. (27) One study showed no significant difference between counting 300, 510 or 1020 cells using the standard method (counting three high power fields including "hot" and "cold" spots and determining the average), but obtained a lower median Ki67 index counting 1200 cells than 510 cells if using the "hot spot" or advancing edge approach. (28) The "cut-off" value in assigning tumours to the luminal A or luminal B groups is not universally agreed upon. The ideal value has been shown to vary from one institution to another, ranging from 3,5% to 34%. (18) The St. Gallen International Expert Consensus in 2011 recommended a cut-off of 14%. (7) At the St. Gallen conference of 2015, most experts favoured a value of 20-29%. (29) There are also different thresholds obtained if studies are conducted using different clinical end-points and sub-sets of patients. (30) It is recommended that the Ki67 cut-off value should be determined separately for each laboratory. (25)

Most of the studies assessing different methods of analysis of the Ki67 index have compared DIA to the counting method, and only a few have included the EB technique. One study addressing the methodology in Ki67 determination showed very poor interobserver and intraobserver reproducibility for the "quick-scan" subjective method of analysis. In the same study, using ocular-grid-guided subjective counts, 13% of cases were classified differently between two pathologists using a Ki67 cut off of 15%, and there were disagreement percentages of 6-22% using other values for Ki67 cut off. They also found that assessment using DIA and computerized point-grid-guided interactive morphometry was reproducible and strongly prognostic. (31) In contrast, a study by Varga et al, in which EB and counting were evaluated, showed more reliable results for EB than for counting (32).

Stalhammer and colleagues evaluated manual counting and DIA scores in the categorisation of breast carcinomas into luminal A and luminal B subtypes. They correlated their results with PAM50 assays, high versus low transcriptomic grade, axillary lymph node status and prognostic value. DIA was found to be the most accurate method, especially in hot spots. (33). Klauschen et al. showed good correlation between manual counting and automated Ki67 scores in a study cohort of over 1000 patients, with a Pearson correlation coefficient of 0.89. The Ki67 manual counting in this study was centralised and strict standardisation of the counting method was applied. (34)

Image analysis software is usually not freely available and requires virtual microscopy scanning. ImmunoRatio©(IR) is a DIA system which is accessed via a web browser. It is free and no additional equipment or software are required. (35) The website includes a Camera adjustment wizard to set the contrast and brightness of the images. The analysis has been found to be optimal using three images taken with a 20x objective and averaging the results. (35)

Aim

The aim of this study was:

1. To compare three different methods of quantifying the proliferation index on Ki67 immunohistochemical stains performed on breast core biopsies.
2. To determine the percentages of intrinsic subtypes using the Ki67 values obtained by the different methods.
3. To assess inter-observer variability in counting by two pathologists.

Methods

This was a retrospective non-interventional cross-sectional study.

Two hundred and four cases of breast carcinoma diagnosed on core biopsy were obtained from the cases enrolled in the Chris Hani Baragwanath South African Breast Cancer (SABC) case-control study

(PI, I Romieu and funded by the World Cancer Research Fund), which included a total of 500 cases. These biopsies were obtained between 23/05/2014 and 13/12/2017. Consecutive cases diagnosed with breast carcinoma and reported at the National Health Service, Chris Hani Baragwanath Hospital histopathology laboratory were included. The sample size was determined using the formula recommended by Watson et al. (36).

Cases with insufficient representation of invasive tumour for counting or IR, where a Ki67 stain was not done or reported, where the patient had been treated with chemotherapy or radiotherapy prior to biopsy and biopsies with extensive necrosis or inflammation were excluded. Biopsies from male patients were also excluded.

The original reports as captured on the National Health Laboratory information system (Trakcare) were examined to determine the Ki67s reported on each case, assumed to have been determined using the EB method. Oestrogen receptor positivity in the initial reporting of these cases is defined as more than 1% positivity according to current recommendations (37). The assessment in the original histopathological report was used and was not repeated.

The slides stained with the Ki67 immunostain (antibody 30-9, Roche diagnostic, Ventana, USA; using the Ventana Benchmark XT machine) were retrieved from the archives.

To assess Ki67 by the counting method, photomicrographs were taken of the Ki67 stained biopsies by the first pathologist (P1). Consecutive fields to include obvious "hot spots" were photographed. The images were saved in Microsoft Powerpoint and a grid overlay was applied to facilitate counting. A total of 500 consecutive tumour cells were counted along the grid lines from left to right; this formed the denominator. The point after the 500th cell was marked. The positively stained tumour cells amongst these 500 cells were then counted in the same way, this formed the numerator. The

percentage Ki67 value was determined by dividing the number of positive cells by 500. This process was repeated by the second pathologist (P2) independently, using the same photomicrographs. P2 was blinded to the P1 findings.

The slides were then interpreted using the IR software. Photomicrographs of three areas for each case were analysed and a mean value obtained, using a 20x microscopic objective. An Olympus BH-2 microscope with an attached Zeiss AxiocamERc 5s camera was used. The same images were used for counting and IR, however, in some cases additional images were required for counting in order to obtain 500 cells.

Statistical analysis

All data was captured on an Excel spreadsheet and analysed using Stata software in conjunction with a statistician.

A Bland and Altman diagram for each pair was constructed. The Bland –Altman plot is a graphical method to compare two measurement techniques. The differences obtained between the two measurements are plotted against the average difference between the two techniques. Limits of agreement are calculated which should not exceed the maximum allowed difference in order to assess the two methods as being in agreement.

For each comparison a Lin's concordance coefficient was determined to assess how well the points lay when plotted with a 45° line drawn though the origin. The Lin's correlation coefficient quantifies the agreement between two measures of the same variable. The value ranges between -1 to 1, with perfect agreement at 1.

Comparisons were done of the number of low Ki67 results obtained using cut-off values of 14% and 20% (St. Gallen conference recommended values for 2011 and 2015 respectively). A Cohen's Kappa

statistic was calculated for each of these. The Cohen's Kappa statistic evaluates the agreement between categorical assessments, taking into account the amount of agreement that would be obtained by chance. The maximum is 1 (perfect agreement) and the minimum 0 (no agreement). Since the most important problem is division of the cases into low Ki67 and high Ki67 groups, values within the range of 10 – 25% were also analysed separately.

Results.

Core biopsies were retrieved from female patients newly diagnosed with invasive mammary carcinoma. The patient ages ranged from 27 years to 86 years with a median of 52 years and an interquartile range of 42 – 60 years. All the patients were black Africans.

Two hundred and four Ki67 immunostains were photographed. Three of these were excluded from counting as the cells were extremely crushed in 2 cases and in the other case images with sufficient clarity were not obtained (excluded by P2). See Fig. 1.

The majority (89.5%) of the tumours were morphologically invasive carcinoma of no special type. Mucinous, lobular, micropapillary, carcinoma with medullary features, apocrine, metaplastic and mixed (no special type and micropapillary) subtypes comprised the remainder. See Table 1.
The histological grades were assessed as grade 1 in 8%, grade 2 in 44% and grade 3 in 47%. It was not reported in 2 cases (1%).

According to the initial biopsy reports (using a Ki67 cut-off of 14%), the intrinsic subtypes were 13 (6.5%) luminal A, 143 (70%) luminal B, 17 (8.5%) HER2 enriched and 31 (15.0%) triple negative. There were no cases falling into the luminal HER2 subtype.

Using a cut-off of 14% the number of cases classified as luminal A by IR, counting by P1 (CP1) and counting by P2 (CP2) would be 9, 14 and 14 respectively. Using a Chi-square test, the p-value=0.687. (Table 2)

Using a cut-off of 20% the number of cases classified as luminal A by EB, IR, CP1 and CP2 would be 34, 22, 29 and 31 respectively. Using a Chi-square test, the p-value=0.347.

Comparison of EB and CP1 (Fig. 1A) showed a Lin's concordance correlation coefficient of 0.698 with a 95% confidence interval of 0.629 -0.767. On a Bland Altman plot, the mean difference obtained was 5.13% with 6.44% of the values outside the limits of agreement. Using a Ki67 cut-off value of 14% (Table. 3), 17 cases were classified as low (below the cut-off value for luminal A vs. luminal B tumours) by CP1 and 14 cases by EB, and there was disagreement in the classification as low or high Ki67 in 21 cases (10.4%), giving a Cohen's Kappa statistic of k=0.267 ($p<0.001$, fair agreement). Using a cut-off of 20%, 32 cases were classified as low by CP1 and 20 cases by EB, and there was disagreement in 28 cases (13,9%), giving a Cohen's Kappa statistic of k=0.387 ($p<0.001$, fair agreement).

Comparison of IR and CP1 (Fig. 1B) showed a Lin's concordance correlation coefficient of 0.790 with a 95% confidence interval of 0.744 -0.854. On a Bland Altman plot, the mean difference obtained was 8.80% with 4.95% of the values outside the limits of agreement. Using a Ki67 cut-off value of 14% (Table 3), 17 cases were classified as low by CP1 and 9 cases by IR, and there was disagreement in the classification as low or high Ki67 in 14 cases (6.9%), giving a Cohen's Kappa statistic of k=0.428 ($p<0.001$, moderate agreement). Using a cut-off of 20%, 32 cases were classified as low by CP1 and 23 cases by IR, and there was disagreement in 21 cases (10,4%), giving a Cohen's Kappa statistic of k=0.560 ($p<0.001$, moderate agreement).

Comparison of EB and IR (Fig. 2) showed a Lin's concordance correlation coefficient of 0.716 with a 95% confidence interval of 0.650-0.783 ($p<0.001$). On a Bland Altman plot, the mean difference

obtained was -3.73 with 5.91% of the values outside the limits of agreement. Using a Ki67 cut-off value of 14% (Table 3), 14 cases were classified as low by EB and 9 cases by IR, and there was disagreement in classification of high or low Ki67 in 13 cases (6,4%), giving a Cohen's Kappa statistic of k=0.403 (p<0.001, fair agreement). Using a cut-off of 20%, 20 cases were classified as low by EB, and 23 cases by IR and there was disagreement in 23 cases (11,3%), giving a Cohen's Kappa statistic of k=0.402 (p<0.001, fair agreement).

Comparison of counting done by two pathologists (Fig. 2) showed a Lin's correlation coefficient of 0.965 with a 95% confidence interval of 0.955 - 0.974. On a Bland Altman plot, the mean difference obtained was -0.76% with 4.98% of the values outside the limits of agreement. Using a Ki67 cut-off value of 14% (Table 3), 16 cases were classified as low by CP1 and 16 cases by CP2, and there was disagreement in classification as low or high Ki67 in 8 cases (4,0%), giving a Cohen's Kappa statistic of k=0.728 (p<0.001, substantial agreement). Using a cut-off of 20%, 31 cases were classified as low by CP1 and 33 cases by CP2, and there was disagreement in 10 cases (5%), giving a Cohen's Kappa statistic of k=0.814 (p<0.001, almost perfect agreement).

When limiting comparisons to samples within the range of Ki67 readings of 10-25% with either method (Table 3), using a cut-off of 14%, eyeballing vs CP1 showed 32.1% disagreement in classification as low or high Ki67, IR vs CP1 31.1% disagreement, EB vs. IR 22.4% disagreement and CP1 vs. CP2 17.4% disagreement. Using a cut-off of 20% EB vs CP1 showed 44.6% disagreement, IR vs CP1 46.7%, EB vs. IR 42.9% disagreement and CP1 vs CP2 21.7% disagreement.

Discussion:

Immunohistochemistry can serve as a surrogate for molecular assessment of intrinsic subtypes in breast carcinoma, therefore many laboratories now routinely perform immunohistochemistry for oestrogen receptors, progesterone receptors, HER-2-neu and Ki67 on tissue from biopsy or excision

specimens. In respect of the latter the gold standard in methodology has not yet been widely agreed upon. We assessed the comparability of different methods of analysis and the inter-observer agreement for the counting method performed by two pathologists.

The histological grade was assessed as grade 1 in 8%, grade 2 in 44% and grade 3 in 47% and 1% not reported. This is in keeping with racial differences that have been described. (38) In an Australian study (24) 33.2% of the cases were grade 1, 34.2 were grade 2 and 32.6% were grade 3. Carey et al reported grade1: 25%, grade2: 29% and grade 3: 46% in African American patients (38). Breast carcinomas have previously been shown to have more aggressive features in general, including higher grade, in African American women than in white American patients. (39) This was also found in a large Ethiopian cohort (40).

The subtyping of the cases in this study based on immunohistochemistry, when compared to international and African/African American figures showed a lower percentage of cases classified as luminal A, and a higher percentage classified as luminal B. (9,42) Averaging the results of the different methodologies in our study classifies 6% as luminal A and 71% as luminal B at a Ki67 cut-off of 14%; and 14% luminal A and 63% luminal B at 20%. International figures have been reported as luminal A 60%, luminal B 10%, Her2-enriched 15% and Basal-like 15% (3). Huo et al analysed the subtypes for African Americans as follows: luminal A 52%, luminal B 27%, HER2 12%, basal 56%. (28)

Thus there is up to a 10-fold decrease in luminal A cases and up to a 7-fold increase in luminal B cases in our study compared to these. A previous study at Chris-Hani Baragwanath Hospital showed similar figures to our study: luminal A 10.9%, luminal B 70.6%, HER2 enriched 6.6% and triple negative 11.9%(41). Whilst this difference may partially reflect geographical differences, it may also be due to the cut-off level for a low Ki67 index. This highlights the need for determination of the most appropriate cut-off level by each laboratory, (25) by correlation with molecular analysis.

The number of luminal A cases identified by each method using Ki67 cut-offs of 14% and 20% does not show a statistical difference by the Chi-square test, however, these values do not give an entirely accurate reflection of the level of agreement as they do not reflect how many of the cases identified are the same. The Cohen's Kappa statistic is more accurate in this regard.

Overall, good correlation was obtained between the counts done by the two pathologists (Lin's concordance correlation coefficient 0.965).

The best correlation between methods was obtained by IR vs. counting (Lin's concordance correlation coefficient 0.790), followed by EB vs. IR (Lin's concordance correlation co-efficient 0.716) and then EB vs. CP1 (Lin's concordance correlation coefficient 0.698).

Determining a low Ki67 with a cut-off of 14%, fair agreement was obtained between EB and CP1, and between EB and IR, with moderate agreement between IR and CP1, and substantial agreement for manual counting by two observers.

Determining a low Ki67 using a cut-off of 20% yielded similar results, except that agreement for manual counting between the two observers was almost perfect.

The level of agreement was poorer when analysing those cases with Ki67 between 10% and 25%. Up to 46,7% of cases would be differently classified by IR vs. CP1. Up to 44,6% of cases would be differently classified according to EB compared to CP1.

Even with a 4% disagreement (CP1 vs. CP2 with 14% cut-off) 8 of the 201 cases would potentially receive suboptimal management. The high disagreement between CP1 and CP2, even when counting the same images, may be due to the fact that even small differences in counting around the cut-off point result in different categorisation of cases.

A similar finding was obtained by Varga et al (32). Five pathologists evaluated the Ki-67 labelling index in carcinomas of varying degrees of differentiation. Good to very good agreement in the values was obtained, however, on analysing the results for grade 2 carcinomas only, there was poor

inter-observer and intra-observer agreement. Attempts were made to eliminate some suggested factors that could result in this discrepancy, but these did not improve the result. This group of researchers concluded that the differences lay in "the realm of cognitive psychology". Aeffner et al have reviewed visual and cognitive traps in manual scoring of biomarkers, such as confirmation bias in which observers tend to lean towards information that favours a certain hypothesis. (43)

In practice Ki67 determination by immunohistochemistry is often preferable to molecular methods. (24) This is especially true in resource challenged settings, where the cost of Ki67 is less prohibitive and can form part of the routine histopathological panel. An added advantage is the ability to exclude non-tumour cells such as inflammatory cells by the reporting pathologist.

Whilst the recommended method of determining the Ki67 index is manual counting, there is no definite "gold standard".

Eyeballing is the most subjective method of analysing Ki67, and digit preference is a factor as values are usually rounded to the nearest 5 or 10, which reduces the accuracy. This is particularly likely to introduce a bias when a cut off value such as 14% is used. However, this is the most convenient method to use and some authors claim that it is more accurate than other methods because it is based on an overall impression of all the tissue stained. (32) Here, the experience of the observer is an important factor. (43)

The major problem encountered in our study with the use of IR was the intensity of the counterstain. The intensity of the counterstain was not adequate in some cases and could not be corrected by altering the degree of contrast especially in the older cases examined. Many of the negative cells were not detectable by the software. If this method were to be used routinely, the staining characteristics would have to be optimised for this purpose. Tumour cells are not always readily distinguishable from non-neoplastic cells such as endothelial, stromal and inflammatory cells using IR, although the size of the tumour cells is used as a guideline.

Zabaglo et al. identified the elimination of subjectivity in the interpretation of the Ki-67 index as an advantage of DIA. (44) Aeffner et al (43) discussed the fact that DIA methods require a gold standard or reference method to establish analytical validity – usually this standard is conventional subjective scores determined by an experienced pathologist. This then is a paradox as DIA has been touted for its reproducibility and objectivity by virtue of its elimination of human bias. A disadvantage of DIA identified by these authors is that it lacks the cognitive complexity needed to appreciate nuances in histopathological features unless specifically trained to do so. The recommendation of the International Ki67 in Breast Cancer Working Group is that the use of DIA has not been proven for routine clinical practice. (27)

Counting is tedious and time consuming, and distractions are problematic. Intra-tumoural heterogeneity can also affect the results obtained, although it is less of a problem in core biopsies than in tumour blocks as the volume of tumour is less. This factor was eliminated in this study by the use of identical photomicrographs which resulted in the same fields being counted by both observers. Care must be taken to count only invasive cells, avoiding areas of in situ neoplasia. The cells to accept as positively staining may also differ between observers, such as very pale staining cells, speckled staining and cells that are partially overlapping. It has been recommended that any staining should be accepted as positive, since differences from one cell to another reflect variable staining patterns in different parts of the cell cycle. (45) Varga et al. found that despite attempting to better define which nuclei should be accepted as positive, there was no improvement in interobserver agreement. (32)

Limitations

The values for the EB method were obtained from the report, reflecting the assessment of different pathologists. Thus, whilst the counting method was done using a standardised approach for all the cases, this may not be true of the EB results.

Since the cases were assessed retrospectively, the immunohistochemical staining was not optimised for the use of IR, and darker counterstaining would probably have improved the accuracy of IR.

Conclusion

The different methods of assessing Ki67 yielded acceptable correlation between results at high levels and low levels. The intermediate levels (10-25%) are problematic and using different methods results in discrepancies which are particularly concerning in the subtyping of luminal cancers into luminal A and luminal B subtypes. Lack of standardisation in the interpretation of the Ki67 index needs to be addressed. Additional efforts to improve accuracy in the Ki67 range of 10-25% are warranted. In light of the research referenced above regarding digital image analysis, this deserves further attention and may be useful under the supervision of pathologists in these "borderline" cases. Ideally, these cases should be further analysed by molecular methods to assist with precise classification, to determine the Ki67 cut off value per laboratory and to assess which Ki67 determination method is most superior. Further development in molecular methods may result in these becoming more practical in the future.

References:

1. Sorlie T, Perou CM, Tibshirani R, et al. Gene expression patterns of breast carcinomas distinguish tumor subclasses with clinical implications. Proc Natl Acad Sci USA 2001;98: 10869–10874. Available at: http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=11553815%5Cnhttp://www.ncbi.nlm.nih.gov/pmc/articles/PMC58566/pdf/pq010869.pdf

2. Perou CM, Sorlie T, Eisen, MB, et al. Molecular portraits of human breast tumours. Nature 2000;406: 747–752.

3. Collins LC. Breast, In: Goldblum JR, McKenney JK, Lamps LW, Myers JL. Eds. Surgical Pathology. 11th Ed. Elsevier; 2018.

4. The Cancer Genome Atlas Network. Comprehensive molecular portraits of human breast tumours. Nature 2012;490: 61-70.

5. Van de Vijver MJ, He YD, van't Veer LJ, et al, A gene-expression signature as a predictor of survival in breast cancer. N Eng J Med 2002;347: 1999-2009.

6. Emad AR, Reis-Filho JS, Ellis IO. Combinatorial biomarker expression in breast cancer. Breast Cancer Res Treat 2010;120: 293-308.

7. Goldhirsch A, Winer EP, Coates A S, et al. Personalizing the treatment of women with early breast cancer: Highlights of the St Gallen international expert consensus on the primary therapy of early breast Cancer. Ann Oncol 2013;9: 2206–2223.

8. Cheang MCU, Chia SK, Voduc D, et al. Ki67 index, HER2 status, and prognosis of patients with luminal B breast cancer. JNatlCancerInst, 2009;101: 736-750.

9. Tang P, Tse GM. Immunohistochemical surrogates for molecular classification of breast carcinoma: A 2015 update. Arch Pathol Lab Med 2016;140: 806–814. Available at: http://www.archivesofpathology.org/doi/10.5858/arpa.2015-0133-RA.

10. Masuda H, Baggerly KA, Wang Y, et al. Differential response to neoadjuvant chemotherapy among 7 triple-negative breast cancer molecular subtypes. Clin Cancer Res 2013;19: 5533–5540.

11. Scholzen T, Gerlach C, Catorelli G. An insider's view on how Ki-67, the bright beacon of cell proliferation, became very popular. A tribute to Johannes Gerdes (1950-2016). Histopathology 2018;73: 191-196.DOI: 10.1111/his.13511

12. Gerdes J, Schwab U, Lemke H, et al. Production of a mouse monoclonal antibody reactive with a human nuclear antigen associated with cell proliferation. Int J Cancer 1983;31: 13-20.

13. Bouzubar N, Walker K J, Griffiths K, et al. (1989). Ki67 immunostaining in primary breast carcinoma: clinical and pathological associations. Br J Cancer 1989;59: 943–947.

14. Cuylen S, Blaukopf C, Politi, AZ, et al. Ki67 acts as a biological surfactant to disperse mitotic chromosomes. Nature 2016;535: 308-312.

15. Allegra CJ, Paik S, Colangelo LH, et al. Prognostic value of thymidylate synthase, Ki-67, and P53 in patients with Dukes' B and C colon cancer: A National Cancer Institute – National Surgical Adjuvant Breast and Bowel Project collaborative study. J Clin Oncol 2003;21, 241-250.

16. Trihia H, Murray S, Price K, et al. Ki-67 expression in breast carcinoma: Its association with grading systems, clinical parameters, and other prognostic factors – A surrogate marker. Cancer 2003;97, 1321-31.

17. Coons SW, Johnson PC, Pearl DK. The prognostic significance of Ki-67 labelling indices for oligodendrogliomas. Neurosurgery 1997;41, 878-884.

18. DeAzambuja E, Cardosa F, DeCastro D, 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. Available at: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2359936&tool=pmcentrez&rendertype=abstract.

19. Acs B, Zambo V, Vizkeleti L, et al. Ki-67 as a controversial predictive and prognostic marker in breast cancer patients treated with neoadjuvant chemotherapy. Diagnostic Pathology, 2017:12: 1-12

20. 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 2008:19: 323-334.

21. Criscitiello C, Disalvatore D, DeLaurentiis 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 2013;23: 69–75. Available at: http://dx.doi.org/10.1016/j.breast.2013.11.007.

22. Sueta A, Yamamoto Y, Hayashi M, et al. Clinical significance of pretherapeutic Ki67 as a predictive parameter for response to neoadjuvant chemotherapy in breast cancer; is it equally useful across tumour subtypes?

23. Ellis MJ, Tao Y, Luo J et al Outcome prediction for estrogen receptor positive breast cancer based on postneoadjuvant endocrine therapy tumour characteristics. J Natl Cancer Inst. 2008;100: 1380-1388

24. Pathmanathan N, Balleine, RL. Ki67 and proliferation in breast cancer. J Clin Pathol 2013;66: 512–6. Available at: http://www.ncbi.nlm.nih.gov/pubmed/23436927.

25. Mei-Yin, C., Polley, C.Y., Leung, L.M., et al. An international Ki67 reproducibility study. J Natl Cancer Inst 2003;105: 1897-1906.

26. Aleskandarany MA, Green AR, Ashankyty I, et al. Impact of intratumoural heterogeneity on the assessment of Ki67 expression in breast cancer. Breast 2016;15: 287-295.

27. Dowsett M, Torste, ON, 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: 1656-1664. Doi.org/10.1093/jnci/djr393

28. Focke CM, van Diest PJ, DeckerT. St. Gallen 2015 subtyping of luminal breast cancers: Impact of different Ki67-based proliferation assessment methods. Breast Cancer Res Treat 2016;159: 257-263 doi:10.1007/s10549-016-3950-5

29. Coates AS, Winer EP, Goldhirsch RD, et.al. Tailoring therapies – improving the management of early breast cancer: St. Gallen international expert consensus on the primary therapy of early breast cancer 2015. Ann Oncol 2015;26: 1533-1546.

30. Denkert C, Loibl S, Muller B, et al. Ki67 levels as predictive and prognostic parameters in pretherapeutic breast cancer core biopsies: A translational investigation in the neoadjuvantgepartrio trial. Ann Oncol 2016;24: 2786–2793.

31. Gudlaugsson E, Skatland I, Jansen EAM, Smaaland R, Shao Z, Malpica A, 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: 1134-1144.

32. Varga Z, Diebold J, Dommann-Scherer C, et.al. How reliable is Ki67 Immunohistochemistry in grade 2 Breast Carcinomas? A QA study of the Swiss Working Group of Breast- and Gynecopathologists. PLoS ONE 2012;7, e37379 doi:10.1371/journal.pone.0037379

33. Stalhammer G, Robertson S, Wedlund L, Lippert M, Rantalainen M, Bergh J et al. Digital image analysis of Ki67 in hot spots is superior to bath manual Ki67 and mitotic counts in breast cancer. Histopathology 2018; 72, 974-989.

34. Klauschen F, Wienert S, Schmidtt WD, Loibl S, Gerber B, Blohmer J et al. Standardised Ki67 diagnostics using automated scoring – clinical validation in the GeparTrio Breast Cancer Study. Clinical Cancer Research. 2014;21(16): 3651-3657.

35. Tuominen VJ, Ruotoistenmaki S, Viitanen A, et al. ImmunoRatio: a publicly available web application for quantitative image analysis of estrogen receptor (ER), progesterone receptor (PR), and Ki-67. Breast Cancer Res 2010;12, R56. Available at: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2949645&tool=pmcentrez&rendertype=abstract%5Cnhttp://www.ncbi.nlm.nih.gov/pubmed/20663194%5Cnhttp://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC2949645.

36. Watson P, Petrie A, Method agreement analysis: A review of correct methodology. Theriogenology 2010;73, 1167-1179.

37. Hammond MEH, Hayes DF, Dowsett M, Allred DC, Hagerty KL et al. American Society of Clinical Oncology/College of American Pathologists Guideline recommendations for Immunohistochemical Testing of Oestrogen and Progesterone Receptors in Breast Cancer. Arch Pathol Lab Med 2010; 134: 907-922.

38. Carey LA, Perou CM, Livasey CA, et al. Race, Breast Cancer Subtypes and Survival in the Carolina Breast Cancer Study. JAMA, 2006:295, 2492-2502.

39. Furberg H, Milikan R, Dressler L, Newman B, Geradts J. Tumour characteristics in African American and white women. Breast Cancer Res Treat. 2001; 68, 33-43

40. Kantelhardt EJ, Zerche P, Mathewos A, Trocchi P, Addissie A, Aynalem A, et al. Breast cancer survival in Ethiopia: A cohort study of 1070 women. Int J Cancer. 2014; 135(3): 702-9 doi: 10.1002/ijc.28691

41. Joffe M, Ayeni O, Norris SA, McCormack VA, et al. Barriers to early presentation of breast cancer among women in Soweto, South Africa. PLoS ONE 2018; 13, e0192071 doi: 10.1371/journal.pone.0192071

42. Huo D, Hu H, Rhie S et al. Comparison of breast cancer molecular features and survival by African and European ancestry in The Cancer Genome Atlas. JAMA Oncol 2017; E1-E10. Doi:10.1001/

43. Aeffner F, Wilson K, Martin NT, Black JC, Hendriks CLL, Bolon B, et al. The gold standard paradox in digital image analysis. Manual versus automated screening as ground truth. Arch Pathol Lab Med. 2017;141(9): 1267-1275 doi: 10.5858/arpa.2016-2016-0386-RA

44. Zabaglo L, Salter J, Anderson H. Comparative validation of the SP6 antibody to Ki67 in breast cancer. J Clin Pathol 2010;63, 800-804.

45. Urruticocoechea A, Smith IE, Dowsett M. Proliferation marker Ki-67 in early breast cancer. J Clin Oncol 2005;23, 7212-20

APPENDIX A

FIGURES

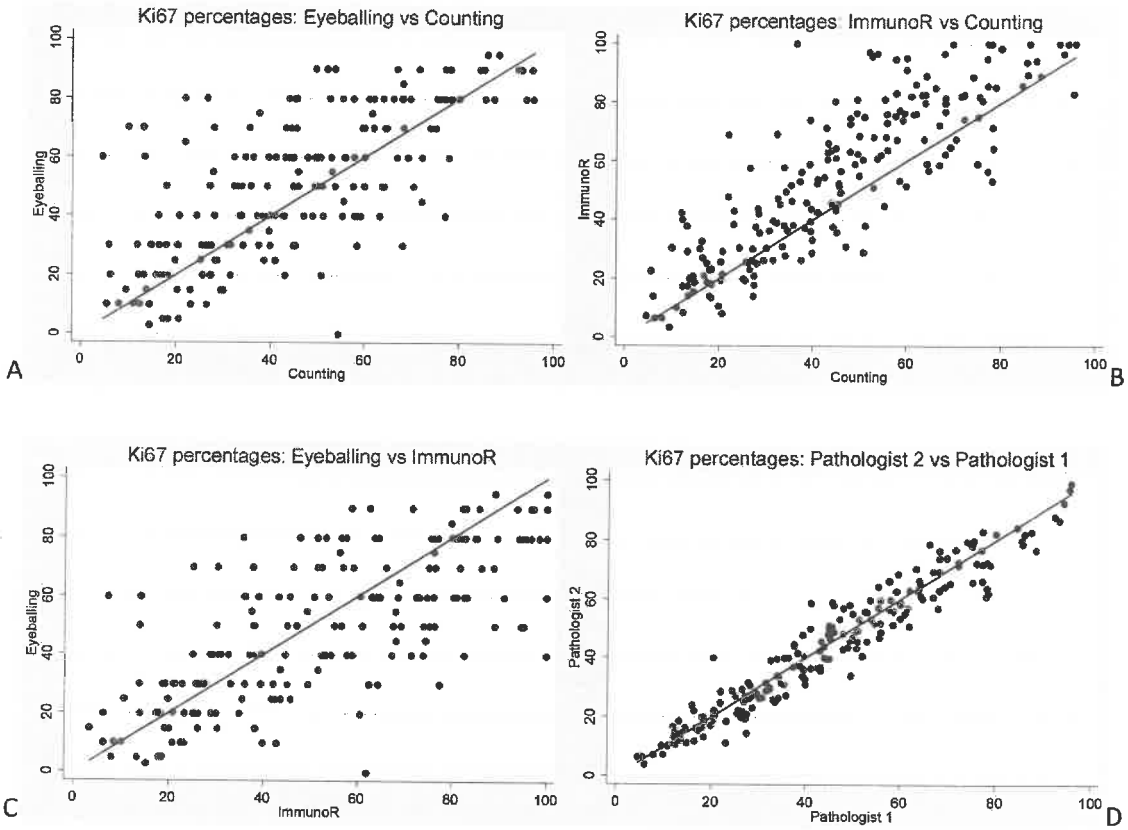
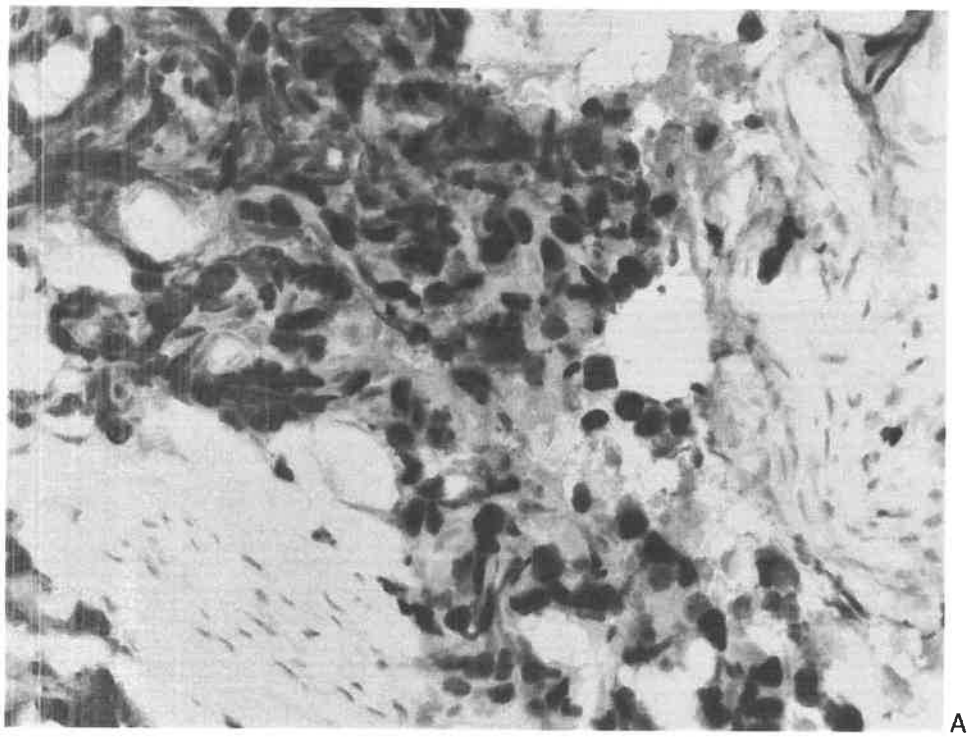
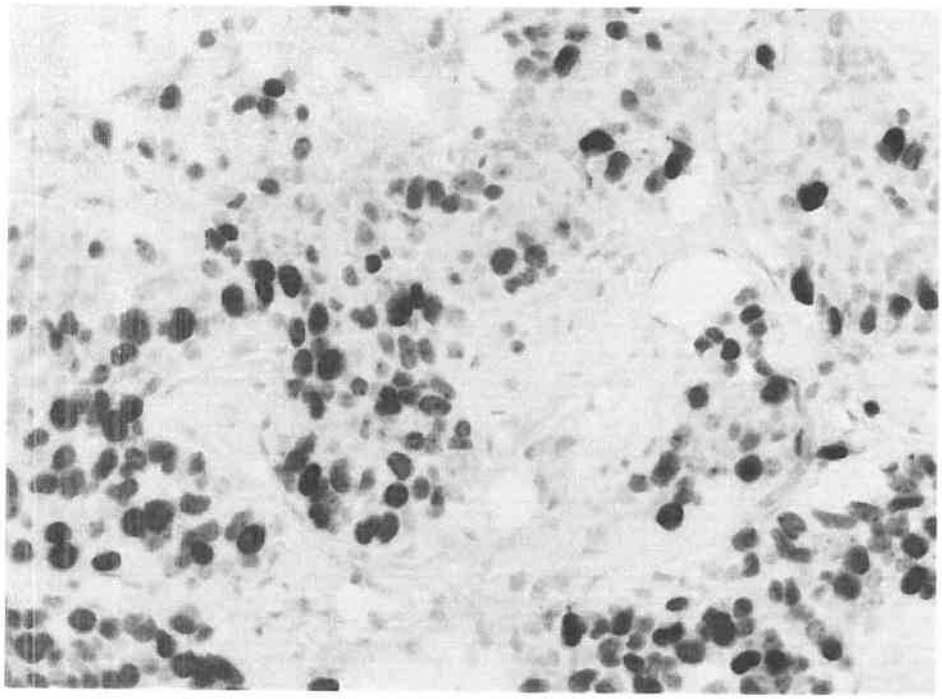


Fig. 1. Scatter plots of Ki67 percentages with a 45° line. A. Eyeballing vs. CP1. B. IR vs CP1. C. Eyeballing vs. IR. D. CP1 vs CP2.



A



B

Fig. 2 Technical problems in Ki67 analysis. Ki67 immunostain at 400x magnification. A. Crush artefact precluding accurate counting. B. Pale counterstain hindering IR analysis.

APPENDIX B - TABLES

Table1. Patient Characteristics

Characteristic	Description	Number (n=201)
Race	Black	201 (100%)
Age	Range	27-86 years
	Median	52 years
	Interquartile range	42-60years
Histological type	No special type	179 (89%)
	Mucinous	7 (3.5%)
	Lobular	6 (3%)
	Micropapillary	3 (1.5%)
	Medullary features	2 (1%)
	Apocrine	2 (1%)
	Metaplastic	1 (0.5%)
	Mixed	1 (0.5%)
Histological grade	Grade 1	16 (8%)
	Grade 2	88 (44%)
	Grade 3	95 (47%)
	Not reported	2 (1%)
Intrinsic subtype	Luminal A	13 (6.5%)
	Luminal B	143 (70%)
	HER2 enriched	17 (8.5%)
	Triple negative	31 (15%)

Table 2. Assessment of numbers of Luminal A and Luminal B cases determined by the different methods.

		Eyeballing (n=203)	IR (n=203)	CP1 (n=202)	CP2 (n=201)	Chi-square test
14% Ki67 Cut-off	Luminal A	13 (6.4%)	9 (4.4%)	14 (6.9%)	14 (7.0%)	P=0.687
	Luminal B	143 (70.4%)	147 (72.4%)	142 (70.3%)	142 (70.2%)	
20% Ki67 Cut-off	Luminal A	34 (16.8%)	22 (10.8%)	29 (14.4%)	31 (15.4%)	P=0.347
	Luminal B	122 (60.1%)	134 (66.0%)	127 (62.9%)	125 (61.7%)	

Table 3. Percentage disagreement between methods when considering all cases or only cases in the range of 10 – 25% (determined by either of the methods).

	All cases			Ki67 10-25% with either method		
	Total cases	<14% cut-off	<20% cut-off	Total cases	<14% cut-off	<20% cut-off
Eyeballing vs counting	202	21 (10.4%)	28 (13.9%)	56	18 (32.1%)	25 (44.6%)
IR vs Counting	202	14 (6.9%)	21 (10.4%)	45	14 (31.1%)	21 (46.7%)
Eyeball vs IR	203	13 (6.4%)	23 (11.3%)	49	11 (22.4%)	21 (42.9%)
CP1 vs. CP2	201	8 (4.0%)	10 (5.0%)	46	8 (17.4%)	10 (21.7%)

APPENDIX C

APPROVED RESEARCH PROTOCOL

Ki67 immunohistochemistry quantification in breast carcinoma: A comparison of visual estimation, counting and Immunoratio©.

1. Introduction

Histopathological analysis of biopsies for breast carcinoma includes immunohistochemical staining for oestrogen receptors (ER), progesterone receptors (PR), human epidermal growth factor receptor (HER-2) and Ki-67. Together with the tumour stage and morphologic assessment of the tumours, these supply information which is routinely used in management decisions and prognostication of breast carcinoma.

Microarray-based molecular analysis of breast carcinomas has shown that breast carcinomas include distinct sub-groups with a spectrum of molecular and biological characteristics (Sorlie, 2001). This was demonstrated by Perou and Sorlie in 2000, who used a hierarchical method of clustering of gene expression to divide breast cancers into subtypes that correlate with their morphologic and clinical features (Perou, 2000). They defined five groups: luminal A, luminal B, normal breast-like, HER2-enriched and basal-like. Subsequently, multiple other groups have been described.

Molecular profiling has led to a more personalised approach to therapy (Emad, 2010). Similar results have been found with genomic analysis (Cancer Genome Atlas Network, 2012) as well as with multiparameter tests including PAM50 (Sorlie et al, 2001) and MammaPrint/BluePrint. Employing these methods for each biopsy evaluated routinely is not viable in many laboratories in view of the financial cost and time constraints, and the St. Gallen conference of 2013 advised the use of immunohistochemical stains for molecular classification (Goldhirsch, 2013).

Analysis of the immunohistochemical stains listed above has been used to serve as a surrogate assessment (Ping Tang, 2016) as follows:

1.Luminal breast cancer

a. Luminal A: ER and PR positive (PR > 20%), HER2 negative and low Ki67.

b. Luminal B: ER+, HER2-, Ki67 > 14%

or ER+, HER2-, Ki67 14-19%, and PR < 20%.

2. HER2 Positive

a.Luminal HER2 Subtype: ER+ (but often not as strongly positive as HER2 negative tumours)

b.HER2+-Enriched subtype: ER-, PR-, HER2+

3.Triple negative.

Basal-like subtype: ER-, PR-, HER2-, Cytokeratin5+ and/or EGFR+

Nonclassified triple negative subtype: Negative for ER, PR, HER2, Cytokeratin5 and EGFR.

Whilst many triple negative breast cancers fall into the basal-like subtype, seven subgroups have been identified (Masuda et al, 2013). These subtypes are predictive of different responses to chemotherapeutic agents (Loibl, 2015).

The use of immunohistochemistry in classifying tumours as luminal A or luminal B subtype is an area of controversy. Molecular methods employed in this regard use genes which assess proliferation (Wirapati et al, 2008).

Ki67 is a proliferation marker which was described in 1984 by Gezer et al. Prior to this,indirect methods such as ^{3}H thymidine determination and flow cytometry had been used for this purpose (Bouzubar, 1989), but were not procedures which could practically be used on a day to day basis. Ki67 reacts with cells in the cell cycle, but not in the quiescent phase.

A high proliferation index as defined by Ki67 staining has been shown to correlate with a poorer prognosis (de Azambujo et al, 2007). It also is an indication of the likely response to chemotherapy – the tumour response is better with a high Ki67 (Criscitiello et al, 2014). Patients with luminal A breast cancers have been found to profit most from hormone therapy without chemotherapy, whereas those with luminal B tumours respond better to chemotherapy (Emad, 2010).

There is widespread controversy regarding the assessment and interpretation of Ki67. Determination of the proliferation index by this method has been found to have unacceptable interobserver and interlaboratory variability (Pathmanathan, 2015). Some authors are of the opinion that routine reporting of Ki67 should not be done at all in view of its lack of reliability (Mei-Yin, 2013).

There are many variables that impact on the use of Ki67 immunostaining. Various methods are used for assessment of the percentage of Ki67 positive cells. These

include "eyeballing" (visually scanning the section and estimating the percentage of positive cells), formally counting the number of cells, and image analysis methods (Pathmanathan, 2015). Contrary to the instinctive idea that counting cells is the most accurate method, one study has found "eyeballing" to be a more reproducible method, although overall interobserver reproducibility is poor (Varga et al, 2012).

Also, it is uncertain whether the highest staining areas ("hot spots") should be assessed, whether average areas should be assessed, whether low staining areas should be assessed or whether multiple areas should be assessed and a mean value calculated. Proponents of the first method argue that the "hot spots" are the areas representing the tumour component most likely to metastasize and therefore the most important component prognostically (Aleskandrarany et al., 2016). Assessment of the invasive edge of the tumour has been recommended (Dowsett et al, 2011).

If counting of positive cells is done, there is a variable approach to the number of cells counted. The International Ki67 in Breast Cancer Working Group recommended a minimal of 500 cells, and ideally 1000 cells to be counted (Dowsett et al, 2011). One study showed no significant difference between counting 300, 510 or 1020 cells using the standard method, but lower figures counting 1200 cells than 510 cells if using the "hot spot" or advancing edge approach (Focke et al, 2016).

The "cut-off" value to be used in assigning tumours to the luminal A or luminal B groups is also not agreed upon. The ideal value has been shown to vary from one institution to another, and these have ranged from 3,5% to 34% (De Azarmbuja, 2007) The St. Gallen International Expert Consensus in 2011 recommended a cut-off of 14% (Goldhirsch et al, 2013). At the St. Gallen conference of 2015, most of the experts favoured a value of 20-29% (Coates et al, 2015). There are also different thresholds obtained if studies are conducted using different clinical end-points and sub-sets of patients (Denckert et al, 2013).

It is recommended that the Ki67 cut-off value should be determined separately for each laboratory (Mei-Yin et al, 2013).

Image analysis software in general is not freely available and necessitates the use of complex software for analysis, as well as virtual microscopy scanning. ImmunoRatio is a digital image analysis system which is accessed via a web browser. It is free and no additional equipment or software is required. (Tuominen et al, 2010). The website includes a Camera adjustment wizard to set the contrast and brightness of the images.

Comparison of Immunoratio results to visual counting showed a good correlation. The analysis has been found to be optimal using images taken with a 20x objective, and averaging the results for three images. (Tuominen et al, 2010).

2. Aims and objectives

Aim

Comparison of three different methods of quantifying the proliferation index on Ki67 immunohistochemical stains performed on breast core biopsies.

Objectives

2.1 A review of the results obtained by 'eyeballing' of breast core biopsies.
2.2 Determination of Ki67 using Immunoratio© software on the same cases.
2.3 Ki67 determined by counting positive cells.
2.4 Comparison of the values obtained between the two methods.
2.5 Determination of the classification of oestrogen-receptor positive cases using the St. Gallen recommended cut-off points of 14% and 20%.

Null hypothesis: There is no significant difference in Ki67 obtained by the different processes.

3. Materials and Methods

This is a retrospective, non-interventional study.

3.1 The study will be conducted in the laboratory setting.
3.2 The original report captured on the NHLS information system (Trakcare) will be examined to determine the Ki67 reported on each case. This is assumed to have been determined by the "eyeballing" method.

3.3 The slides stained with Ki67 immunostain will be retrieved. The slides have been stained according to the laboratory Standard Operating Procedure (See Appendix A).

3.4 The Ki67 value will be obtained using the counting method. 500 cells will be counted for each case in consecutive fields. Any degree of staining will be included. The counting will be corroborated by the first supervisor.

3.5 The slides will be interpreted using the Immunoratio© software. Photomicrographs of three areas for each case will be analysed and a mean result obtained. These areas will include a "hot spot", a "cold spot" and an intermediate area. A 20x microscope objective will be used.

3.6 An Olympus BH-2 microscope with an attached Zeiss AxiocamERc 5s camera will be used to photograph the cases. The camera will be calibrated according to the Immunoratio© Camera adjustment Wizard.

3.7 The data will be entered on an Excel spreadsheet.

3.8 For objective 2.5 the cases will be subtyped using the St. Gallen Ki67 cutoff values of 14% and 20% in order to determine the number of cases which would be classified as luminal A in each case.

4 Inclusion and exclusion criteria.

4.1 Inclusion Criteria.

4.1.1 Cases diagnostic of invasive breast carcinoma, luminal subtype on breast core biopsy.

4.2 Exclusion Criteria.

4.2.1 Biopsies with only duct carcinoma in situ or microinvasion.

4.2.2 Biopsies with representation of insufficient tumour for three photomicrographs of different areas.

4.2.3 Cases in which a Ki67 stain has not been done.

4.2.4 Tumours which have been previously treated with chemotherapy or radiotherapy.

4.2.5 Tumours which are extensively necrotic or severely inflamed.

4.2.6 Biopsies with artefactual changes hampering evaluation.

5 Sample size and selection.

201 breast carcinoma cases satisfying the selection criteria will be included. These will be obtained from the cases enrolled in the Chris Hani Baragwanath SABC study. Sample size calculation: The formula recommended by Watson et al (2010) has been used as follows:

$$1 + \frac{8(1.96)^2(1-\rho_1)^2(1+\rho_1)^2}{2W_\rho^2}$$

Where ρ_1 is the anticipated value of the interclass coefficient (ICC), and W_ρ the acceptable width of the 95% confidence interval. Using a ρ_1=0.8 and W_p=0.1 a sample size of 201 is needed.

6 Statistical analysis.

Comparisons will be done to evaluate both the repeatability of the cell counting results (researcher vs supervisor) and the reliability of the 'eyeballing' results ('eyeballing' vs Immunoratio© software and 'eyeballing' vs positive cell count). Initially, systematic differences between the two groups will be determined using a Wilcoxon signed ranks test as data are expected to be non-parametric. Thereafter, a Bland and Altman diagram will be drawn to examine the agreement between the paired readings. A Bland and Altman diagram will also demonstrate if there is a funnel effect present where the magnitude of the effect has an influence on the degree of the agreement. Finally, when assessing the reliability of the different techniques, Lin's concordance correlation coefficient will be calculated which is a modified version of the Pearson correlation coefficient that assesses how well points lie on a 45° line drawn through the origin.

7. Budget and funding.

No funding is required.

8 Ethics

The study is a retrospective study conducted on archived slides in the NHLS laboratory. The identifying details will be known only to the student and supervisor and all information will be de-identified.

The Department of Anatomical Pathology has been granted blanket approval by the Committee for Research on Human Subjects (Medical) of the University of the Witwatersrand to conduct such retrospective studies (M10744). Nevertheless, a separate application will be submitted to the committee as soon as the protocol is approved for project specific clearance.

9 Time line

Month	1	2	3	4	5	6	7	8	9	10	11	12
Literature review	x	x										
Protocol presentation			x									
Protocol assessment				x	x							
Ethics application				x	x							
Data collection						x	x	x				
Data analysis								x	x			
Writing up										x	x	
Submission												x

APPENDIX A

Appendix A: procedure for immunohistochemistry tests using the Ventana Benchmark XT machine.

1. Equipment and reagents
    a. Equipment: Benchmark XT
    b. Reagents:
        i. EZprep
        ii. LCS: Liquid Cover slippeSSC IN bottle 3: 18l dH20 + 2lSSC

iii. Reaction buffer CC1: Cell Conditioning 1 in bottle 5: undiluted CC2: Cell conditioning 2.
iv. Antibodies
v. Ultra-View DAB kit
vi. Opti-view kit

c. Reagent preparation:
i. EZ prep in bottle number 1: 18l dH2O + 2l EZ prep solution
ii. LCS in bottle 2: undiluted
iii. SSC in bottle 3: 18l dH2O + 2l SSC solution
iv. Reaction buffer in bottle 4: 18l dH2O + 2l Reaction buffer
v. CC1 in bottle 5: undiluted
vi. CC2

2. Procedural steps
a. Cut the blocks according to the immunohistochemistry request form.
b. Print the bar-coded slide labels which include the test to run, laboratory number and date.
c. Label the slides with the appropriate printed labels.
d. Open the machine to load the slides with the barcodes facing towards the centre and close it.
e. Remove the antibodies from the fridge and load them onto the machine.
f. Check if the bulk solutions need to be filled up before starting the run.
g. Check the waste container and empty it if necessary.
h. Click the run button to start the run and enter the number of slides you loaded and the end time for the run.
3. Ki67 antibody used: 30-9, Roche diagnostic, Ventana, USA.

References

Aleskandarany, M.A., Green, A.R., Ashankyty, I., et al. (2016) Impact of intratumoural heterogeneity on the assessment of Ki67 expression in breast cancer. *Breast cancer research and treatment,* 158(2), p.287-295.

Bastien, R.R.L., Rodriguez-Lescure, A., Ebbert, M.T.A., et al.(2012) PAM50 Breast Cancer Subtyping by RT-qPCR and Concordance with Standard Clinical Molecular Markers; *BMC Medical Genomics*, 5(44).

Bouzubar, N., Walker, K. J., Griffiths, K., et al. (1989). Ki67 immunostaining in primary breast carcinoma: clinical and pathological associations, *British Journal of Cancer,* 59, p.943–947.

The Cancer Genome Atlas Network. (2012) Comprehensive molecular portraits of human breast tumours. *Nature,* 490, p.61-70.

Coates, A.S., Winer, E.P., Goldhirsch, R.D., et.al. (2015) Tailoring therapies – improving the management of early breast cancer: St. Gallen international expert consensus on the primary therapy of early breast cancer 2015. *Annals of Oncology*, 26(8) p.1533-1546.

Criscitiello, C., Disalvatore, D., DeLaurentiis, M., et al. (2013) 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*, 23(1), p.69–75. Available at: http://dx.doi.org/10.1016/j.breast.2013.11.007.

deAzambuja, E., Cardosa, F., DeCastro, D., et al., (2007). Ki-67 as prognostic marker in early breast cancer: a meta-analysis of published studies involving 12,155 patients. *British journal of cancer*, 96(10), p.1504–13. Available at: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2359936&tool=pmcentrez&rendertype=abstract.

Denkert, C., Loibl, S., Muller, B., et al. (2013) Ki67 levels as predictive and prognostic parameters in pretherapeutic breast cancer core biopsies: A translational investigation in the neoadjuvant gepartrio trial. *Annals of Oncology*, 24(11), p.2786–2793.

Emad, A.R., Reis-Filho, J.S., Ellis, I.O. (2010) Combinatorial biomarker expression in breast cancer. *Breast Cancer Research and Treatment;* 120, p.293-308.

Focke, C.M., van Diest, P.J., Decker,T. (2016).St. Gallen 2015 subtyping of luminal breast cancers: Impact of different Ki67-based proliferation assessment

methods. *Breast Cancer research and Treatment*, 159(2), p.257-263 doi:10.1007/s10549-016-3950-5

Goldhirsch, A.,Winer, E. P., Coates, A. S., et al. (2013) Personalizing the treatment of women with early breast cancer: Highlights of the St Gallen international expert consensus on the primary therapy of early breast Cancer. *Annals of Oncology*, 24(9), pp.2206–2223.

Loibl, S.(2015). PG 8.01 Primary systemic therapy for clinicians: medical and research perspectives. *The Breast*, 24, pp.S13–S14. Available at: http://linkinghub.elsevier.com/retrieve/pii/S0960977615700305.

Masuda, H., Baggerly, K.A., Wang, Y., et al.(2013) Differential response to neoadjuvant chemotherapy among 7 triple-negative breast cancer molecular subtypes. *Clinical Cancer Research*, 19(19), pp.5533–5540.

Mei-Yin, C., Polley, C.Y., Leung, L.M., et al., 2013. An international Ki67 reproducibility study. *Journal of the National Cancer Institute*, 105, pp. 1897-1906.

Pathmanathan, N. &Balleine, R.L.(2013) Ki67 and proliferation in breast cancer. *Journal of clinical pathology*, 66(6), pp.512–6. Available at: http://www.ncbi.nlm.nih.gov/pubmed/23436927.

Perou, C.M., Sorlie, T., Eisen, M.B., et al.(2000). Molecular portraits of human breast tumours. *Nature*, 406(6797), pp.747–752.

Sorlie, T., Perou, C.M., Tibshirani, R., et al.(2001) Gene expression patterns of breast carcinomas distinguish tumor subclasses with clinical implications. *Proceedings of the National Academy of Sciences of the United States of America*, 98(19), pp.10869–10874. Available at: http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=11553815%5Cnhttp://www.ncbi.nlm.nih.gov/pmc/articles/PMC58566/pdf/pq010869.pdf

Tang, P. &Tse, G.M.(2016). Immunohistochemical Surrogates for Molecular Classification of Breast Carcinoma: A 2015 Update. *Archives of Pathology & Laboratory Medicine*, 140(8), pp.806–814. Available at: http://www.archivesofpathology.org/doi/10.5858/arpa.2015-0133-RA.

Tuominen, V.J., Ruotoistenmaki, S., Viitanen, A.,et al.,(2010)ImmunoRatio: a publicly available web application for quantitative image analysis of estrogen receptor (ER), progesterone receptor (PR), and Ki-67. *Breast cancer research : BCR*, 12(4), p.R56. Available at: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2949645&tool=pmcentrez

&rendertype=abstract%5Cnhttp://www.ncbi.nlm.nih.gov/pubmed/20663194%5Cnhttp://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC2949645.

Wason, P., Petrie, A., (2010) Method agreement analysis: A review of correct methodology. *Theriogenology,* 73(2010) 1167-1179.

Wirapati, P., Sotiriou, C., Kunkel., S., et al. (2008) Meta-analysis of gene expression profiles in breast cancer: toward a unified understanding of breast cancer subtyping and prognosis signatures. *Breast Cancer Research*, 10(4), p.R65. Available at: http://breast-cancer-research.com/content/10/4/R65/abstract%5Cnhttp://breast-cancer-research.com/content/10/4/r65%5Cnhttp://breast-cancer-research.com/content/pdf/bcr2124.pdf.

APPENDIX D

ETHICS CLEARANCE CERTIFICATE

UNIVERSITY OF THE WITWATERSRAND JOHANNESBURG

R14/49 Dr E van den Berg

HUMAN RESEARCH ETHICS COMMITTEE (MEDICAL)
CLEARANCE CERTIFICATE NO. M180574

NAME: (Principal Investigator)	Dr E van den Berg
DEPARTMENT:	School of Pathology Department of Anatomical Pathology NHLS
PROJECT TITLE:	Ki67 immunohistochemistry quantifiation in breast carcinoma: a comparison of visual estimation, counting and immunoratio
DATE CONSIDERED:	Ad hoc
DECISION:	Approved unconditionally
CONDITIONS:	
SUPERVISOR:	Drs R Mohanlal and R Duarte
APPROVED BY:	Professor CB Penny, Chairperson, HREC (Medical)
DATE OF APPROVAL:	06/06/2018

This clearance certificate is valid for 5 years from date of approval. Extension may be applied for.

DECLARATION OF INVESTIGATORS

To be completed in duplicate and **ONE COPY** returned to the Research Office Secretary on 3rd floor, Phillip V Tobias Building, Parktown, University of the Witwatersrand, Johannesburg.

I/We fully understand the conditions under which I am/we are authorised to carry out the above-mentioned research and I/we undertake to ensure compliance with these conditions. Should any departure be contemplated from the research protocol as approved, I/we undertake to resubmit to the Committee. **I agree to submit a yearly progress report**. The date for annual re-certification will be one year after the date of convened meeting where the study was initially reviewed. In this case, the study was initially reviewed in **2018/05/01** and will therefore be due in the month of **2018/05/01** each year. Unreported changes to the application may invalidate the clearance given by the HREC (Medical).

Principal Investigator Signature _____ Date _____

PLEASE QUOTE THE PROTOCOL NUMBER IN ALL ENQUIRIES

APPENDIX E

TURN IT IN REPORT

a0014170:Eunice_MMed_Turnitin.docx

ORIGINALITY REPORT

7%	2%	3%	5%
SIMILARITY INDEX	INTERNET SOURCES	PUBLICATIONS	STUDENT PAPERS

PRIMARY SOURCES

1	Submitted to University College London Student Paper	1%
2	Submitted to National postgraduate Medical College of Nigeria Student Paper	1%
3	www.science.gov Internet Source	1%
4	Submitted to Charles Sturt University Student Paper	<1%
5	theses.gla.ac.uk Internet Source	<1%
6	archivesofpathology.org Internet Source	<1%
7	Joy L. Barber, Alexsandra Zambrano-Perez, Øystein E. Olsen, Fevronia Kiparissi et al. "Detecting inflammation in inflammatory bowel disease — how does ultrasound compare to magnetic resonance enterography using standardised scoring systems?", Pediatric	<1%

	Radiology, 2018 Publication	
8	M. C. U. Cheang. "Ki67 Index, HER2 Status, and Prognosis of Patients With Luminal B Breast Cancer", JNCI Journal of the National Cancer Institute, 05/12/2009 Publication	<1%
9	Submitted to Higher Education Commission Pakistan Student Paper	<1%
10	www.dovepress.com Internet Source	<1%
11	Stephen Stathis, Bec Litchfield, Paul Letters, Ivan Doolan, Graham Martin. "A Comparative Assessment of Suicide Risk for Young People in Youth Detention", Archives of Suicide Research, 2008 Publication	<1%
12	Sangjung Park, Sunyoung Park, Jungho Kim, Sungwoo Ahn, Kwang Hwa Park, Hyeyoung Lee. "Assessment of Ki-67 for Predicting Effective Prognosis in Breast Cancer Subtypes", Biomedical Science Letters, 2018 Publication	<1%
13	Submitted to 9561 Student Paper	<1%

14	Gustav Stålhammar, Stephanie Robertson, Lena Wedlund, Michael Lippert, Mattias Rantalainen, Jonas Bergh, Johan Hartman. "Digital image analysis of Ki67 in hot spots is superior to both manual Ki67 and mitotic counts in breast cancer", Histopathology, 2018 Publication	<1%
15	Gudlaugsson, Einar, Ivar Skaland, Emiel A M Janssen, Rune Smaaland, Zhiming Shao, Anais Malpica, Feja Voorhorst, and Jan P A Baak. "Comparison of the effect of different techniques for measurement of Ki67 proliferation on reproducibility and prognosis prediction accuracy in breast cancer :", Histopathology, 2012. Publication	<1%
16	Michi Morita, Rin Yamaguchi, Maki Tanaka, Gary M Tse et al. "Two progressive pathways of microinvasive carcinoma: low-grade luminal pathway and high-grade HER2 pathway based on high tumour-infiltrating lymphocytes", Journal of Clinical Pathology, 2016 Publication	<1%
17	Noureddin Nakhostin Ansari, Soofia Naghdi, Hoda Moammeri, Shohreh Jalaie. "Ashworth Scales are unreliable for the assessment of muscle spasticity", Physiotherapy Theory and	<1%

	Practice, 2009 Publication	
18	"ESP Abstracts 2015", Virchows Archiv, 2015 Publication	<1%
19	Submitted to The University of Manchester Student Paper	<1%
20	"Breast", Modern Pathology, 01/2008 Publication	<1%
21	Submitted to University of Greenwich Student Paper	<1%

Exclude quotes	On	Exclude matches	Off
Exclude bibliography	On		

APPENDIX F

LIST OF CORRECTIONS

MMED corrections – list

Examiner 1

1. Similar studies and their findings have been discussed in the introduction. Page 4, line17 paragraph 2 –page 5, line 7, paragraph 1
2. Value for Ki67 included. Page 1. Line21 Point 1a
3. "PR+" added. Page 1. Line22. Point 1b
4. Predictive and prognostic value of Ki67 discussed. Page 2, line20, paragraph 3 to page3, line 17, paragraph2
5. Aims and objectives. Added: "To determine the percentages of intrinsic subtypes using the Ki67 obtained by the different methods." Page 5. "Aims" point 2. Line18
6. "Cross-sectional study" added to methods section. Page 5, line 23.
7. Paragraph 4 from the methods section moved to introductory paragraphs. Page 5 paragraph 5, Line 25
8. "SABC" written out as "South African Breast Cancer" study. Page 5, line 26, paragraph 1
9. Time period for selection of cases stated. Page 6, line2, paragraph 1
10. Total number of cases from which they were selected stated. Page 6, line1, paragraph 1
11. Inclusion criteria stated. Page 6, line2-4, paragraph 1
12. Breast carcinoma in males added in exclusion criteria. Page 6 paragraph 2 line 7 - 10
13. Statement regarding oestrogen positivity clarified. Page 6, line 14-16, paragraph 3
14. Description of Ki67 counting method expanded. Page 6, line 21, paragraph 5 – page 7, line 3, paragraph1.
15. What does the Ki-67 value represent? Has been discussed on page 2, lines 6-18, paragraph 2.
16. Use of "eyeballing" results from the original reports was stated as a limitation. Page 14, line22-24, paragraph 3
17. Stated that no cases fitted into the luminal Her2 category. Page 8, line24, paragraph 5
18. Digital preference in "eyeballing" method has been commented on. Page 13, line 12-17, paragraph 3
19. Confidence intervals in tables 2 and 3 discussed with biostatistician – the tables give absolute numbers (frequencies) in each category therefore there can be no confidence intervals.
20. Discrepancies in the values quoted in the "results" section compared to Tables 2 and 3 have been corrected. Page 8, line 22, paragraph 4 – page 10, line19, paragraph3
21. Error corrected "16 cases were classified as low by CP1 and 16 by CP2" Page 10, line 10, paragraph2
22. Histological grades discussed in the light of other studies. Page 11, line 4-10, paragraph 2
23. Data from elsewhere has been quoted re. the spread of subtypes. Page 11, line 12 - 24, paragraph 2
24. Reference to another study with poor agreement in mid-ranges. Page 12, line 22, paragraph 8- page 13, line 5, paragraph 1
25. Additional studies comparing digital image analysis and counting discussed. Page 5, lines 1 – 7, paragraph 1.
26. Comparison of interobserver variability has been separated from the comparison of the different methods. Page 12, lines 5-6, paragraph 2.
27. A limitations paragraph has been added Page 14, line 21, paragraph3

Examiner 2

1. Figures for spread of breast cancers between subtypes included, as well as African/African American figures. Page 11, lines 12-24, paragraph 3-4
2. "Proponents of the first method" changed to "proponents of the 'hot spot' method". Page 4, line1, paragraph 1
3. "Standard method" defined. Page 4, lines 7-8, paragraph 1
4. "Lower figures counting 1200 cells than 510 cells..." rephrased. Page 4, line8, paragraph1
5. Clarified use of the same photos for counting and Immunoratio. Page 7, lines 7-9, paragraph 2
6. Stated that consecutive cases used. Page 6, lines 2 – 4, paragraph 1
7. Statistical methods explained. Page 7, line 14 to page 8, line 5
8. Abbreviation "EB" defined. Page 3, line 23, paragraph 3
9. Explanation attempted for 21.7% discrepancy between CP1 and CP2 in 10-25% range. Page 12, lines 18 – 21, paragraph7
10. Numbers of cases falling into the 10-25% range have been included in table 3. Page 23
11. Study showing correlation between molecular and surrogate immunohistochemical subtyping quoted. Page 1, lines 16 – 18, paragraph 2
12. Standardisation (gold standard) discussed in the conclusion. Page 15, lines 4 – 14, paragraph 2.