A Newly Proposed Semi-Automated Method of Grading
Invasive Lobular Carcinoma: A Unifying Concept and
Correlation with Prognostic Markers and Patient Survival

Emily Stevens,1 Bruce F. Kimler,2 Marilyn K. Davis,1 Fang Fan,1 Patricia Thomas,1
Xiao-Yun Wang,1 Ivan Damjanov,1 and Ossama W. Tawfik1
1Pathology and Laboratory Medicine Department and 2Radiation Oncology Department,
The University of Kansas Medical Center, Kansas City, Kansas

Abstract. Invasive ductal carcinoma (IDC) of the breast is currently graded according to the Nottingham
modification of the Scarff-Bloom-Richardson system (SBR). This system involves subjective evaluation of
3 morphologic features: tubule formation, nuclear pleomorphism, and mitosis. Our recently proposed
semi-automated Nuclear and Proliferation Index [N+P] grading system for IDC has demonstrated
agreement among grades and prognostic markers with better prediction of patient survival than the SBR
system. Our present objective is to expand the utilization of the N+P system to grading invasive lobular
carcinoma (ILC). Fifty-eight ILC cases were evaluated by the SBR and N+P systems. The 2 systems were
compared in terms of correlation with patient survival, tumor size, grade, angiolymphatic invasion, lymph
node status, ploidy status, and ER, PR, Her-2, p53, EGFR, and Bcl-2 staining. The N+P and SBR systems
demonstrated overall agreement when correlated with clinical and prognostic parameters. Twenty-four of
30 tumors initially classified as SBR Grade II were down-graded to N+P I. Three of 26 tumors initially
classified as SBR Grade I were up-graded to N+P II. Grading of ILC provides valuable predictive and
prognostic information. The N+P grading system for ILC decreases the element of subjectivity for assessing
mitotic activity and appears to be superior to the SBR system in predicting patient survival.

Keywords: invasive lobular carcinoma, tumor grading, breast cancer

Introduction

Invasive mammary carcinoma is a histologically
heterogeneous disease, with invasive ductal
carcinoma (IDC) of no special type representing
the majority of carcinomas. Invasive lobular
carcinoma (ILC) comprises approximately 10-15%
of breast cancers and appears to have a biology that
is distinct from other carcinoma types [1-5]. As
ILC is less common than IDC, fewer data have
been reported that address the unique biologic
features of ILC in the context of clinical outcome.

Breast carcinomas are currently graded by the
Nottingham modification of the Scarff-Bloom-
Richardson system (SBR), as recommended by the
World Health Organization (WHO) [6]. This
system has become a popular and widely used
grading scheme with proven prognostic significance
for IDC [7]. In contrast, application of the SBR
grating system to ILC is not routinely performed
[8]. Furthermore, correlations of the SBR grade
with prognostic markers, response to adjuvant
treatments, and long-term survival are not well
defined for ILC. The SBR system requires subjective
evaluation of tubule formation, nuclear pleo-
morphism, and mitosis. It is hindered by a lack of
precision in assessing all 3 parameters, with the
most impact in evaluating mitotic frequency [9-
12]. Additionally, grading with the SBR system has

Address correspondence to Ossama W. Tawfik, M.D., Ph.D.,
Department of Pathology and Laboratory Medicine, Univ. of
Kansas Medical Center, 3901 Rainbow Boulevard, Kansas
City, KS 66160, USA; tel 913 588 1185; fax 913 588 8780;
e-mail otawfik@kumc.edu.

Available online at www.annclinlabsci.org
25
been shown to have suboptimal reproducibility [13,14]. Data collected by our group and others demonstrate that utilization of an automated MIB-1 count can be beneficial in assessing the proliferative activity of invasive and in-situ ductal carcinoma [15,16]. Using nuclear grade and Ki67 (MIB-1) proliferative activity, we developed a new grading system designated as the Nuclear and Proliferation Index [N+P] system. The components of the N+P system are nuclear grade and proliferation index. Nuclear grade is defined using nuclear pleomorphism as in the SBR system; however, the proliferation index is based on the percentage of cells expressing MIB-1 by immunohistochemistry (IHC). In IDC, the new system demonstrated overall agreement with the SBR system when correlated with histologic and prognostic parameters [16]. Furthermore, the N+P system appears to be superior to the SBR system with better correlation with patient survival and less subjectivity in assessing proliferative activity.

Our overall objective is to introduce a unifying method for grading breast cancer, applicable to all forms of invasive and in situ carcinomas, that provides a more complete and reliable assessment of their biologic phenotypes and clinical behaviors while minimizing subjectivity. In this paper, we expand our studies to include grading ILC. We assess whether automated image analysis reduces subjectivity in grading such tumors and whether the N+P grading system correlates with patient survival as well as relevant tumor biomarkers.

### Materials and Methods

#### Patient cohort
A total of 788 primary breast carcinomas was examined, consisting of 650 (82.5%) IDC, 58 (7.4%) ILC, 38 (4.8%) mixed ductal and lobular carcinoma, and 42 cases (5.3%) of the other types, such as colloid, tubular, cribriform, medullary, papillary, and metaplastic squamous cell carcinomas. This study was approved by the institutional research committee of the Kansas University Medical Center. Written informed consent was not obtained from patients as it was not required by our institutional research committee for this retrospective, de-identified study. Histologic tumor samples were obtained from core needle biopsy, lumpectomy, and mastectomy specimens. The ILC samples were taken from 52 lumpectomy/mastectomy specimens and 6 core biopsies. Histopathologic parameters, including histologic type, histologic grade, and nuclear grade, were recorded for all patients. All tumors were graded using the the Nottingham modification of the Scarff-Bloom-Richardson system (SBR) [7] and the recently described N+P system [15]. Briefly, the nuclear grade was scored from 1 to 3 using the SBR system. Automated MIB-1 count was scored in 3 categories: score 1 (≤9%), score 2 (10% to 25%), and score 3 (>25%). The rationale for the grading categories for MIB-1 count is based on our previous experience with grading invasive ductal carcinoma by the N+P system [15]. Our analysis showed that segregation of MIB-1 quantification into 3 categories resulted in a fairly even distribution between grades. The highest level of MIB-1 expression significantly differed from the low and intermediate levels of expression (p = 0.022; log-rank test). We tested whether shifting the cutpoints by a few percentage points either way altered the prognostic ability, but there was no improvement in the separation of the curves.

N+P grade I was defined as tumors having nuclear scores of 1 or 2 and MIB-1 scores of 1 or 2. N+P grade II includes tumors having either nuclear scores of 3 and MIB-1 scores of 1 or 2, or tumors with MIB-1 scores of 3 and nuclear scores of 1 or 2 (total score of 4-5). N+P grade III included tumors having both nuclear and MIB-1 scores of 3 (Table 1).

### Table 1. N+P scheme for grading lobular carcinoma of the breast.

<table>
<thead>
<tr>
<th>Feature</th>
<th>Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nuclear grade</td>
<td></td>
</tr>
<tr>
<td>Small regular uniform nuclei with little variation</td>
<td>1</td>
</tr>
<tr>
<td>Moderate nuclear variation</td>
<td>2</td>
</tr>
<tr>
<td>Marked variation in size and shape of nuclei</td>
<td>3</td>
</tr>
<tr>
<td>Automated MIB-1 count</td>
<td></td>
</tr>
<tr>
<td>≤9%</td>
<td>1</td>
</tr>
<tr>
<td>10-25%</td>
<td>2</td>
</tr>
<tr>
<td>&gt;25%</td>
<td>3</td>
</tr>
<tr>
<td>N+P grading:</td>
<td></td>
</tr>
<tr>
<td>Nuclear grade 1 or 2 and MIB-1 scores 1 or 2</td>
<td>grade 1</td>
</tr>
<tr>
<td>Nuclear grade 3 and MIB-1 score 1 or 2 OR</td>
<td>grade 2</td>
</tr>
<tr>
<td>MIB-1 score 3 and nuclear grade 1 or 2</td>
<td></td>
</tr>
<tr>
<td>Both nuclear grade 3 and MIB-1 score 3</td>
<td>grade 3</td>
</tr>
</tbody>
</table>
Additional parameters including patients’ age, tumor size, angiolymphatic invasion, and regional lymph node status were also recorded for lumpectomy/mastectomy specimens.

Patient outcome data were obtained from the Kansas Tumor Registry. Specifically, the date of death was recorded and used to calculate overall survival. We elected to evaluate overall survival, rather than disease-specific survival, because incomplete data and discrepancies regarding the cause of death are common.

Immunohistochemical techniques. Tissue blocks containing the most representative and best-preserved tumor areas were selected for IHC and ploidy studies. Immunohistochemical analysis was performed on tissue fixed with neutral buffered formalin. Estrogen receptor alpha (ERα), progesterone receptor (PR), p53, epidermal growth factor receptor (EGFR), MIB-1, Bcl-2, Her-2, and ploidy analyses were performed on all specimens using a DAKO autostainer (DAKO, Carpinteria, CA). Her-2 antibody was detected using the HercepTest (DAKO) according to the manufacturer’s protocol. Tawfik et al [15] provides information regarding the antibodies, vendors, antibody concentrations, and epitope retrieval and detection methods. Hematoxylin was used as the counterstain and appropriate positive and negative controls were included. Nuclear morphology, tubule formation, and the MIB-1 proliferation index (PI) were analyzed and assigned scores.

Quantification of immunohistochemistry. Positive IHC reactions were defined as a dark brown reaction on the cell membrane for Her-2 and EGFR, positive nuclear staining for MIB-1, ERα, PR, and p53, and positive cytoplasmic staining for Bcl-2. For PI of MIB-1, the percentage of nuclei with immunopositivity was determined using the PI program of the Cell Analysis System (CAS) 200 image analyzer (Bacus Laboratory, Chicago, IL) for the period from 1991 to 2001 or with the Clarient Automated Cellular Imaging System (ACIS™) after 2001 (San Juan Capistrano, CA). For each specimen, 5 to 10 areas with the highest staining intensity were selected for quantification. An average score for all selected areas was then calculated. For ERα, PR, and p53, either the CAS-200 or ACIS systems was used for automated counts. Manual microscopy was used to score tumor staining with antibodies to EGFR and Bcl-2. Her-2 staining was quantified using a score of 0 or 1+ to indicate a negative result (absent to weak, according to the new College of American Pathologists/ American Society of Clinical Oncology [CAP/ASCO] guidelines), and 2+ (equivocal), or 3+ to represent positive staining. Results were validated using the Her-2 scoring system of the ACIS machine. Staining ≥10% of the tumor cells with antibodies to EGFR or Bcl-2 was considered positive, whereas for p53 a count >5 was considered positive.

Statistical analysis. Overall frequencies and percentages were summarized for tumor grade (N+P and SBR systems), ERα, PR, p53, EGFR, Bcl-2, Her-2, vascular invasion, and node positivity. The frequencies of each variable stratified by the grading system were calculated, and their relationships to each grading system were evaluated using the Chi-square test. Summaries of biomarker expression by N+P system stratified by SBR system were also given. Distributions of continuous variables were analyzed by the non-parametric Wilcoxon test. The log-rank test was used to compare overall survival across the 3 grades for the N+P and SBR systems independently. All tests were 2-sided. A p-value of <0.05 was considered as significant, with no correction for multiple comparisons.

Results

The SBR and N+P systems both demonstrated overall agreement when correlated with the clinical and prognostic parameters including age, tumor size, grade, angiolymphatic invasion, lymph node status, ER, PR, Her-2, p53, EGFR, BCL-2, and ploidy status (Table 2). All but 2 cases were graded

Table 2. Relationship between grade (SBR vs N+P), histopathology, and immunohistochemistry of invasive lobular carcinoma.

<table>
<thead>
<tr>
<th>Variable</th>
<th>N</th>
<th>Positive for SBR grade (%)</th>
<th>Positive for N+P grade (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>(median/ cutoff for positivity)</td>
<td></td>
<td>I</td>
<td>II/III</td>
</tr>
<tr>
<td>Mean age, yr</td>
<td>58</td>
<td>60</td>
<td>59</td>
</tr>
<tr>
<td>Tumor size, cm</td>
<td>52</td>
<td>1.9</td>
<td>2.0</td>
</tr>
<tr>
<td>Vascular invasion present</td>
<td>58</td>
<td>12%</td>
<td>31%</td>
</tr>
<tr>
<td>Positive lymph node metastasis</td>
<td>35</td>
<td>27%</td>
<td>30%</td>
</tr>
<tr>
<td>Aneuploid</td>
<td>44</td>
<td>26%</td>
<td>56%</td>
</tr>
<tr>
<td>ER expressing (10%)</td>
<td>56</td>
<td>89%</td>
<td>90%</td>
</tr>
<tr>
<td>PR expressing (10%)</td>
<td>57</td>
<td>65%</td>
<td>61%</td>
</tr>
<tr>
<td>EGFR expressing (10%)</td>
<td>33</td>
<td>0%</td>
<td>9%</td>
</tr>
<tr>
<td>Her-2 expressing (2+, 3+)</td>
<td>50</td>
<td>12%</td>
<td>4%</td>
</tr>
<tr>
<td>Bcl-2 expressing (10%)</td>
<td>36</td>
<td>93%</td>
<td>82%</td>
</tr>
<tr>
<td>p53 expressing (5%)</td>
<td>50</td>
<td>8%</td>
<td>24%</td>
</tr>
</tbody>
</table>

The 2 grading systems demonstrated similar frequencies for the different histologic grades and general agreement with each other for all of the biomarkers studied.
as either I or II by both systems. Of the 30 tumors classified as SBR grade II, 24 (80%) were downgraded to N+P grade I. In contrast, only 3 of 26 (12%) tumors classified as SBR grade I were upgraded to N+P grade II. Fifty-five percent of ILC tumors were graded as II/III by the SBR system, while only 19% were graded as II/III by the N+P system (Table 3).

Comparison of the N+P II/III and SBR II/III tumors revealed distinct differences and trends between the 2 groups. There was borderline significance in the difference between age at diagnosis for N+P low grade (I) vs N+P high grade (II/III) tumors. No such difference was seen for the SBR system. Also, although there was no difference in tumor size, the N+P high grade (II/III) tumors tended to have increased incidence of angiolymphatic invasion (46% vs 31%) and nodal metastasis (27% vs 19%), compared to the SBR high grade (II/III) tumors. Lastly, there were statistically significant differences in EGFR and p53 expression, with higher frequencies of expression in N+P high grade (II/III) vs N+P low grade (I) tumors. No such differences were observed for the SBR system.

Fig. 1 shows that patients with IDC or ILC have similar overall survival curves for 12-yr duration. As in our previous study of IDC [15], the N+P grading system for ILC appears to provide improved separation between the survival curve for low grade (grade I) vs higher grade tumors (grades II and III), although the difference is not statistically significant, given the small number of patients (p = 0.14 by log-rank test). Minimal difference (p = 0.57) was observed in the survival curves for low grade vs high grade tumors obtained by the SBR system (Fig. 3).

![Invasive Breast Cancer](image1)

**Fig. 1.** Overall survival of patients with invasive ductal carcinoma and invasive lobular carcinoma. The survival curve of patients with invasive lobular carcinoma is identical to that of patients with invasive ductal carcinoma.

![N+P Grading System (Lobular)](image2)

**Fig. 2.** Overall survival of patients with invasive lobular carcinoma classified by the N+P grading system. The N+P grading system for ILC appears to provide improved separation between the survival curve for low grade (grade I) vs higher grade tumors (grades II and III), although the difference is not statistically significant, given the small number of patients (p = 0.14 by log-rank test).

### Table 3. Comparison of the SBR to the N+P grading systems for Invasive lobular carcinoma of the breast.

<table>
<thead>
<tr>
<th>N+P Grading System</th>
<th>I</th>
<th>II</th>
<th>III</th>
</tr>
</thead>
<tbody>
<tr>
<td>I</td>
<td>47 (81%)</td>
<td>23§</td>
<td>24*</td>
</tr>
<tr>
<td>II</td>
<td>10 (17%)</td>
<td>3†</td>
<td>6§</td>
</tr>
<tr>
<td>III</td>
<td>1 (2%)</td>
<td>0†</td>
<td>0†</td>
</tr>
<tr>
<td>Total</td>
<td>58 (100%)</td>
<td>26 (45%)</td>
<td>30 (52%)</td>
</tr>
</tbody>
</table>

*Down-graded from SBR to N+P system. §No change from SBR to N+P system. †Up-graded from SBR to N+P system.
Discussion

Invasive lobular carcinoma (ILC) is a well known histologic subtype that represents 10-15% of invasive mammary carcinomas [1,5,17]. Much less is known about the biology of ILC than its more common counter type, IDC [5]. Although our study and others have shown that ILC and IDC have similar overall survival patterns [1,5,18,19], they differ in their biological features. For example, ILCs are unique in their hormonal pattern, with a higher percentage of tumors showing ER and PR positivity, compared to IDC [8]. This finding, however, is not predictive of a favorable outcome for ILC as it is for IDCs with a similar hormonal pattern [1,8].

Relevant histopathologic information such as tumor size, lymph node status, and histologic grade are critical in calculating the Nottingham prognostic index, which oncologists use to categorize breast cancer patients in need of chemotherapy [20]. Recently, additional information such as hormonal status, Her-2 status, and gene expression profiling data are being routinely considered in decision making for the management of breast cancer patients [21-24]. Traditional histologic grading of breast cancers has been shown to be limited by subjectivity. We recently proposed the N+P grading system, which combines nuclear grade with automated quantitative MIB-1 index, to better predict overall survival for patients with invasive and in situ ductal carcinoma [15,16]. Our newly proposed system combines well-established knowledge of breast tumor grading with recent technology utilizing IHC and image analysis.

Previously we showed that SBR grade II IDCs, which have identical overall survival to grade I tumors, could be separated into favorable and unfavorable subgroups by the N+P system. In the present study we addressed the much debated topic of ILC grading. Recent studies have proposed the potential utility of the SBR system in grading ILC, showing good correlation between the different grades and a variety of prognostic markers [8,25,26]. In this study, we not only confirm the need for ILC grading, but also show that the N+P grading system is potentially better than the SBR system. Most importantly, the overall survival curves are better separated as compared to the SBR system. A major effect of the N+P system is to downgrade previously SBR grade II to grade I tumors, a change that elucidates the differential survival of the respective groups. All but 2 cases fell into a low or intermediate grade by both grading systems. This finding may be due in part to the bland nuclei of ILC, which are usually small, uniform, and round to oval with hyperchromasia (ie, nuclear grades 1 or 2). One could argue the potential utility of a 2-tiered instead of a 3-tiered system for grading ILC. Tumors would either be low grade (corresponding to N+P grade I) or high grade (N+P grades II/III). Further studies are recommended to evaluate the validity of this proposed system.

The SBR system consists of 3 histologic criteria: tubule formation, nuclear pleomorphism, and mitotic activity [7]. Resistance to using the SBR system to grade ILC originates from the perception that excessive subjectivity exists within the system. The absence of tubule formation and the inherent low to intermediate nuclear grade in ILC skews the SBR score. In fact, our results show that more than 50% of ILCs tumor graded by the SBR system fell into histologic grade II.

Beyond providing improved survival prediction, the N+P model also correlates better with hormonal and prognostic markers. ER, PR, EFRG, p53, and ploidy are vital information in respect to the
treatment and prognosis of IDC [27]. The same markers have less prognostic impact on ILC [1]. In a recent publication by Bane et al [8], the histologic grade of ILC was correlated with biomarkers and prognostic factors; a major finding was that ER and PR were highly expressed in ILCs irrespective of grade. Our data show a clearer separation of hormonal status when the N+P grading system was used. Furthermore, the expression of favorable (Bcl-2) and unfavorable prognostic factors (EGFR, p53, aneuploidy) correlated better with the N+P grading system. A notable finding in regard to Her-2 was its decreased expression as grades increased for both grading systems. Studies have shown increased deletions of chromosome 17 (which corresponds to the Her-2 gene) in ILC cases compared to IDC [1,27]. In agreement with the observations of Bane et al [8], we found that higher grade N+P ILC tumors were larger in size, with higher incidence of angiolymphatic invasion and increased rate of nodal metastasis.

In conclusion, grading ILC is clinically important, since it provides valuable predictive and prognostic information. However, to date, no universally accepted grading system exists for non-ductal carcinomas. Our long-term goal is the utilization of the N+P system for all invasive breast carcinomas. In this study, we show the N+P system to be superior to the SBR system in ILC, as it correlates better with patient survival. It is also automated in one component, resulting in less subjectivity in assessing mitotic activity. The success we have obtained with applications of N+P to ILC and IDC further supports our hope that the N+P system may unify the grading of invasive and in situ breast carcinomas. Additional research is needed to evaluate correlations of the N+P histologic grade with different treatment modalities, responses to therapy, and disease-free survival in a larger number of cases.

References


