Heterogeneity Between Ducts of the Same Nuclear Grade Involved by Duct Carcinoma In Situ (DCIS) of the Breast

Rutgers University has made this article freely available. Please share how this access benefits you.
Your story matters. [https://rucore.libraries.rutgers.edu/rutgers-lib/30309/story/]

This work is the VERSION OF RECORD (VoR)
This is the fixed version of an article made available by an organization that acts as a publisher by formally and exclusively declaring the article "published". If it is an "early release" article (formally identified as being published even before the compilation of a volume issue and assignment of associated metadata), it is citable via some permanent identifier(s), and final copy-editing, proof corrections, layout, and typesetting have been applied.

Citation to Publisher Version: Miller, Naomi A., Chapman, Judith-Anne W., Christens-Barry, William A., Fu, Yuejiao, Yuan, Yan, Lickley, H. Lavina & Axelrod, David E. (2010). Heterogeneity Between Ducts of the Same Nuclear Grade Involved by Duct Carcinoma In Situ (DCIS) of the Breast. Cancer Informatics 2010(9), 209-216.


Terms of Use: Copyright for scholarly resources published in RUcore is retained by the copyright holder. By virtue of its appearance in this open access medium, you are free to use this resource, with proper attribution, in educational and other non-commercial settings. Other uses, such as reproduction or republication, may require the permission of the copyright holder.

Article begins on next page
Heterogeneity Between Ducts of the Same Nuclear Grade Involved by Duct Carcinoma In Situ (DCIS) of the Breast

Naomi A. Miller1, Judith-Anne W. Chapman2, Jin Qian3, William A. Christens-Barry4, Yuejiao Fu5, Yan Yuan6, H. Lavina A. Lickley7 and David E. Axelrod8

1Department of Pathology, University Health Network, 200 Elizabeth St., Eaton 11-444, Toronto, Ontario, Canada. M5G 2C4. 2NCIC Clinical Trials Group, Queens University, 10 Stuart Street, Queens University, Kingston, Ontario, Canada. K7L 3N6. 3Math Elite Inc., 135 Arbour Crest Dr. NW, Calgary, AB, Canada, T3G 4L2. 4Equipoise Imaging LLC, 4009 St. Johns Lane, Ellicott City, MD, USA 21042. 5Department of Mathematics and Statistics, York University, 4700 Keele Street, Toronto, Ontario, Canada. M3J 1P3. 6Department of Population Health Research, Cancer Epidemiology, Prevention and Screening, Alberta Health Services, 1331-29th Street NW, Calgary, AB, Canada, T2N 4N2. 7Henrietta Banting Breast Centre, Women's College Hospital, 76 Grenville Street, 7th floor, Toronto, Ontario, Canada. M5S 1B2. 8Department of Genetics and Cancer Institute of New Jersey, Rutgers University, 604 Allison Road, Piscataway, NJ 08854-8082, USA. Corresponding author email: axelrod@biology.rutgers.edu

Abstract

Purpose: Nuclear grade of breast DCIS is considered during patient management decision-making although it may have only a modest prognostic association with therapeutic outcome. We hypothesized that visual inspection may miss substantive differences in nuclei classified as having the same nuclear grade. To test this hypothesis, we measured subvisual nuclear features by quantitative image cytometry for nuclei with the same grade, and tested for statistical differences in these features.

Experimental design and statistical analysis: Thirty-nine nuclear digital image features of about 100 nuclei were measured in digital images of H&E stained slides of 81 breast biopsy specimens. One field with at least 5 ducts was evaluated for each patient. We compared features of nuclei with the same grade in multiple ducts of the same patient, and compared features of nuclei with the same grade in two ducts of different patients using 2-sided t-tests (P ≤ 0.05). Also, we compared image features for nuclei in patients with single grade to those with the same grade in patients with multiple grades using t-tests.

Results: Statistically significant differences were detected in nuclear features between ducts with the same nuclear grade, both in different ducts of the same patient, and between ducts in different patients with DCIS of more than one grade.

Conclusion: Nuclei in ducts visually described as having the same nuclear grade had significantly different subvisual digital image features. These subvisual differences may be considered additional manifestations of heterogeneity over and above differences that can be observed microscopically. This heterogeneity may explain the inconsistency of nuclear grading as a prognostic factor.

Keywords: breast DCIS, heterogeneity, nuclear grade, image cytometry
Introduction

Nuclear grade is one of the pathological factors reported for duct carcinoma in situ (DCIS) of the breast and may influence patient management.\(^1\) Nuclear grading is the assignment of a numerical value to reflect various nuclear characteristics such as pleomorphism, size, nucleoli (whether present, single or multiple), chromatin features (diffuse, coarse or vesicular), and mitotic activity. Assignment of nuclear grades has been found to be more consistent than architectural patterns.\(^2\) Nuclear grade may identify patients with DCIS at higher risk of recurrence after local excision.\(^3\)

However, in our previous study, grade was not found to have a significant association with either local DCIS recurrence, or the development of invasive disease, regardless of whether grade was assessed as worst or predominant.\(^4\) In this previous study, nearly 50% of the patients had DCIS with more than one nuclear grade; the different grades were either in different ducts in the same area, or in different ducts in different areas, and in some instances different grades were noted in the same duct. This heterogeneity in nuclear grading may have contributed to the apparent lack of statistical significance of grading in the development of invasive disease or DCIS recurrence. Some other recent studies have reported similar heterogeneity of nuclear grading in DCIS.\(^5,6\) Nevertheless, while there is increasing recognition that DCIS is heterogeneous, the significance of grading heterogeneity, and the possible effect of this on prognosis is not clearly understood. In addition, in clinical practice, there is still an expectation that DCIS in a patient is assigned a single nuclear grade.

In the context of individualized patient management, if patients with DCIS of different grades are treated differently, the issue of heterogeneity and the effect of this on assignment of a single grade is particularly important.\(^7,8\) The inherent subjectivity in grading DCIS should also be considered. Several grading systems have been proposed,\(^3,9,10\) but there is a lack of international consensus about the appropriate pathologic grading system to use for DCIS.

Since we previously did not find that nuclear grade had a prognostic effect when many patients had mixed nuclear grade, we assessed undetectable differences in routine microscopy with quantitative image cytometry features, and found that some of these features are significantly associated with the development of invasive disease or DCIS recurrence.\(^11,12\)

This demonstration of prognostic relevance for quantitative cytometry suggested a role for digital image cytometry in evaluating nuclear features of DCIS. In this communication, we utilize image cytometry to assess nuclear heterogeneity by looking at whether a patient with a single grade exhibited differences between ducts, and whether the presence of more than one nuclear grade in the same patient affects differences in nuclear features in ducts of the same grade. We focused attention on patients with grades that might influence decision-making: grade 2 nuclei in patients with only grade 2, compared to grade 2 nuclei from patients with both grade 2 and grade 3, and the analogous situation for grade 3 nuclei.

Materials and Methods

Patients and specimens

Out of the full cohort of 124 patients with DCIS that were studied previously,\(^4\) this study was restricted to the subgroup of 88 patients who had undergone a lumpectomy alone without adjuvant radiotherapy. The presentation of these 88 patients is as follows: 58 presented mammographically, 24 clinically (palpable mass and/or nipple discharge), and in 6 patients the presentation was unknown. Details of specimen handling are as previously reported.\(^4\) The cohort of patients was accrued between 1979 and 1994, the specimens were handled in a manner consistent with the standards at the time: the specimens were sampled directed by the gross appearance of the specimen and the position of any localizing needle, but were not submitted in toto. The tissue was formalin fixed and had not been frozen. The slides evaluated by morphometry were prepared using tissue sections of uniform thickness, approximately 3–4 microns, and stained with hematoxylin and eosin. Grades were assigned based on review of the entire slides of the case. The DCIS was graded into three grades: low, intermediate and high grade (grades 1, 2, 3 respectively) based on nuclear size and appearance of nuclear chromatin and nucleoli.\(^13–15\) When more than one grade was present, all grades, the worst (highest) grade and the predominant (most extensive) grades were recorded. This project was reviewed and approved by the Insti-
Image analysis

For each patient, one low power microscopic field was selected in which there were a minimum of 5 ducts containing DCIS. At high power (40×), digital images of DCIS were obtained: five computer images were obtained from each of these 5 selected ducts. Image features were measured for each of approximately 20 representative nuclei per duct, for a total of approximately 100 nuclei per patient. The nuclear grade(s) were recorded for each duct. For each nucleus, 39 feature values were determined in three categories. (i) Morphometry: area, perimeter, ellipse major axis, ellipse minor axis, ellipticity (major axis/minor axis), shape form factor (4 × π × area/perimeter squared), and roundness b (4 × area/π × ellipticity squared). (ii) Densitometry: mean density, standard deviation of density, modal density, minimum density, maximum density, sum density (mean density × area, used instead of I.O.D. of NIH-Image), range density. (iii) Markovian texture features were calculated from the Markovian co-occurrence matrix of pixel densities with a step size of 2. They were angular second moment, contrast, correlation, variance, inverse difference moment, entropy, entropy, variance, initial entropy, final entropy, entropy, sum entropy, difference entropy, coefficient of variation, peak transition probability, diagonal variance, diagonal moment, second diagonal moment, product moment, and triangular symmetry. Additional texture features, calculated from the binned histogram of pixel gray scale values, included histogram mean, histogram variance, histogram skewness, and histogram kurtosis. Further details of the digital image analysis method were reported previously.11,12 For this study, satisfactory computer images were obtained and data extracted for 81 patients. The reproducibility of the image measurements obtained was assessed in two ways, with repeated measurements of the same nucleus and with duplicate measurements made at different times of randomly selected nuclei from seven randomly selected patients.

Hardware

A custom image cytometry system was assembled which consisted of a CCD camera attached to a bright field microscope and linked to a desktop computer with a frame grabber card. Images of nuclei were acquired and stored as follows: hematoxylin and eosin stained slides were viewed with a bright field microscope (Wild model M20), 40× N.A. 0.75 objective, 1.25× phototube, 530–590 nm band pass green filter, detected with an 8 bit monochrome CCD camera (Sonyo model VDC3874) connected to a video monitor (RCA TC1112) and a frame grabber card (60 HZ Data Translation Quick Capture model DT2255) in a desktop computer (Apple Macintosh model IIci, 12 MB RAM, 80 MB hard disk), and stored as uncompressed TIFF files on removable media (Zip 100 disks). Ten frames were averaged and acquired using NIH-Image software (v. 1.57, written by Wayne Rasband, obtained from the internet by anonymous FTP). Each TIFF formatted image was 640 × 480 pixels, with 256 gray levels. The resulting pixel images were isotropic, with an effective size of 0.25 microns × 0.25 microns. Segregated nuclear images were of modest resolution, typically containing 800 to 1600 square pixels. Sizes were calibrated with a B&L stage micrometer. Optical density of pixel gray values was standardized and camera response calibrated with a set of neutral density filters (50, 25, and 12.5% transmission).

Software

NIH-Image v.1.62b34-Arnv software (modified from http://rsb.info.nih.gov/) and StatView v. 5.01 statistical package (BrainPower, Calabasas, CA, USA) were used to measure and calculate DNA densitometric and nuclear morphometric features using a Mac G4 computer. TextureCalc v. 1.1ax, software (written by W. C.-B.) was used to rebin 256 gray levels into 8 intervals and to calculate texture features from the Markovian gray level co-occurrence matrices. Programs written in SAS release 6.12 for the Macintosh (SAS Institute, Inc., Cary, NC, USA) were used to format data and BMDP PC Dynamic Version 7.0 (Statistical Solutions, Sagua, MA, USA) was used for statistical analysis.

Statistical analysis

We looked at intra-DCIS heterogeneity, i.e. whether there were differences between ducts of patients who...
had the same nuclear grade, Figure 1. For each patient, the 39 digital image feature data were pooled across all nuclei in a duct to yield for the duct, 39 summary feature values [mean image feature/standard error of the mean (s.e.m.)]. We examined whether the values of each continuous image feature were significantly different by grade. If there was no evidence against equal variances across grades ($P \geq 0.10$, by Levene test), then an analysis of variance (ANOVA) F-test was used; however, if there was evidence of heterogeneous variance by the Levene test ($P < 0.10$), the Welch’s adaptation of the t-test which does not assume equal variances, was used to test for homogeneity of image feature values between ducts with the same nuclear grade(s). Then, when there was significant evidence of different image feature value(s) in different ducts, pair-wise t-tests were used to test for differences in image feature value(s) between ducts with the same nuclear grade.

We looked at inter-person heterogeneity of image analysis features, i.e. whether image features for DCIS of a particular grade in a patient with a single nuclear grade differed from those of DCIS of the same grade in a patient with more than one grade, Figure 2. As before, the image feature data were pooled for each patient across all nuclei in ducts with the same nuclear grade, to yield summary feature values of mean/s.e.m. for each of the 39 image features, for nuclei of the same nuclear grade. Two-sided Student t-tests were used to test for differences between grade 2 nuclei in patients with only grade 2 nuclei compared to grade 2 nuclei in patients with both grade 2 and grade 3 nuclei. Also, Student t-tests were used to test for differences in grade 3 nuclei in patients with only grade 3 nuclei compared to patients with both grade 2 and grade 3 nuclei. Comparisons of digital image analysis features for patients with grade 1 nuclei (grade 1, grades 1 and 2, and grades 1, 2, and 3) were not considered because of the low numbers of patients in each of these categories.

Authors contributions
D.E.A. acquired the data, was involved in design of the study, analysis and interpretation of results, drafting and preparing the final manuscript. J.W.C. was involved in design of the study, analysis and interpretation of results, drafting and preparing the final manuscript. N.M. was involved in design of the study, pathological review of specimens, interpretation of results, and preparing the final manuscript. W.C.-B. wrote the software used to extract the image features and reviewed the manuscript. J.Q., Y.Y and Y.F. were involved in analysis of data and reviewed...
the manuscript. H.L. was involved in acquisition of specimens through surgical management of patients and reviewed the manuscript.

Results
In order to determine the reliability of the data extracted by image analysis of nuclei, two kinds of measurements were made. First, repeated measurements were made of one feature (i.e. area) of one nucleus. The coefficient of variation of 150 measurements was 3.4%. Second, for 10 nuclei of each of 7 patients, duplicate measurements were made of all 39 features. Duplicate measurement were made at different times. The percent of features in pairs of measurements that were not statistically different (two tailed paired t-test, 0.05 level of significance) ranged from 90% to 97% for different nuclei.

The distribution of the number of patients and their nuclear grade, or grades, is shown in Table 1. Of the total of 81 patients, 47 had a single grade and 34 had more than one grade. Figure 3 shows an example of more than one grade in the same patient. Where there was more than one grade in the same patient, the different grades were in different ducts in some cases or in the same duct in other cases.

Comparisons were made between image features of nuclei from different ducts of the same nuclear grade(s) in the same patient. Some patients had a single grade and other patients had more than one grade. Figure 1 illustrates an example of comparisons in a patient with only one grade. Table 2 shows the results of pair-wise comparisons for 2, 3, 4 or 5 ducts with the same grade (grade 2 compared with grade 2, and grade 3 compared with grade 3), along with the number of significant t-test that were expected and observed. For each of the comparisons, except that with 2 ducts, the observed number of significant features exceeded the expected. Therefore, statistically significant differences were detected between nuclei of the same grade in different ducts of the same patient.

Comparisons were made between image features for grade 2 nuclei in patients with only grade 2 nuclei with grade 2 nuclei in different patients with both grades 2 and 3, Figure 2. All 39 image features were significantly different, \( P \leq 0.05 \) in two sided t-tests. Similarly, comparisons were made between grade 3 nuclei in patients with only grade 3 with grade 3 nuclei in patients with both grades 2 and 3. There were significant differences in 28 of the 39 image features. There were significant differences in each type of image feature assessed: 6/7 morphometric, 5/7

Table 1. Number of patients by nuclear grade.*

<table>
<thead>
<tr>
<th>Nuclear grade or grades</th>
<th>Number of patients</th>
</tr>
</thead>
<tbody>
<tr>
<td>Grade 1 only</td>
<td>1</td>
</tr>
<tr>
<td>Grade 2 only</td>
<td>27</td>
</tr>
<tr>
<td>Grade 3 only</td>
<td>19</td>
</tr>
<tr>
<td>Grades 1 and 2</td>
<td>8</td>
</tr>
<tr>
<td>Grades 2 and 3</td>
<td>25</td>
</tr>
<tr>
<td>Grades 1, 2, and 3</td>
<td>1</td>
</tr>
<tr>
<td>Total</td>
<td>81</td>
</tr>
</tbody>
</table>

*Lumpectomy alone patients with image analysis feature data used in assessments of DCIS heterogeneity.
densitometric, and 17/25 texture features. Therefore, statistically significant differences were detected between nuclei of the same grade in ducts of different patients.

Nuclei of the same grade in the same patient had differences in image features regardless of whether there was a single grade in the duct or multiple grades in the duct. In addition, nuclei of the same grade in different patients differed by whether the ducts had a single or multiple grades.

**Discussion**

There is increasing recognition of heterogeneity within tumors of many different tissues, including heterogeneity within DCIS of the breast. This study extends the previous observations on intratumoral heterogeneity of DCIS by documenting that quantitative differences can exist even between ducts that appear to have the same nuclear grade.

We compared image features for nuclei in ducts with the same nuclear grade within the same patient and found statistically significant differences. Also, differences were detected between ducts of the same nuclear grade in different patients, in which one patient had a single grade and the other patient had more than one grade. Image analysis of digital images of biopsy specimens was able to extract quantitative subvisual information about nuclei that was found to be statistically different.

The high replicability of repeated measurements of nuclear features by image analysis implies that the statistically significant differences reported in this study are unlikely to be due to measurement error, but rather represent real differences between nuclei.

Nuclear grade has been found to be associated with both risk of DCIS recurrence and progression to invasive carcinoma. Based on such results, nuclear grade is a required component of the pathologic evaluation and reporting of DCIS. Traditional nuclear grading depends on visual inspection and subjective judgement. Image cytometry can detect additional subvisual information and the extracted data is amenable to objective statistical analysis. This additional information has been used in the assessment of biopsy specimens of many tissues, including in situ and invasive carcinoma of the breast. In previous studies, we showed that image cytometric features were significantly associated with risk of DCIS recurrence and development of invasive cancer.

Here we show that image cytometry can characterize interductal heterogeneity, the difference between ducts with the same nuclear grade. Since nuclear grade is one of the factors considered in the management of DCIS, not accounting for interductal heterogeneity may have clinical implications.

**Table 2.** Significant differences in image features between ducts with same nuclear grade.

<table>
<thead>
<tr>
<th>Number of ducts same nuclear grade</th>
<th>No. t-tests per patient</th>
<th>Significantly different t-tests per patient*</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Expected</td>
</tr>
<tr>
<td>5</td>
<td>390</td>
<td>20</td>
</tr>
<tr>
<td>4</td>
<td>234</td>
<td>12</td>
</tr>
<tr>
<td>3</td>
<td>117</td>
<td>6</td>
</tr>
<tr>
<td>2</td>
<td>39</td>
<td>2</td>
</tr>
</tbody>
</table>

*Two-sided t-tests, with 5% significance level.
lack of association between nuclear grade and patient outcome in a previous report.4

Interductal heterogeneity also has implications for studies of DCIS at the molecular and cellular levels. Some of these studies are based on analysis of a single patient sample, either a mixture of cells from a “representative” region, or a small number of selected cells from a region obtained by laser capture microdissection. Several studies have compared DCIS to normal tissue, to invasive carcinoma, or to metastatic carcinoma from the same patient by gene expression,35–39 protein expression,40 microsatellite markers,40,41 loss of heterozygosity,43,44 gene amplification or deletion (CGH),45,46 and nuclear image features.29 Many of these studies included paired samples of DCIS and other lesions from the same patient; however, it is often not clear how many samples of DCIS were assessed and therefore whether the sampling method would account for the kind of interductal heterogeneity reported here. Without characterizing multiple samples from different ducts from the same patient, it is not clear if the differences found between the single DCIS sample and the other invasive or metastatic lesion of the same patient would also have been found between multiple samples of DCIS of the same patient.

Interductal heterogeneity can also be a concern in analysis of samples in tissue microarrays. Tissue microarrays often include multiple samples from the same patient. The reproducibility of measurements of pairs of samples has been demonstrated.47,48 However, if the samples analyzed in tissue microarrays come from the same region of tissue, these samples may not reflect the heterogeneity existing in the patient’s tumor.49

Our results suggest that studies of DCIS at the molecular and cellular levels should incorporate analysis of multiple samples from different areas of tissue demonstrating DCIS in order to account for the range of molecular and cellular diversity that may exist between the different ducts within each patient.

In summary, digital image analysis, previously used to quantitatively characterize premalignant and malignant specimens, may reveal subvisible information useful for diagnosis and prognosis of breast and other tumors. In this communication, we used image analysis to reveal heterogeneity between ducts of breast DCIS of the same nuclear grade.

Disclosures
This manuscript has been read and approved by all authors. This paper is unique and not under consideration by any other publication and has not been published elsewhere. W.C.-B. is President and Chief Scientist at Equipoise Imaging LLC, all other authors and peer reviewers of this paper report no conflicts of interest. The authors confirm that they have permission to reproduce any copyrighted material.

Acknowledgements
We thank Dr. Edward B. Fish and Marilyn Link for their extensive work on the clinical database; Drs. D. August, Y. Gusev, R. Sklarew, D. Foran, and W. Hanna for helpful discussions; Dr. W. Sofer for equipment; A. Khokher and A. Kagan for laboratory assistance; and Dr. G. Heiman for suggestions on the manuscript.

D.E.A was supported by National Institutes of Health (NIH) grant U56 CA113304 and the New Jersey Commission for Cancer Research 1076-CCR-S0. J.W.C. is supported by a program grant from the Canadian Cancer Society to the NCIC Clinical Trials Group.

References


