Nuclear Grade Plus Proliferation Grading System for Invasive Ductal Carcinoma of the Breast

Validation in a Tertiary Referral Hospital Cohort

Maria Papadimitriou, MD,1 Sarantos Kaptanis, MD, JD,2 Evangelos Polymeropoulos, MD,3 Grigorios Mitsopoulos, MD, FRCPath,4 Dimitris Stogiannis, PhD,5 Chrys Caroni, PhD,6 George Vaiopoulos, MD, PhD,7 Joannis G. Panayiotides, MD, PhD,8 and Petros Karakitsos, MD, PhD8

From the 1Laikon Hospital, 1st Internal Medicine Department, University of Athens Medical School, Athens, Greece; 2Barts Cancer Institute, Queen Mary, University of London, London, England; 3Athens Red Cross Hospital, Athens, Greece; 4Ealing Hospital NHS Trust, London, England; 5Department of Epidemiology & Public Health, University College London, London, England; 6Department of Mathematics, National Technical University of Athens, Athens, Greece; 7University of Athens Medical School, Athens, Greece; and 8Department of Cytopathology, University of Athens Medical School, “Attikon” University Hospital, Athens, Greece.

Key Words: Invasive breast carcinoma grading; MIB-1 (Ki-67) immunohistochemistry; Automation; Survival; Nottingham Prognostic Index

ABSTRACT

Objectives: For patients with invasive breast cancer, management decisions are informed by tumor grade according to the Nottingham Grading System (NGS), either on its own or as part of the Nottingham Prognostic Index (NPI). A system retaining the nuclear grade element but substituting the two subjective components, mitosis count and tubule formation, of the NGS with a proliferation index based on Ki-67 (MIB-1) has been proposed (nuclear grade plus proliferation [N+P] grading).

Methods: We validated the prognostic value of this grading system on a population of 322 women.

Results: N+P grading resulted in more grade I tumors (47.9% vs 4.5%) and fewer grade II (32% vs 51.5%) and grade III (20.1% vs 44%) tumors compared with NGS. The NPI calculated based on N+P grade had a similar association with survival (P < .001; odds ratio, 1.729) as the NPI calculated on the basis of the NGS grade (P < .001; odds ratio, 1.668).

Conclusions: The N+P system seems equivalent to the NGS system.

Therapeutic decisions in patients with breast cancer largely depend on their classification in a manner that predicts their biology: clinical staging (American Joint Committee on Cancer TNM stages I-IV1), histologic grading (1-32), and molecular biomarker-based classifications (estrogen receptor [ER],3 progesterone receptor [PR],4 human epidermal growth factor receptor 2 [HER2] gene,5 and recently multigene prognostic indices6-7).

Histologic grading of breast cancer is commonly done according to the Nottingham (Elston-Ellis) modification of the Scarff, Bloom, and Richardson grading system,8 also known as the Nottingham Grading System (NGS). In the past, concerns over reproducibility of grading systems have led to reluctance to use grades to aid therapeutic decisions.9 In 2007, Tawfik et al10 introduced the nuclear grade plus proliferation (N+P) grading system, a more objective grading system that is claimed to better represent breast cancer biology.

The N+P grade substitutes MIB-1 count for tubule formation and mitotic count, with the percentage of the total number of breast cancer cells with nuclear staining classified into three categories: 9% or less, 10% to 25%, and more than 25%, respectively. N+P grade I is defined as a tumor having nuclear grade 1 or 2 and MIB-1 25% or less. N+P grade II refers to a tumor with nuclear grade 3 and MIB-1 25% or less, or nuclear grade 1 or 2 with a MIB-1 more than 25%. N+P grade III describes a tumor having a nuclear grade 3 and MIB-1 more than 25%.

The initial results with the N+P grading system showed that it is precise, reliable, and easy to use. MIB-1 assessment
has been shown to outperform mitotic count in interobserver agreement in other cancers. Nonetheless, the inventors of the system did not test whether the N+P grading system might confer additional prognostic information on patient survival when tumor size or lymph node status is taken into account—parameters used to determine prognosis in the Nottingham Prognostic Index (NPI). N+P also has not been evaluated with regard to the prognostic information it confers on disease-free survival.

The present study sought to compare the additional prognostic information that the N+P grading system may confer over the NGS when other relevant clinical parameters are taken into account. It also sought to establish whether the N+P grading system is valid in predicting disease-free survival, as well as overall survival.

### Materials and Methods

#### Patient Cohort

This was a retrospective study. After obtaining institutional board ethics approval (13/22-10-2010), we reviewed medical records of 322 patients treated surgically for breast cancer, from 2003 to 2007, in the Greek tertiary hospital “Elena.” These included pathology reports and complete follow-up records. Completion date for the follow-up of all patients was February 28, 2011. Complete data were available for 319 patients, and these were analyzed further. Median follow-up was 65 months (minimum: 7 months; maximum: 97 months).

Collection of pathology specimens was according to institutional protocol: briefly, mastectomy and wide local excision specimens were formalin fixed in 10% formalin after the surgical procedure. Following paraffin embedding, the tumor specimens were examined by H&E staining. All specimens were evaluated by two consultant pathologists (J.G.P. and P.K.) who reached consensus on tumor grading according to the Nottingham modification of the Scarff-Bloom-Richardson grading system.

#### Immunohistochemistry

Tissue blocks containing representative tumor areas were selected for immunohistochemistry (IHC). IHC was performed on tissue fixed with 10% neutral buffered formalin. Paraffin-embedded tissue blocks were cut to 5-μm sections, deparaffinized in xylene, treated with 100% ethanol, and heat treated in a microwave (0.01 M citrate buffer, pH 6.0 for 15 minutes) for antigen retrieval. Hormone receptors (estrogen, progesterone), Her-2/neu, and Ki-67 were examined on all specimens using monoclonal antibodies and commercially available monoclonal antibodies. Table 1. ER and PR were graded as positive if more than 10% of the tumor cells exhibited nuclear overexpression. HER2/neu was evaluated based on the scoring guidelines of DAKO; scores of 0 and 1+ were considered negative, and 2+ and 3+ were considered positive. At least three sections per tumor were examined. Ki-67 proliferation index was measured using the CAS-200 image analyzer (Becton-Dickinson, San Jose, CA). An average score of 10 areas with the highest staining intensity from each specific lesion was calculated quantitatively using the PI program of the image analyzer.

#### Subtypes

We evaluated tumor subtypes using an accepted histologic classification,12 classifying ER+ or PR+, HER2− as luminal A; ER+ or PR+, HER2+ as luminal B; ER−, PR−, HER2+ as HER2 overexpressing; and ER−, PR−, HER2− as triple negative.

#### Prognostic Indices

NPI was calculated using the original formula: NPI = [0.2 × tumor size in cm] + [number of lymph nodes involved: 0 = 1, 1-3 = 2, >3 = 3] + [grade of tumor], where grade was either the NGS or N+P grade. We used the NPI to stratify our patients into three groups: good prognosis (NPI ≤3.4), moderate prognosis (NPI 3.41-5.4), or poor prognosis (NPI >5.4).

#### Statistical Analysis

We constructed a database wherein the following parameters were included: age, type of surgery (mastectomy or wide local excision), date of operation, ER and PR status, ErbB-2 (HER2) status, tumor size (measured on gross pathologic examination), number of lymph nodes, histologic grade, nuclear grade, Ki-67 staining (percent and classification), history of familial breast cancer, adjuvant chemotherapy, adjuvant hormonal therapy, relapse, chemotherapy treatment on relapse, death, and total length of follow-up. We incorporated the three variables: tumor size, number of lymph nodes, and histologic grade in the NPI.

All statistical analysis was performed on SPSS version 22.0 (SPSS, Chicago, IL). We performed univariate Cox analysis of all variables to find which of the parameters in our database could be included in our model and, following
that, built multivariate Cox analysis models both for overall survival and disease-free survival. Our analysis compares the NPI using the NGS grading system and the NPI using the N+P grading system in a unique patient cohort.

Results

After reviewing the database, 319 patients, all female, were included in the analysis due to missing data on three patients. Patient demographics are presented in Table 2. Importantly, more patients in our cohort underwent breast-conserving surgery than mastectomy (52.4%), and most (68.7%) were placed in the moderate prognosis group according to the NPI. Of the patients, 74.3% were ER positive, 66.8% were PR positive, and 33.9% were HER2 positive. Most patients had adjuvant chemotherapy (94.7%) and hormone therapy (83.1%). The anti-HER2 antibody trastuzumab (Herceptin) was not available for treatment of primary breast cancer in Greece during that period.

The Ki-67 (MIB-1) proliferation index split patients into three groups: 9% or less (24.5%), 10% to 25% (48.6%), and more than 25% (27.0%). We used these data and nuclear grade to calculate the N+P score for all patients in our cohort and the equivalent NPI, and we compare our findings with the standard NGS grading in Table 3. Although the two systems gave similar results for grade and prognostic index, more patients were placed in the good prognosis group using the N+P score (12.5% vs 1.9%), corresponding to more grade I tumors (47.9% vs 4.5%) and less grade II (32% vs 51.5%) and grade III (20.1% vs 44%) tumors.

The importance of prognostic indices is their correlation with patient survival and recurrence (ie, prediction of outcomes) Table 4. At the end of follow-up (median, 65 months), 284 (89.0%) patients were still alive, and 236 (74.0%) patients were disease free. There were six cases of local recurrence, seven recurrences in the contralateral breast, five in the regional lymph nodes, and 64 instances of metastatic disease.

We performed univariate Cox regression to evaluate the effect of patient-specific, disease-specific, and treatment-specific factors on survival and recurrence Table 5. Age (better prognosis with older age, \( P = .018 \)), ER (better prognosis with positive ER, \( P = .003 \)), PR (better prognosis...
with positive PR, \( P = .004 \), tumor size (better prognosis with smaller tumor size, \( P = .036 \)), number of positive lymph nodes (better prognosis with less involved lymph nodes, \( P = .015 \)), lymph node stage (\( P = .012 \)), tumor grade calculated with the NGS system (better prognosis with lower grade, \( P = .039 \)), Ki-67 (MIB-1 classes) (\( P = .017 \)), tumor grade calculated with the N+P system (better prognosis with lower grade, \( P = .006 \)), NPI calculated with the NGS grade (\( P < .001 \)), NPI calculated with the N+P grade (\( P < .001 \)), prognostic groups defined by the NPI calculated with the NGS grade (\( P < .001 \)), prognostic groups defined by the NPI calculated with the N+P grade (better prognosis with smaller scores, \( P < .001 \)), IHC-defined subtypes (from better to worse prognosis: luminal B, luminal A, triple negative, HER2 overexpressing, \( P = .011 \)), and adjuvant hormone treatment (better prognosis for those who received hormone treatment, \( P = .007 \)) were statistically significant factors affecting survival.

Number of positive lymph nodes (\( P = .011 \)), lymph node stage (\( P = .014 \)), tumor grade (NGS) (\( P = .004 \)), nuclear grade (\( P = .006 \)), tumor grade (N+P) (\( P = .002 \)), NPI (NGS) (\( P < .001 \)), NPI (NGS) groups (\( P = .006 \)), NPI (N+P) (\( P < .001 \)), NPI (N+P) groups (\( P = .001 \)), and IHC-defined subtype classification (\( P = .037 \)) were statistically significant factors affecting recurrence.

We found that the N+P-based NPI is not inferior to the NGS-based NPI in defining prognostic groups.
Multivariate Cox regression was performed using statistically significant variables and removing variables that derive from others and create problems of multicollinearity. Thus, we selected age, ER status, PR status, and the relevant NPI for comparisons. The prognostic indices retain their statistical significance in the multivariate models. The value of prognostic indices in the clinic is in guiding management decisions for adjuvant chemotherapy and frequency of outpatient follow-up; in this respect, the N+P NPI model proves slightly superior to the NGS NPI model, clearly separating good and moderate prognosis patients from poor prognosis ones both for overall and disease-free survival.

Discussion

The NGS is endorsed by the World Health Organization, the College of American Pathologists, the American Joint Committee on Cancer Staging manual, the European Union, and the Royal College of Pathologists. It assigns a score of 1 to 3 each to tubule formation (>75% = 1, 10%-75% = 2, <10% = 3, commonly abbreviated as T), nuclear pleomorphism (an indicator of cellular differentiation; small regular uniform cells = 1, moderate nuclear size and variation = 2, marked nuclear variation = 3, commonly abbreviated as N), and mitotic activity (an indicator of proliferation; with 0-9 mitoses/high-power field [hpf] = 1, 10-19 mitoses/hpf = 2, >20 mitoses/hpf = 3, commonly abbreviated as M). A tumor with a total score of 3 to 5 is categorized as grade 1, one with a score of 6 or 7 as grade 2, and one with a score of 8 or 9 as grade 3. This system is reported to have significant interobserver agreement and correlates well with both disease-free and overall survival. At 10-year follow-up, grade 1 cancers are expected to have 85% survival, grade 2 about 60%, and grade 3 only 45%.

The NGS has a modest degree of reproducibility, although continuous standardization efforts have led to
better results over time. Nonetheless, histologic grade is and remains a partially subjective parameter. Thus arises the need for a more objective grading system that is easier to use at the same time: the N+P system.

Antigen Ki-67 is a cellular marker for proliferation. Ki-67, a nonhistone protein, is expressed in proliferating cells through all active phases of the cell cycle (G1, G2, S, M) but is absent from resting cells (G0). Several antibodies against Ki-67 have been assessed, the most commonly used being MIB-1 (DAKO), which, unlike the original Ki-67 antibody, can be used on formalin-fixed, paraffin-embedded tissue.15

Measurement of Ki-67 and receptor status by image analysis is widely reported in the literature16,17 and has been shown to closely correlate with visual analysis.11,18 The advantage of automated, quantitative methods such as the CAS-200 is their objectivity and speed.19 Assessment of Ki-67 and other immunohistochemical markers by automated systems has known weaknesses. Interobserver agreement as low as 56% has been reported,20 and there is controversy over Ki-67 assessment by both image analysis and visual inspection.21 Even though CAS-200 is an image analysis system developed in the 1980s and there are newer, possibly more accurate, methods available or under development today,22 it remains in use in a variety of institutions.

Ki-67 has been assessed as a prognostic factor in a number of studies,23-27 but it is not part of routine clinical evaluation in breast cancer.28 There is evidence that well-established immunohistochemical markers (ER, PR, HER2, Ki-67) combining hormone receptors with a proliferation index (the IHC4 score) could confer similar prognostic information to the 21-gene Genomic Health recurrence score (GHI-RS, also known as Oncotype DX) at a fraction of the cost.29 The cutoff points for Ki-67 used in our study and previously suggested by Tawfik et al10 are far from universal, and more research is needed to evaluate these as prognostic markers. To date, this is the largest patient series from Greece evaluating Ki-67 and derived indexes in breast cancer prognosis. Strengths of our study include population homogeneity since our cohort comes from a single tertiary hospital in a relatively short period of 4 years, no losses to follow-up, and robust, consistent laboratory methods in the assessment of biomarkers from a single university laboratory.

Our study has several weaknesses. It is a cohort study and is subject to inherent bias since patient management is not rigorously controlled as in a clinical trial. Apparent indicators of this are that most (95%) of our patients had adjuvant chemotherapy. There is controversy over Ki-67 assessment by both image analysis and visual inspection21; our study did not include a comparison of the automated evaluation with visual assessment. Data collection, although prospective, did not include factors that may have influenced outcomes—namely, type of chemotherapy regimens and adjuvant radiotherapy. Our follow-up period, although greater than 5 years on average, may be considered short for breast cancer, a disease with an excellent outcome with current treatment modalities; in fact, our series has 89% overall survival and 74% disease-free survival. We encourage additional prospective study of a more controlled patient group.

In conclusion, we have demonstrated the N+P system to be equivalent to the traditional NGS system in the grading of breast cancer in a Greek patient cohort, both as an individual predictor and as part of the NPI. At the same time, the N+P system is simpler and easier to use.

References

Corresponding author: Sarantos Kaptanis, MD, JD, Centre for Tumour Biology, Barts Cancer Institute, Queen Mary, University of London, John Vane Science Centre, Charterhouse Square, London EC1M 6BQ, UK; sarandos@gmail.com.

This research received institutional board ethics approval from the University of Athens (13/22-10-2010) and was supported by Greek tertiary hospital “Elena.”


