The Genomic Grade Assay Compared With Ki67 to Determine Risk of Distant Breast Cancer Recurrence

Michail Ignatiadis, MD, PhD; Hatem A. Azim Jr, MD, PhD; Christine Desmedt, PhD; Isabelle Veys, MD; Denis Larsimont, MD; Roberto Salgado, MD; Maria B. Lyng, PhD; Giuseppe Viale, MD; Brian Leyland-Jones, MD; Anita Giobbie-Hurder, MS; Rosita Kammler, MD; Patrizia Dell’Orto, ScD; Françoise Rothé, PhD; Ioanna Laïos, ScD; Henrik J. Ditzel, MD, PhD; Meredith M. Regan, ScD; Martine Piccart, MD, PhD; Stefan Michiels, PhD; Christos Sotiriou, MD, PhD

IMPORTANCE The Genomic Grade Index (GGI) was previously developed, evaluated on frozen tissue, and shown to be prognostic in early breast cancer. To test the GGI in formalin-fixed, paraffin-embedded breast cancer tumors, a quantitative reverse transcriptase polymerase chain reaction assay was developed and named the Genomic Grade (GG). The GG assay has the potential to increase the clinical application of the GGI, but robust demonstration of the clinical validity of the GG assay is required.

OBJECTIVE To evaluate the prognostic ability of the GG assay to detect breast cancer recurrence compared with centrally reviewed immunohistochemical testing of Ki67 antigen proliferation.

DESIGN, SETTING, AND PARTICIPANTS This is an internationally collaborative substudy of a large phase 3 4-arm adjuvant trial. Patients had endocrine receptor-positive, node-positive, or node-negative nonmetastatic primary breast cancer. Patients included in this study had available formalin-fixed, paraffin-embedded samples of their primary tumors and were randomized to either a 5-year tamoxifen monotherapy arm or a 5-year letrozole monotherapy arm. Associations between either GG assay results or log2-transformed Ki67 data and survival end points were evaluated using Cox regression models stratified for chemotherapy use; the 2 vs 4 arm randomization option; and endocrine therapy assignment with and without adjustment for clinicopathological parameters, including centrally reviewed histological grade, hormone receptors, and ERBB2 (formerly HER2 or HER2/neu). The likelihood ratio statistic was used to assess the added prognostic value.

INTERVENTIONS Central evaluation and comparison, blinded for clinical information, of the GG assay, breast cancer histological grade, and Ki67.

MAIN OUTCOMES AND MEASURES Distant recurrence-free interval (DRFI).

RESULTS Genomic Grade assay data were obtained in 883 breast cancer samples (62%). At a median follow-up of 8.1 years, 84 (10%) had distant recurrences. Increasing GG or Ki67 were both significantly associated with lower DRFI and added independent prognostic information to the clinicopathological prognostic factors. In patients with early node-negative breast cancer who were endocrine-only treated, 38% were GG1 with a 10-year DRFI of 99% (95% CI, 97%-100%), and 18% were histological grade 1 with a 10-year DRFI of 100% (95% CI, 100%-100%). For GG equivocal patients, the 10-year DRFI was 94% (95% CI, 90%-98%), and for GG3 patients, the 10-year DRFI was 87% (95% CI, 80%-94%).

CONCLUSIONS AND RELEVANCE Either the GG assay or centrally reviewed Ki67 significantly improves clinicopathological models to determine distant recurrence of breast cancer. Compared with the histological grade, the GG assay can identify a higher proportion of endocrine-only treated patients with very low risk of distant recurrence at 10 years.

TRIAL REGISTRATION clinicaltrials.gov Identifiers: NCT00004205 and NCT00004205

Author Affiliations: Author affiliations are listed at the end of this article.

Corresponding Author: Christos Sotiriou, MD, PhD, Breast Cancer Translational research Laboratory (BCTL), Institut Jules Bordet, Boulevard de Waterloo, 121, 1000 Brussels, Belgium (christos.sotiriou @bordet.be).

Published online December 3, 2015. Last corrected on September 8, 2016.
Estragon receptor (ER)-positive, node-negative breast cancer is the most frequently diagnosed breast cancer. Patients are frequently treated with adjuvant chemotherapy, although most would have a favorable outcome with endocrine therapy alone. Standard clinicopathological characteristics such as age, tumor size, nodal status, and histological grade (HG), either considered alone or in the context of tools such as Adjuvant! Online (version 8.0; http://www.adjuvantonline.com) and/or the Nottingham prognostic index, have helped clinicians decide the administration of adjuvant chemotherapy. While these tools have shown high sensitivity in identifying patients with high-risk breast cancer, they often overestimate the risk of recurrence of other cancers resulting in the administration of chemotherapy to patients who derive very little benefit.

In the past decade, several genomic tests have been developed to provide additional prognostic information to clinicopathological models and have further refined the decision-making process regarding adjuvant chemotherapy. In a meta-analysis including more than 3000 patient samples, our research group has shown that proliferation is the main driver of the prognostic performance of these genomic tests. Traditionally, cancer proliferation can be assessed using HG or the immunohistochemical testing of nuclear antigen Ki67. Yet, these tools have shown poor reproducibility among pathology laboratories. A recent international reproducibility study on immunohistochemical testing of Ki67 included some of the most experienced laboratories in the world and further showed substantial variability in Ki67 scoring.

Our group has developed a 97-gene signature—the Genomic Grade Index (GGI)—that can accurately and reproducibly quantify tumor proliferation. We previously showed that it can separate HG2 breast tumors into distinct GGI1 and GGI3 groups with different clinical outcomes. The GGI was developed and tested on frozen tissue, but to increase clinical application, we developed the Genomic Grade (GG) assay, a quantitative reverse transcriptase polymerase chain reaction tool that can reliably evaluate breast cancer tumors from formalin-fixed paraffin embedded (FFPE) specimens.

In this study, we aimed to validate the prognostic performance of the GG assay by comparing it with the centrally reviewed immunohistochemical testing of Ki67 in patients in the Breast International Group (BIG) 1-98 trial. This substudy was approved by the International Breast Cancer Study Group (IBCSG) Biological Protocols Working Group. In this substudy, only patients randomized to the tamoxifen alone and letrozole alone arms were included (n = 4922). The database lock took place in 2010, 12 years after trial initiation. Ribonucleic acid (RNA) extracted from FFPE primary breast tumor specimens was available for 1440 patients (29%) considered evaluable for this substudy.

At a Glance
- The clinical validity of the Genomic Grade (GG) quantitative reverse transcriptase polymerase chain reaction previously developed on frozen tissue has not yet been demonstrated.
- The prognostic value of the GG assay was assessed and compared with centrally reviewed immunohistochemical testing of Ki67 and histological grade in the Breast International Group (BIG) 1-98 phase 3 adjuvant trial.
- Either increasing GG or Ki67 was significantly associated with lower distant relapse free intervals and added independent prognostic information to standard clinicopathological factors.
- The GG assay can be used to identify patients with early breast cancer who have excellent prognosis for endocrine therapy only, therefore avoiding unnecessary adjuvant chemotherapy.

Genomic Grade Quantitative Reverse Transcriptase Polymerase Chain Reaction Assay

The development of the GG assay from the original 97 gene GGI was previously presented. Detailed information is provided in the eMethods in the Supplement. Briefly, we selected a subset of genes from the original GGI signature based on their performance in assessing HG and prognosis using 15 independent, publicly available microarray data sets. The GG assay is composed of 9 genes: 6 reporter genes, of which 4 (MCM10, CCNB2, ASPM, PTTG1) are overexpressed in HG1 tumors and 2 (FRY, CX3CR1) are overexpressed in HG1 tumors, and 3 reference genes (GUS, TBP, RPLP0).

We then assessed the concordance between the GG assay and the GGI in 44 paired frozen and FFPE samples identified at the department of pathology in Institut Jules, Bordet, Brussels, Belgium (eMethods in the Supplement). We found a high correlation (r^2 = 0.87) between the results obtained on frozen and FFPE samples. This was subsequently validated in 336 consecutive patients with available FFPE samples diagnosed between January 2004 and December 2010 with early node-negative, ER-positive breast cancer. The GG was successfully determined in 336 samples (96%).

Ribonucleic acid was extracted from 1-mm FFPE tumor cores (at least 50% tumor cellularity) at the Emory University School of Medicine using an FFPE RNA kit (Omega Biotek). According to the manufacturer's instructions, 300 ng of RNA were required, and quantification was performed using nanodrop technology. The RNA from the 1440 evaluable samples was shipped to Ipsogen Laboratories for central evaluation of the GG assay, blinded for clinical information.

Methods

This is a translational research substudy of the BIG 1-98 phase 3 trial. The BIG 1-98 trial compared 5 years of endocrine treatment with tamoxifen alone, letrozole alone, tamoxifen followed by letrozole, or letrozole followed by tamoxifen in 8010 breast cancer patients. Detailed descriptions of the BIG 1-98 trial and patient eligibility criteria was previously published.

In brief, eligible patients had to be postmenapausal with positive endocrine receptor status and either node-positive or node-negative nonmetastatic breast cancer. Informed consent was obtained from all patients at study entry. The study was approved by the ethics committees of all participating sites. This
Normalized gene expression was obtained using the normalized log copy number method (NlogCN) according to the following formula: \( N\text{logCN} = (C_{ti} - \text{intercept})/\text{slope} - \text{Mean}(C_{ti} - \text{intercept})/\text{slope} \), where \( i \) is the \( i^{th} \) reporter gene and \( r \) the \( r^{th} \) reference gene of the GG assay. Results for each sample were reported as both an absolute GG value and an assignment to the GG1 or GG3 category based on the previously defined cutoff.\(^{13}\) We considered cases with a GG falling into the 95% CI around the GG1 or GG3 cut-off as equivocal.

### Pathology Evaluation

Retrospective collection of the FFPE primary tumor specimens for central pathology review was performed as a part of the BIG 1-98 trial.\(^{14}\) Central pathology evaluation was performed at the IBCSG central pathology laboratory blinded for clinical information. This included HG using the Elston-Ellis modification of the Scarff-Bloom-Richardson grading system,\(^{16}\) in addition to ER, progesterone receptors, and Ki67 labeling index using immunohistochemical analysis as previously reported.\(^{17,18}\) The Ki67-low and Ki67-high subgroups were defined using the median as a cut-off (median vs > median). ERBB2 (formerly HER2 or HER2/neu) status was evaluated centrally and samples were considered positive if the ERBB2 gene-to-chromosome 17 ratio was ≥2 by fluorescence in situ hybridization (FISH) or ERBB2 expression was found to be greater than or equal to 3 by immunohistochemistry.\(^{19}\)

### Statistical Methods

All end points were defined according to Hudis et al.\(^{20}\) The primary end point for this analysis was distant recurrence-free interval (DRFI), which is defined as the time from random assignment to the earliest time of distant metastases. The DRFI was censored at either the time of death without a cancer event or the last follow-up.

Secondary end points included invasive disease-free survival (DFS) and overall survival (OS). Disease-free survival was censored at the last follow-up, while OS was censored at the date last known alive. All analyses were performed according to the intention-to-treat principle.

The associations between continuous GG assay or log2-transformed Ki67 and DRFI, DFS, and OS were evaluated using likelihood ratio tests in Cox regression models stratified for chemotherapy use, the 2- vs 4-arm randomization option, and either tamoxifen or letrozole endocrine therapy assignment, with and without adjustment for the clinicopathological model. The clinicopathological model included age (continuous variable), log2 tumor size (continuous variable), nodal status (0 vs 1-3 vs ≥4), ER (percentage of cells as a continuous variable), log2 progesterone receptors (percentage of cells as a continuous variable), centrally reviewed HG (1 vs 2 vs 3) and ERBB2 status (negative vs positive). The clinicopathological model was defined by the linear predictor using these covariates in the stratified Cox model. For subgroup analyses (ie, patients without nodal involvement) the corresponding covariate (eg, nodal status) was omitted from the linear predictor.

The same analysis was performed by replacing the linear predictor of the clinicopathological model with Adjuvant! Online scores\(^3\) in the stratified Cox model. In all cases, the added prognostic value was assessed using the likelihood ratio test statistic.

We calculated the estimated 10-year relapse-free survival by Adjuvant! Online for all patients. For HG, we used the results of the central evaluation. Comorbidities were considered average for age for all patients.

### Results

#### GG Evaluation and Patient Characteristics in the GG Assay Substudy

The GG was successfully obtained in 883 of 1440 evaluable samples (61%). Reasons for nonevaluable samples included poor RNA quantity and/or quality, preanalytical failure for 341 samples (24%), and technical failures related to the robotic handling process in 216 samples (15%) (Figure 1).

Table 1 describes the characteristics of patients included in the GG assay substudy compared with the remaining patients of the tamoxifen and letrozole monotherapy arms of the BIG 1-98 trial.\(^{14}\) Of 883 patients included in the GG assay substudy, 435 patients (49%) were treated with tamoxifen, while the remaining were treated with letrozole. After excluding patients with missing information on centrally performed pathology evaluation, there were more patients with HG3 (27.3% vs 16.4%; \( P < .001 \)), high Ki67 (47.9% vs 28.4%; \( P < .001 \)), progesterone receptor-positive tumors (90.7% vs 59.5%; \( P < .001 \)), and node-negative (59.1% vs 56%; \( P = .07 \)) tumors in the GG assay substudy. The 8-year DRFI was 90% (95% CI, 88%-92%) for the GG assay substudy vs 86% (95% CI, 85%-88%) for the remaining patients (eFigure in the Supplement).

The GG assay classified 878 patients with centrally reviewed HG as either GG1 (36%) or GG3 (24%) (eTable 1 in the Supplement). The remaining patients (40%) were classified as GG equivocal. When we focus on the 502 patients with HG2 tumors, the GG assay classified 40% as GG1 and 16% as GG3, while 44% of HG2 patients were classified as equivocal.
The prognostic value of the different parameters included in the clinicopathological model are provided in Table 2 in the Supplement.

In a univariate analysis including all patients, the GG (as a continuous variable), centrally reviewed Ki67 (as a continuous variable), and the clinicopathological model were all significantly associated with DRFI (Table 2).

In the multivariate analysis, both the GG assay and Ki67 significantly improved the prognostic performance of the clinicopathological model to predict DRFI (Table 2). Similar results were observed when substituting the clinicopathological parameters with Adjuvant! Online scores (eTable 3 in the Supplement). Adding the GG assay to Ki67 appeared to improve the prognostic model, but adding Ki67 to the GG assay did not. However, when either continuous GG assay results or centrally reviewed Ki67 was included in the clinicopathological model, the addition of the other did not add further prognostic information (Table 2).

Survival Analysis

All Patients
The prognostic value of the different parameters included in the clinicopathological model are provided in eTable 2 in the Supplement.

In a univariate analysis including all patients, the GG (as a continuous variable), centrally reviewed Ki67 (as a continuous variable), and the clinicopathological model were all significantly associated with DRFI (Table 2).

In the multivariate analysis, both the GG assay and Ki67 significantly improved the prognostic performance of the clinicopathological model to predict DRFI (Table 2). Similar results were observed when substituting the clinicopathological parameters with Adjuvant! Online scores (eTable 3 in the Supplement). Adding the GG assay to Ki67 appeared to improve the prognostic model, but adding Ki67 to the GG assay did not. However, when either continuous GG assay results or centrally reviewed Ki67 was included in the clinicopathological model, the addition of the other did not add further prognostic information (Table 2).

Similar analyses were performed looking into the association between the GG assay, Ki67, and the secondary survival end points DFS and OS (eTable 4 in the Supplement). Univariate analysis indicated that both the GG assay and Ki67 were significantly associated with DFS and OS, yet were not able to add important prognostic information to clinicopathological results. Of note, in the absence of the clinicopathological model, adding the GG assay to Ki67 marginally improved the prediction of DFS ($\chi^2 = 3.6; P = .06$) and OS ($\chi^2 = 6.0; P = .01$) but not the reverse.

Node-Negative Patients Treated With Endocrine Therapy Alone (N = 467)
We analyzed the prognostic value of the centrally reviewed HG, GG, and Ki67 on DRFI in node-negative patients who received only endocrine therapy as adjuvant systemic therapy.
The Genomic Grade Assay for Breast Cancer Recurrence

Original Investigation  Research

With Ki67 ≤13% or less are considered Ki67-low, and patients with tumors with Ki67 >13% are considered Ki67-high. All patients were treated with endocrine therapy alone. A similar outcome was observed for patients with Ki67 greater than 13% are considered Ki67-high. Eq indicates equivocal; GG, Genomic Grade; HG, histological grade.

Table 2. Added Value of GG Assay and Ki67 as Continuous Variables Predicting DRFIa

<table>
<thead>
<tr>
<th>Analysis</th>
<th>No. Patients</th>
<th>Events</th>
<th>Likelihood Ratio (χ²)</th>
<th>P Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Univariate</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CP model alone</td>
<td>848</td>
<td>79</td>
<td>45.50</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>GG alone</td>
<td>883</td>
<td>84</td>
<td>21.71</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Ki67 alone</td>
<td>862</td>
<td>78</td>
<td>16.86</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Multivariate</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CP + GG vs CP</td>
<td>848</td>
<td>79</td>
<td>4.59</td>
<td>.03</td>
</tr>
<tr>
<td>CP + Ki67 vs CP</td>
<td>833</td>
<td>75</td>
<td>5.35</td>
<td>.02</td>
</tr>
<tr>
<td>Ki67 + GG vs Ki67</td>
<td>862</td>
<td>78</td>
<td>8.20</td>
<td>.004</td>
</tr>
<tr>
<td>GG + Ki67 vs GG</td>
<td>862</td>
<td>78</td>
<td>3.20</td>
<td>.07</td>
</tr>
<tr>
<td>CP + GG + Ki67 vs CP + GG</td>
<td>833</td>
<td>75</td>
<td>2.38</td>
<td>.12</td>
</tr>
<tr>
<td>CP + GG + Ki67 vs CP + Ki67</td>
<td>833</td>
<td>75</td>
<td>1.74</td>
<td>.18</td>
</tr>
</tbody>
</table>

Abbreviations: CP, clinicopathological model; DRFI, distant recurrence-free interval; ER, estrogen receptor; GG, Genomic Grade; PgR, progesterone receptor.

* Linear predictor of Cox regression model including age, tumor size, nodal status, histological grade, ER, PgR, and ERBB2 stratified for chemotherapy use, the 2- vs 4-arm randomization option, and endocrine therapy assignment (tamoxifen or letrozole).

(Figure 2). The 10-year DRFI was 100% (95% CI, 100%-100%) for patients with HG1, 94% (95% CI, 91%-97%) for patients with HG2, and 90% (95% CI, 84%-96%) for patients with HG3. The 10-year DRFI was 99% (95% CI, 97%-100%) for patients with GG1, 94% (95% CI, 90%-98%) for patients with GG2, and 97% (95% CI, 80%-94%) for patients with GG3. The 10-year DRFI was 97% (95% CI, 94%-99%) for patients with Ki67-low tumors (Ki67 ≤13%) and 91% (95% CI, 86%-95%) for patients with Ki67-high tumors (Ki67 >13%). All 3 tests were able to define distinct groups of patients with different prognoses. The GG assay was able to identify 38% of patients with risk of distant recurrence at 10 years of only 1% with endocrine therapy alone. A similar outcome was observed for patients with HG1, but these were only 18% of patients.

Figure 3

All patients were treated with endocrine therapy alone. Patients with tumors with Ki67 13% or less are considered Ki67-low, and patients with tumors with Ki67 greater than 13% are considered Ki67-high. Eq indicates equivocal; GG, Genomic Grade; HG, histological grade.

Discussion

Our study indicated that both the GG assay and centrally reviewed Ki67 can add relevant prognostic information to classic clinicopathological factors in determining risk of distant recurrence. The association between the GG assay and outcome appeared to be more robust in predicting distant recurrences rather than DFS, which is highly relevant given that distant relapses remain the main reason adjuvant systemic therapy is prescribed. In addition, the GG assay was able to identify a relatively large group of node-negative patients with just 1% risk of distant recurrence at 10 years when treated with endocrine...
therapy alone. Also, it showed higher discriminative power to identify distinct groups with different prognosis within patients with HG2 compared to Ki67.

In the past few years, several other genomic tests have been retrospectively tested in patients treated in large randomized trials. They showed similar findings that a genomic test could add relevant prognostic information to classic prognostic parameters. Our study provides evidence that the GG assay is another tool that can be considered.

The clinical use of genomic tests has been a subject of debate over the past few years, and there is skepticism regarding whether the available data are enough to allow these tools in clinical practice to determine the need for adjuvant chemotherapy. In our study, patients with GG, node-negative tumors had a 10-year DRFI of 99% (95% CI, 97%-100%) with just 5 years of endocrine therapy. Thus, the addition of chemotherapy is unlikely to provide any meaningful added benefit in these patients, especially with respect to associated morbidity. Of note, similar results were observed at 10 and 15 years with Endopredict and PAM50 ROR. However, the GG assay is one of only 3 tests, alongside the Oncotype Dx assay and MammaPrint assay, where clinical use is being investigated in a prospective phase 3 trial. The ASTER70s trial (NCT01564056) has a target accrual of around 2000 patients and aims to determine the benefit of adjuvant chemotherapy in elderly patients according to GG assay values. The GG assay is the only genomic test that addresses such questions in the elderly population where several competing morbidities exist, and the absolute value of adjuvant chemotherapy is always in question. Acknowledging that around 40% of patients with breast cancer are 65 years or older and that women older than 70 years have the highest risk of developing breast cancer, the GG assay could emerge as an important tool to aid treatment decisions in daily practice.

Our study has various strengths. Patients were treated in a large phase 3 randomized trial with adequate prospective data collection. It is one of few studies that provide a direct comparison between a genomic test and centrally reviewed Ki67. While the addition of the GG assay to the clinicopathological model and Ki67 did not add significant prognostic information, it is important to note that other genomic tests like the MammaPrint assay and Breast Cancer Index were not previously investigated against Ki67. Nevertheless, acknowledging the variability of Ki67 even among expert laboratories and the uncertainty regarding the cutoff to define highly proliferative tumors, it is becoming more challenging to rely on Ki67 for daily decision making. The GG assay can thus provide an extra level of certainty in the decision-making process.

On the other hand, a limitation of this study is the high failure rate. This was largely owing to poor RNA quantity or quality of the collected samples. Of note, no similar failure rates were observed in the testing data set or more importantly in the ongoing prospective ASTER70s trial (NCT01564056) where the GG assay was not determined in only 10 out of the first 1272 included patients (Dr Etienne Brain, MD, PhD, ASTER70s primary investigator, email communication, March 2015).

Conclusions

This is the first validation of the prognostic ability of the GG assay in a prospective randomized phase 3 trial. The test showed that it can refine the prediction of risk of distant recurrence and is able to identify a larger proportion of patients who have excellent DRFI at 10 years with just 5 years of endocrine therapy compared with centrally reviewed immunohistochemical testing of Ki67 and HG. The feasibility of performing the assay on FFPE samples, and the ongoing testing in a randomized phase 3 trial, would make the case for the GG assay emerging as a useful genomic tool for clinical use.
The Genomic Grade Assay for Breast Cancer Recurrence

Original Investigation Research

Role of the Funder/Sponsor: The sponsors had no role in the design and conduct of the study; collection, management, analysis, and interpretation of the data; preparation, review, or approval of the manuscript; and decision to submit the manuscript for publication.

Additional Contributions: We would like to thank Hélène Peyro-Saint-Paul, MD, and Fabienne Hermite, PhD, for their contributions. We would also like to thank the women, physicians, nurses, and data managers who participated in this clinical trial; the many pathologists who submitted tumor blocks, the BIG1-98 Steering Committee; Novartis for funding of the clinical trial and of the collection of tumor blocks; the IBCSG for the design of the trial, coordination, data management, medical review, statistical support; the IBCG central pathology office for collection and processing of tumor blocks; and the BIG-98 Collaborative Group. We also thank the IBCSG for sharing data and collaborating on this study.

REFERENCES


