Human Pathology (2021) 108, 51–59

Human PATHOLOGY

www.elsevier.com/locate/humpath

Original contribution

Breast cancer tumor heterogeneity has only little impact on the estimation of the Oncotype DX® recurrence score using Magee Equations and Magee Decision Algorithm™☆

Annabelle Remoué MD[a], Virginie Conan-Charlet MD[a], Laura Deiana MD[b], Alexandra Tyulyandina MD[c], Pascale Marcorelles MD, PhD.[a], Ulrike Schick MD, PhD[c], Arnaud Uguen MD, PhD[a,d,*]

[a] *CHRU Brest, Department of Pathology, Brest, F-29220, France*
[b] *CHRU Brest, Institute of Oncology and Hematology, Brest, F-29220, France*
[c] *CHRU Brest, Department of Radiotherapy, Brest, F-29220, France*
[d] *Univ Brest, Inserm, CHU de Brest, LBAI, UMR1227, Brest, France*

Received 16 October 2020; accepted 15 November 2020

Available online 24 November 2020

Keywords:
Breast cancer;
Oncotype DX® assay;
Magee Equations;
Tumor heterogeneity;
Tissue microarray

Summary Oncotype DX® assay is used to guide therapeutic decisions in early-stage invasive breast carcinoma but remains expensive. Magee Equations (MEs) and Magee Decision Algorithm (MDA) predict the Oncotype DX® recurrence score (RS) on the basis of histopathological parameters. The influence of intratumor heterogeneity on MEs and MDA remains uncertain. We compared Ki-67, estrogen and progesterone receptors, and human erb-b2 receptor tyrosine kinase 2 (HER2) status on tissue microarray cores with the corresponding findings on the whole slides to calculate MEs scores and to decide if Oncotype DX® testing was required as per MDA in two sets of 175 and 59 tumors, without and with Oncotype DX® results, respectively. Agreements in the interpretation of Ki-67, estrogen and progesterone receptors, and HER2 status were very good between limited areas and whole-slide analyses. This resulted also in very good agreements about the results of MEs and MDA. For 7 of 175 (4%) and 3 of 59 (5.1%) cases, MEs and MDA results in different tumor areas would have changed the indication to perform or not perform Oncotype DX® assays. Oncotype DX® RSs were significantly correlated with MEs and MDA results, but among cases initially predicted to have an RS $\leq$25 using MDA, 3 of 34 cases (8.8%) had in fact an RS >25. Tumor heterogeneity appears to have little impact on the estimation of the Oncotype DX® RS using MEs and MDA and would have permitted to avoid half of Oncotype DX® assays in our series. Caution is nevertheless required in

☆ Disclosures: None.

* Corresponding author. Department of Pathology, University Hospital Morvan, 5, Foch Avenue, 29609, Brest, France
E-mail address: arnaud.uguen@chu-brest.fr (A. Uguen).

https://doi.org/10.1016/j.humpath.2020.11.006

discarding Oncotype DX® assay in cases with ME scores >18 associated with low mitotic activity.

1. Introduction

Breast cancers are among the most common malignancies in the world, and invasive breast carcinoma remains a frequent cause of cancer-related death [1,2]. Decades of breast cancer−dedicated research has led to major therapeutic advances. Surgery, radiotherapy, chemotherapy, endocrine therapy, targeted therapies, and immunotherapies are now available in patients with various types and stages of breast cancers, and different therapeutic schemes are used depending on clinical, pathological, and molecular parameters. Pathological examination of breast cancer samples is crucial to guide treatment. Indeed, it allows completion of the diagnostics of invasive carcinoma and determination of its histopathological subtype as well as histoprognostic and theranostic parameters based on morphological features (eg, Nottingham prognostic score based on gland formation, nuclear pleomorphism, and mitotic count per 10 high-power microscopic fields), immunohistochemistry (IHC) analyses (eg, Ki-67 index of prognostic significance, expression of estrogen receptor [ER], progesterone receptor [PR], and hormone receptors [HRs], and human erb-b2 receptor tyrosine kinase 2 [HER2] protein expression by tumor cells), and molecular analyses (eg, *HER2* gene copy quantification using *in situ* hybridization [ISH] in case of equivocal HER2 IHC and multigene expression reverse-transcriptase polymerase chain reaction [RT-PCR] tests for prognostic significance and therapeutic decision) [3−10].

Multigene expression RT-PCR prognostic assays are used to guide decisions on adjuvant systemic chemotherapy in women with HR-positive and HER2-negative early-stage invasive breast carcinomas on the basis of the estimation of their risk of locoregional and distant recurrences and survival. Only the Oncotype DX® (Genomic Health, Inc., Redwood City, CA) 21-gene assay has been clinically validated for predicting the benefit of adding adjuvant chemotherapy to endocrine therapy to further reduce the risk of recurrence based on the calculation of a *recurrence score* (RS). Indeed, in patients with pT1b/pT1c or pT2, lymph node−negative, HR-positive, HER2-negative tumors, a low RS (RS < 11) was not associated with any benefit of adding chemotherapy to endocrine therapy, whereas patients with a high RS (>30) derived a clear benefit from adjuvant chemotherapy. The benefit of adjuvant chemotherapy was also significant in case of an intermediate RS in postmenopausal women aged ≤50 years. The value of the RS to predict benefit of chemotherapy plus endocrine therapy versus endocrine therapy alone in patients with lymph node−positive, HR-positive, HER2-negative tumors has also been demonstrated (very small benefit in case of a low RS and obvious benefit in case of a high RS) [11−17]. In this manner, Oncotype DX® assay is now frequently used by oncologists to guide the administration of chemotherapy. Nevertheless, this assay remains expensive (about $3000 per analysis in the USA and €1850 in France) and often requires outsourcing of molecular analysis to expert centers [18,19]. For this reason, alternative methods that would provide an RS correlated with Oncotype DX® assay, ideally nonexpensive and easy to implement in pathology laboratories, are needed to decrease the number of multigene expression RT-PCR analyses.

In this attempt, multivariate models called Magee Equations (MEs) have been developed to estimate the Oncotype DX® assay RS on the basis of routinely reported histopathology and biomarkers. Multiple linear regression analysis was performed to model the prediction of the Oncotype DX® assay RS by histological Nottingham score, immunohistochemical Ki-67 index, tumor size (in centimeters), H-scores for ER (IHC) and PR (IHC), and HER2 status (negative, equivocal, or positive, IHC and ISH). Three models were built based on different hypotheses and data availability: the first regression model included all available parameters (including the Ki-67 index) for prediction of the Oncotype DX® RS (ME1). The second regression model was similar to the first but did not include Ki-67 (ME2). The third regression model included only semiquantitative immunohistochemical expression levels for ER, PR, HER2, and Ki-67 (ME3) [18−25]. This latter ME3 has been notably demonstrated to be of predictive value to predict pathological complete response to neoadjuvant chemotherapy in ER-positive, HER2-equivocal, or HER2-negative breast tumors [24]. An algorithm (called Magee Decision Algorithm™ [MDA]) using MEs and tumor mitotic activity has been recently proposed to safely avoid about 70% of Oncotype DX® assays in patients with ME scores <18 or between 18 and 25 with a mitotic score of 1 (ie, Oncotype DX® assay RS predicted to be ≤25) and in patients with ME scores ≥31 (ie, Oncotype DX® assay RS predicted to be >25), with no consequence in terms of patients' outcomes [18,25].

Given that predicting the results of the Oncotype DX® assay RS and avoiding several assays on the basis of this prediction appear feasible using analysis of routine histopathological and immunohistochemical/ISH breast cancer slides, the potential influence of tumor heterogeneity within breast cancer tumors on the results of MEs may be

of crucial consequences for making molecular testing and therapeutic choices in patients with breast cancers [26−31]. By sampling only part of a tumor/slide section specimen, tissue microarray (TMA) consists in a tissue analysis method that is intrinsically limited by potential tumor heterogeneity. This limitation makes also TMA a usable way to compare randomly selected areas or areas selected on specific pathological criteria with whole-tissue sections. Estimating the consequences of tumor heterogeneity explored through TMA approaches in terms of Ki-67 IHC, ER IHC, PR IHC, and HER2 analyses on the results of MEs and the prediction of the Oncotype DX® assay RS through MDA is precisely the aim of the present study.

2. Materials and methods

2.1. Case selection

Cases included were patients with invasive breast adenocarcinomas operated at the Brest University Hospital (France) from 2010 to 2020.

A first set of samples not associated with Oncotype DX® analyses was included to build TMA blocks using the 3DHistech TMA Grand Master automated tissue microarrayer (3DHISTECH, Budapest, Hungary), with 2 spots (500-µm-diameter cores) randomly selected within one donor formalin-fixed paraffin-embedded (FFPE) tumor block per patient. For this first set, in each patient, data about these 2 spots were compared with those of the whole tumor as mentioned at the initial 0diagnosis in terms of Ki-67 IHC, ER IHC, PR IHC, and HER2 analyses and ME-MDA results.

A second set of patients with available Oncotype DX® results were included. The Ki-67 IHC, ER IHC, PR IHC, and HER2 diagnostic slides from the same FFPE tumor block that served for Oncotype DX® assay were digitalized using a 3DHistech Panoramic Midi scanner (3DHistech). CaseViewer software (3DHistech) was used for simulating two 500-µm-diameter TMA cores within tumor tissue on the basis of Ki-67 immunohistochemical digitalized slides (*Minimal Ki-67 proliferative area* [MiniKi-67] and *Maximal Ki-67 proliferative area* [MaxiKi-67]) each time that substantial difference between two tumor areas was obvious (otherwise, two random tumor areas with the same Ki-67 proliferative index were selected). *MiniKi-67* and *MaxiKi-67* areas were then reported on HER2, ER, and PR immunohistochemical digital slides. Data about MiniKi-67 and MaxiKi-67 areas as well as whole slides were compared in terms of ME-MDA results and correlated with the Oncotype DX® RS (manual quantification of tissue sections with no automatic analysis).

Clinical and pathological data and, when available, Oncotype DX® RS results (outsourced analyses) of cases analyzed for diagnostic purposes (notably tumor sizes and Nottingham scores required for ME calculation in addition to Ki-67 IHC, ER IHC, PR IHC, and HER2 analysis results) were collected from pathological reports. All patients provided a written informed consent for research use of their tumor samples that were included in a registered tumor tissue collection and used in compliance with our institutional guidelines (CHRU Brest, CPP n° DC − 2008−214).

2.2. IHC and bright-field ISH analyses

IHC and ISH analyses were processed on 3-µm-thick TMA slides and diagnostic whole slides using the Ventana Benchmark Ultra® automated slide preparation system (Roche Diagnostics, Meylan, France) using anti-ER (prediluted, clone SP1; Roche Diagnostics), anti-PR (prediluted, clone 1E2; Roche Diagnostics), anti-HER2 (prediluted, clone 4B5; Roche Diagnostics), and anti-KI-67 (1:50 dilution, clone MIB1; Dako, Glostrup, Denmark) antibodies and the Inform Cocktail HER2/chr17 Probe kit (Roche Diagnostics).

For the different areas and slides, Ki-67 analyses were interpreted in terms of percentage of stained tumor nuclei within tumor areas (0−100), and ER and PR IHC *H-scores* were calculated on the basis of the percentage of stained cells (0−100) and their intensity of staining (0 = none, 1 = weak, 2 = moderate, 3 = strong staining intensities), and HER2 status was interpreted as negative (score = 0, score = 1+, or score = 2+ using HER2 IHC with gene copies <4 per cell using ISH), positive (score = 2+ using HER2 IHC with gene copies >6 per cell using ISH or score 3+ using IHC), or equivocal (score = 2+ using HER2 IHC with gene copies between 4 and 6 per cell using ISH).

2.3. ME result calculations and statistical analyses

ME score calculations were performed as follows as published by the Department of Pathology of University of Pittsburgh Medical Center (built-in calculation tool at https://path.upmc.edu/onlineTools/mageeequations.html): ME1 score = 15.31385 + Nottingham score*1.4055 + ER IHC*(-0.01924) + PR IHC*(-0.02925) + (0 for HER2 negative, 0.77681 for HER2 equivocal, 11.58134 for HER2 positive) + tumor size (cm)*0.78677 + Ki-67 index*0.13269; ME2 score = 18.8042 + Nottingham score*2.34123 + ER IHC*(-0.03749) + PR IHC*(-0.03065) + (0 for HER2 negative, 1.82921 for HER2 equivocal, 11.51378 for HER2 positive) + tumor size (cm)*0.04267; ME3 score = 24.30812 + ER IHC*(-0.02177) + PR IHC*(-0.02884) + (0 for HER2 negative, 1.46495 for HER2 equivocal, 12.75525 for HER2 positive) + Ki-67*0.18649 [23]. MDA was applied to class each area analysis result in terms of class I result (Oncotype DX® assay RS predicted to be ≤25 and multigene assay not needed, ie, three ME scores <18 or at least one ME score between 18 and 25 with a mitotic score of 1), class III result (Oncotype DX® assay RS predicted to be >25 and multigene assay

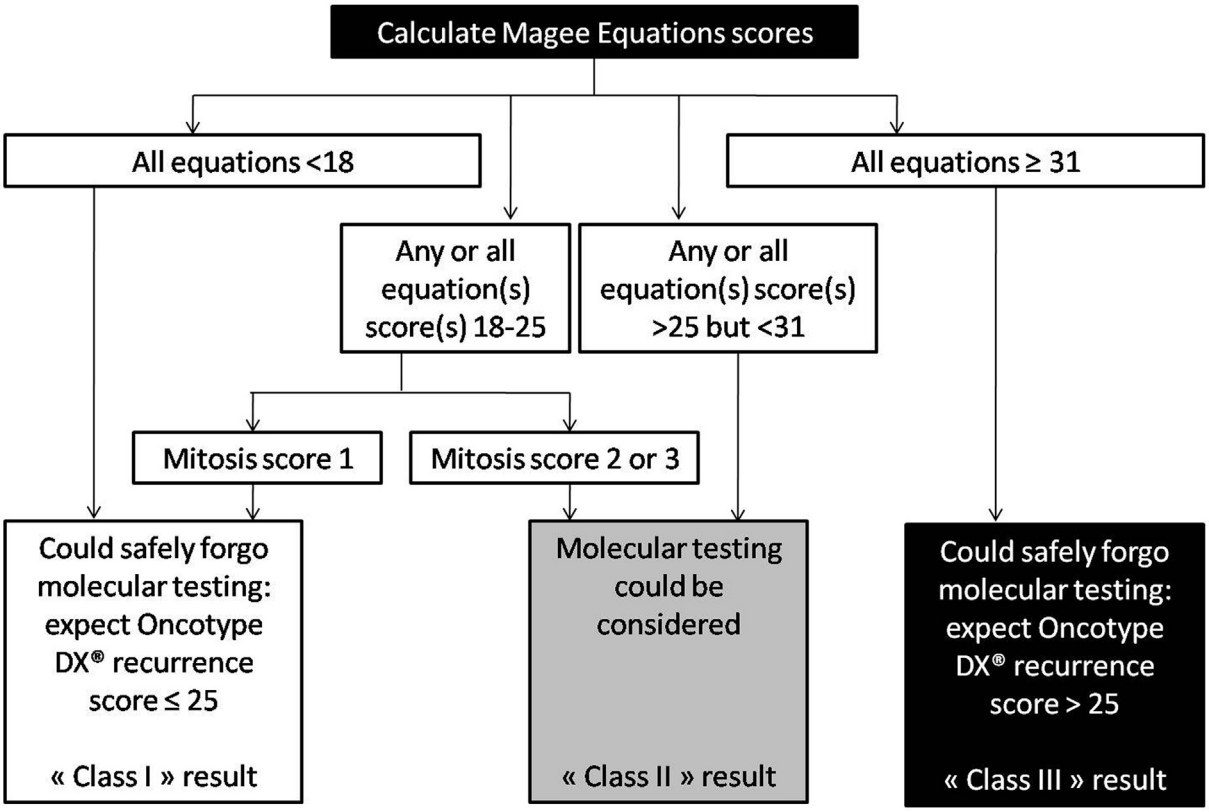

Fig. 1 Adapted Magee Decision Algorithm™ with classes I to III results (adapted from the study by Bhargava et al. [18,25]).

not needed, ie, three ME scores ≥31), or class II result (Oncotype DX® assay could be considered, ie, at least one ME score between 18 and 25 with a mitotic score of 2 or 3 or at least one ME score >25 but <31) [18,25]. See Fig. 1 for adapted MDA.

Statistical analyses were performed using MedCalc Statistical Software, version 13.2.2 (MedCalc Software bvba, Ostend, Belgium; http://www.medcalc.org; 2014). The level of significance was set at $P < 0.05$. The weighted Kappa statistic test (quadratic weights) was used to quantify the agreement for the HER2 status (negative, equivocal, positive) and the class I, II, and III MDA results. Intraclass correlation coefficients (ICCs; for single measures reflecting the reliability of the ratings for one, typical, single observer) were used to calculate the agreement between the measurements of KI-67 IHC, ER IHC, and PR IHC analyses and the ME scores. Comparisons were performed between whole-tumor slide diagnostic data and TMA core data (first tumor set) and between whole-slide data and MiniKi-67 and MaxiKI-67 areas (second tumor set). The values of Kappa strength agreements and ICCs were interpreted as follows: <0.20, poor; 0.21−0.40, fair; 0.41−0.6, moderate; 0.61−0.80, good; and 0.81−1.00, very good agreement [32]. For the second set of tumors with available Oncotype DX® results, Spearman's rank correlation test was used to study the correlation between ME scores and the Oncotype DX® RS, and the percentages of discrepant cases between MDA classes and the Oncotype DX® RS (ie, cases with ME class I result but Oncotype DX® RS >25 and cases with ME class III result but Oncotype DX® RS ≤25) were calculated on the basis of whole-slide data, MiniKi-67 area, and MaxiKI-67 area analyses, as well as the percentage of class discrepancies between whole-slide data, MiniKi-67 area, and MaxiKI-67 area analyses.

3. Results

3.1. Cases included

The first set of cases included 175 women, with a mean age of 59.3 years (range = 30−88 years) with invasive breast cancers, mostly invasive carcinomas of no special type (82.9%), from stage IA to IIIC at initial diagnosis. None of these cases operated between 2010 and 2014 had an Oncotype DX® test.

The second set of cases included 59 women operated between 2017 and 2020 (mean age of 58.4 years, range = 34−73 years) for invasive breast cancer, mostly invasive carcinomas of no special type (86.4%), staged IA to IIB. These patients were considered because they had Oncotype

DX® tests with available results. Among these Oncotype DX® test results, 47 (79.7%) had an RS ≤25, and 12 (20.3%) had an RS >25.

See Table 1 for summary and supplementary data for details about clinical, pathological, and molecular features of the two sets of tumors.

3.2. Agreement between TMA cores and whole-slide analyses in the first set of tumors and impact on ME score calculation

Among cases from the first set of tumors, agreements between TMA-based and whole slides were very good for

Table 1 Summary of the clinical, pathological, and molecular parameters of the two sets of patients.

Parameters	First set of patients (tissue microarray included) (n = 175)	Second set of patients (with Oncotype DX® analyses) (n = 59)
Age (mean, range) (years)	59.3 (30–88)	58.4 (34–73)
Laterality of breast cancer (%)		
Right	83 (47.4%)	31 (52.5%)
Left	92 (52.6%)	28 (47.5%)
Histological subtype (%)		
Invasive carcinoma of no special type (NST)	145 (82.9%)	51 (86.4%)
Invasive lobular carcinoma	20 (11.4%)	7 (11.9%)
Mixed invasive NST and lobular carcinoma	2 (1.1%)	0
Invasive micropapillary carcinoma	2 (1.1%)	1 (1.7%)
Mucinous carcinoma	2 (1.1%)	0
Tubular carcinoma	1 (0.6%)	0
Metaplastic carcinoma	2 (1.1%)	0
Neuroendocrine carcinoma	1 (0.6%)	0
Stage (%)		
IA	79 (45.1%)	39 (66.1%)
IB	9 (5.1%)	6 (10.2%)
IIA	46 (26.3%)	9 (15.3%)
IIB	22 (12.6%)	5 (8.5%)
IIIA	15 (8.6%)	0
IIIB	2 (1.1%)	0
IIIC	2 (1.1%)	0
Nottingham grade (%)		
Grade I (scores 3, 4, or 5)	41 (23.4%)	7 (11.9%)
Grade II (scores 6 or 7)	86 (49.1%)	47 (79.7%)
Grade III (scores 8 or 9)	48 (27.4%)	5 (8.5%)
Estrogen receptors (%)		
Negative (ie, weak to no labeling <1% of tumor cells)	25 (14.3%)	0
Positive (ie, at least weak labeling of ≥1% of tumor cells)	150 (85.7%)	59 (100%)
Progesterone receptors (%)		
Negative (ie, weak to no labeling <1% of tumor cells)	31 (17.7%)	9 (15.3%)
Positive (ie, at least weak labeling of ≥1% of tumor cells)	144 (82.3%)	50 (84.7%)
Ki-67 proliferation index (%)		
<20%	93 (53.1%)	24 (40.7%)
≥20%	82 (46.9%)	35 (59.3%)
HER2 status		
Negative	150 (85.7%)	58 (98.3%)
Equivocal	5 (2.9%)	1 (1.7%)
Positive	20 (11.4%)	0
Oncotype DX® analyses (%)		
RS ≤ 25	–	47 (79.7%)
RS > 25	–	12 (20.3%)

Abbreviations: RS, recurrence score; HER2, human erb-b2 receptor tyrosine kinase 2.

every IHC analyses and also for ME score calculation with minimal (albeit very good) agreement for Ki-67 analyses (ICC = 0.9085; see Table 2 for details). For 7 (4%) cases, the interpretation of ME scores on the basis of TMA or whole-slide data resulted in a different class status following the MDA (4 changes between class I and class II in cases 1#039, 1#064, 1#097, and 1#122; 3 changes between class II and class III in cases 1#135, 1#144, and 1#169) and would have led to different decisions in terms of performing (ie, class II interpretation) or not performing (ie, class I or class III interpretation) a molecular Oncotype DX® test (see Fig. 2A for summary).

3.3. Impact of heterogeneous proliferative activity on ME scores and correlation with the Oncotype DX® RS in the second set of tumors

Among cases from the second set of tumors, agreements between MiniKi-67 areas, MaxiKi-67 areas, and whole-slide analyses remained very good for ME score calculation despite voluntarily trying to maximize the heterogeneity of the Ki-67 proliferative index by selecting the minimally (MiniKi-67) and maximally (MaxiKi-67) proliferative areas within the tumor sections. As in the first set of tumors, agreements remained also very good for ER IHC, PR IHC, and HER2 status (see Table 3 for details). Comparing MDA class results between MiniKi-67 and whole slides, between MaxiKi-67 and whole slides, and between MiniKi-67 and MaxiKI-67 area analyses resulted in 3 (5.1%, cases 2#12, 2#13, and 2#42), 3 (5.1%, cases 2#12, 2#13, and 2#48), and 2 (3.4%, cases 2#42 and 2#48) class changes, respectively (all between class I and class II). Class I results (ie, results that could safely avoid molecular testing because of the expected Oncotype DX® RS predicted to be ≤25) were obtained in 34, 32, and 32 cases with MiniKi-67, MaxiKi-67, and whole-slide data, respectively.

The results of the 3 MEs about MiniKi-67, MaxiKi-67, and whole-slide data were all significantly correlated with the Oncotype DX® RS ($P < 0.0001$, detailed data not shown). Among cases with class I results as per MDA, 3 of 34 cases (8.8%, cases 2#48, 2#53, and 2#57) had a Oncotype DX® RS >25 as per MiniKi-67 data, with 2 of 32 cases (6.25%, cases 2#53 and 2#57) with MaxiKi-67 data and whole-slide data (see Fig. 2B for summary). All of these discrepant cases had the three ME results >18 and were classified as class I results because of score 1 mitotic activity.

Table 2 Agreements between TMA core− and whole-slide−based analyses in the first set of tumors.

Parameters	Kappa/intraclass correlation coefficient (ICC) [95% CI]	Agreement
Ki-67	ICC = 0.9085 [0.8786−0.9313]	Very good
Estrogen receptor H-score	ICC = 0.9501 [0.9333−0.9627]	Very good
Progesterone receptor H −score	ICC = 0.9749 [0.9664−0.9813]	Very good
HER2 status	Kappa = 1.0	Very good
Magee Equation 1 score	ICC = 0.9881 [0.9840−0.9912]	Very good
Magee Equation 2 score	ICC = 0.9889 [0.9851−0.9918]	Very good
Magee Equation 3 score	ICC = 0.9812 [0.9747−0.9860]	Very good
Magee Decision Algorithm™ class	Kappa = 0.95 [0.913−0.986]	Very good

Abbreviations: HER2, human erb-b2 receptor tyrosine kinase 2; CI, confidence interval.

4. Discussion

Given, on the one hand, the need to provide accurate prognostic data to guide therapeutic decisions in patients with breast cancers and, on the other hand, the high cost of multigene expression RT-PCR prognostic assays such as the Oncotype DX® ones, attempts to avoid unnecessary tests in some patients for whom prognostic information can be derived from other costless pathological parameters routinely collected remain interesting. Among these attempts, ME calculation interpreted through MDA is known to avoid about 70% of molecular Oncotype DX® assays by predicting the RS ≤25 or >25 [18]. In our study, no tumor analyzed using Oncotype DX® assay had a MDA-predicted RS >25 (ie, 3 ME scores >31), but 32−34 tumors had a MDA class I result predicting the RS ≤25 among 59 tumors analyzed in clinical practice. In this manner, choosing to perform or not perform Oncotype DX® assay on the basis of the MDA result would have permitted to avoid 54−57% of Oncotype DX® assays, consisting in an economy of €59,200−€62,900 on €109,150 of the total cost for the 59 assays (as per the cost of this assay in the French healthcare system).

Nevertheless, 2−3 patients would not have been selected as requiring Oncotype DX® assays because of MDA predicting the RS ≤25 and would not have received accurate prognostic and therapeutic decision because of a calculated Oncotype DX® RS >25. As a result, avoiding some tests in some patients would result in inadequate therapeutic management in 6.25−8.8% of patients in our series. For this reason, great caution is required when using MDA as guidance to perform or not perform an Oncotype DX® assay. Of note, in our series, discrepancies between the Oncotype DX® RS and MDA prediction were only observed in some cases with 2−3 ME scores between 18 and 25 and a mitotic score of 1 as per the Nottingham grading system. Further studies will be necessary to decrease the proportion of cases with false estimation of

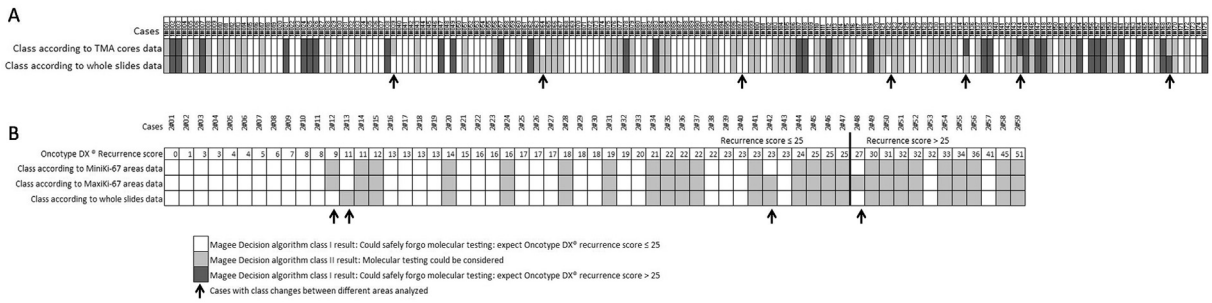

Fig. 2 Graphical summary of the Magee Decision Algorithm™ results in the first set of tissue microarray (TMA)−included tumors (TMA cores and whole-slide data) (A) and in the second set of Oncotype DX®−tested tumors with areas selected on different Ki-67 proliferative indices (MiniKI-67, MaxiKi-67, and whole-slide data) (B).

Oncotype DX ® RS $\leq$ 25 among cases with MEs results >18 and low mitotic activity.

In our series, we only observed little influence of tumor heterogeneity on the evaluation of ER IHC, PR IHC, and Ki-67 IHC analyses as well as HER2 status and, as a consequence, on the calculation of the ME scores and their integration in MDA, with only 3−5% of class changes from one interpretation to another. A limitation in our study is that we did not evaluate the influence of tumor heterogeneity on the evaluation of the Nottingham score and of the histopathological measure of the maximal tumor diameter on Hematoxylin and Eosin (HE) tissue sections as we only focused on complementary IHC (and when required ISH) data. Further study would be necessary to confirm the consistency of the interpretation of these parameters, but literature data already point high reproducibility in this field [33,34].

5. Conclusion

MEs and MDA reflect how morphological histopathological parameters can help to predict cancer prognosis and molecular status without requiring expensive molecular assays. At the era of digital pathology and increasing *pathomics* applications, the use of morphological data can represent a major source of inexpensive parameters to better stratify breast cancer for diagnostic, prognostic, and theranostic purposes in future [35−38]. The choice of analyzing one to all tumor tissue sections in this field with potential associated tumor heterogeneity−related issues

Table 3 Agreements between selected areas on proliferative index− and whole-slide−based analyses in the second set of tumors.

Parameters	MiniKi-67 versus MaxiKi-67 areas		MiniKi-67 area versus whole slide		MaxiKi-67 area versus whole slide	
	Kappa/intraclass correlation coefficient (ICC) [95% CI]	Agreement	Kappa/ICC [95% CI]	Agreement	Kappa/ICC 95% CI]	Agreement
Ki-67[a]	ICC = 0.5136 [0.2984−0.6792]	Moderate[a]	ICC = 0.7061 [0.5508−0.8142]	Good[a]	ICC = 0.8160 [0.7088−0.8863]	Very good[a]
Estrogen receptor H-score	ICC = 0.9297 [0.8846−0.9576]	Very good	ICC = 0.9256 [0.8780−0.9551]	Very good	ICC = 0.8985 [0.8350−0.9384]	Very good
Progesterone receptor H-score	ICC = 0.9641 [0.9404−0.9785]	Very good	ICC = 0.9466 [0.9118−0.9679]	Very good	ICC = 0.9660 [0.9434−0.9796]	Very good
HER2 status	Kappa = 1.0	Kappa = 1.0	Kappa = 1.0	Kappa = 1.0	Kappa = 1.0	Kappa = 1.0
Magee Equation 1 score	ICC = 0.9415 [0.9035−0.9648]	Very good	ICC = 0.9483 [0.9146−0.9689]	Very good	ICC = 0.9524 [0.9213−0.9714]	Very good
Magee Equation 2 score	ICC = 0.9647 [0.9414−0.9789]	Very good	ICC = 0.9536 [0.9231−0.9721]	Very good	ICC = 0.9580 [0.9304−0.9748]	Very good
Magee Equation 3 score	ICC = 0.8974 [0.8333−0.9377]	Very good	ICC = 0.9283 [0.8823−0.9567]	Very good	ICC = 0.9311 [0.8868−0.9585]	Very good
Magee Decision Algorithm™ class	Kappa = 0.931 [0.838−1.0]	Very good	Kappa = 0.896 [0.782−1.0]	Very good	Kappa = 0.897 [0.784−1.0]	Very good

Abbreviations: HER2, human erb-b2 receptor tyrosine kinase 2; CI, confidence interval.

[a] Areas with different Ki-67 staining were selected whenever it was feasible, trying to maximize the difference of the Ki-67 proliferative index between *MiniKi-67* and *MaxiKi-67* areas.

will remain a key for the development of these new image-based digital analyses and will still require validation studies beyond our present work, in which tumor heterogeneity about investigated parameters appeared minimal and of little influence to the results of MDA.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.humpath.2020.11.006.

Acknowledgments

The authors would like to acknowledge surgeons, gynecologists, and the technical staff of the Department of Pathology, Brest University Hospital, Omnium group, the Rotary International 'Jetons (le) Cancer', grant 2019, and the local tumor tissue biobank BB-0033-00037 (*CRB Santé/Tumorothèque de Brest*) for their collaboration in this study. A.R. and A.U. contributed to conception and design of the study; A.R., V.C.-C., and A.U. contributed to acquisition and analysis of data; A.R., V.C.-C., L.D., A.T., U.S., P.M., and A.U. contributed to drafting of the manuscript and figures.

References

[1] Niño de Guzmán E, Song Y, Alonso-Coello P, Canelo-Aybar C, Neamtiu L, Parmelli E, et al. Healthcare providers' adherence to breast cancer guidelines in Europe: a systematic literature review. Breast Canc Res Treat 2020;181:499−518.
[2] Siegel RL, Miller KD, Jemal A. Cancer statistics. CA Cancer J Clin 2019;69:7−34.
[3] Gradishar WJ, Anderson BO, Balassanian R, Blair SL, Burstein HJ, Cyr A, et al. NCCN guidelines insights: breast cancer, version 1.2017. J Natl Compr Canc Netw 2017;15:433−51.
[4] Gradishar WJ, Anderson BO, Balassanian R, Blair SL, Burstein HJ, Cyr A, et al. Invasive breast cancer version 1.2016, NCCN clinical practice guidelines in oncology. J Natl Compr Canc Netw 2016;14:324−54.
[5] Gradishar WJ, Anderson BO, Abraham J, Aft R, Agnese D, Allison KH, et al. Breast cancer, version 3.2020, NCCN clinical practice guidelines in oncology. J Natl Compr Canc Netw 2020;18:452−78.
[6] Goetz MP, Gradishar WJ, Anderson BO, Abraham J, Aft R, Allison KH, et al. NCCN guidelines insights: breast cancer, version 3.2018. J Natl Compr Canc Netw 2019;17:118−26.
[7] Elston CW, Ellis IO. Pathological prognostic factors in breast cancer. I. The value of histological grade in breast cancer: experience from a large study with long-term follow-up. Histopathology 1991;19:403−10.
[8] Allison KH, Hammond MEH, Dowsett M, McKernin SE, Carey LS, Fitzgibbons PL, et al. Estrogen and progesterone receptor testing in breast cancer: American society of clinical oncology/college of American pathologists guideline update. Arch Pathol Lab Med 2020;144:545−63.
[9] Wolff AC, Hammond MEH, Allison KH, Harvey BE, Mangu PB, Bartlett JMS, et al. Human epidermal growth factor receptor 2 testing in breast cancer: American society of clinical oncology/college of American pathologists clinical practice guideline focused update. Arch Pathol Lab Med 2018;142:1364−82.
[10] Dowsett M, Nielsen TO, A'Hern R, Bartlett J, Coombes RC, Cuzick J, et al. Assessment of Ki67 in breast cancer: recommendations from the international Ki67 in breast cancer working group. J Natl Cancer Inst 2011;103:1656−64.
[11] Paik S, Shak S, Tang G, Kim C, Baker J, Cronin M, et al. A multigene assay to predict recurrence of tamoxifen-treated, node-negative breast cancer. N Engl J Med 2004;351:2817−26.
[12] Dowsett M, Cuzick J, Wale C, Forbes J, Mallon EA, Salter J, et al. Prediction of risk of distant recurrence using the 21-gene recurrence score in node-negative and node-positive postmenopausal patients with breast cancer treated with anastrozole or tamoxifen: a TransATAC study. J Clin Oncol 2010;28:1829−34.
[13] Mamounas EP, Tang G, Fisher B, Paik S, Shak S, Costantino JP, et al. Association between the 21-gene recurrence score assay and risk of locoregional recurrence in node-negative, estrogen receptor-positive breast cancer: results from NSABP B-14 and NSABP B-20. J Clin Oncol 2010;28:1677−83.
[14] Albain KS, Barlow WE, Shak S, Hortobagyi GN, Livingston RB, Yeh IT, et al. Prognostic and predictive value of the 21-gene recurrence score assay in postmenopausal women with node-positive, oestrogen-receptor-positive breast cancer on chemotherapy: a retrospective analysis of a randomised trial. Lancet Oncol 2010;11:55−65.
[15] Paik S, Tang G, Shak S, Kim C, Baker J, Kiw W, et al. Gene expression and benefit of chemotherapy in women with node-negative, estrogen receptor-positive breast cancer. J Clin Oncol 2006;24:3726−34.
[16] Tang G, Shak S, Paik S, Anderson SJ, Costantino JP, Geyer Jr CE, et al. Comparison of the prognostic and predictive utilities of the 21-gene Recurrence Score assay and Adjuvant! for women with node-negative, ER-positive breast cancer: results from NSABP B-14 and NSABP B-20. Breast Canc Res Treat 2011;127:133−42.
[17] Sparano JA, Gray RJ, Makower DF, Pritchard KI, Albain KS, Hayes DF, et al. Adjuvant chemotherapy guided by a 21-gene expression assay in breast cancer. N Engl J Med 2018;379:111−21.
[18] Bhargava R, Clark BZ, Carter GJ, Brufsky AM, Dabbs DJ. The healthcare value of the Magee Decision Algorithm™: use of Magee Equations™ and mitosis score to safely forgo molecular testing in breast cancer. Mod Pathol 2020;33:1563−70.
[19] Le référentiel des actes innovants hors nomenclature de biologie et d'anatomocytopathologie (RIHN). RIHN 2020. https://solidarites-sante.gouv.fr/IMG/xlsx/ahn_rihn_v2020.xlsx. [Accessed 8 June 2020].
[20] Nitz U, Gluz O, Christgen M, Kates RE, Clemens M, Malter W, et al. Reducing chemotherapy use in clinically high-risk, genomically low-risk pN0 and pN1 early breast cancer patients: five-year data from the prospective, randomised phase 3 West German Study Group (WSG) PlanB trial. Breast Canc Res Treat 2017;165:573−83.
[21] Stemmer SM, Steiner M, Rizel S, Geffen DB, Nisenbaum B, Peretz T, et al. Clinical outcomes in ER+ HER2 -node-positive breast cancer patients who were treated according to the Recurrence Score results: evidence from a large prospectively designed registry. NPJ Breast Cancer 2017;3:32.
[22] Flanagan MB, Dabbs DJ, Brufsky AM, Beriwal S, Bhargava R. Histopathologic variables predict Oncotype DX recurrence score. Mod Pathol 2008;21:1255−61.
[23] Klein ME, Dabbs DJ, Shuai Y, Brufsky AM, Jankowitz R, Puhalla SL, et al. Prediction of the Oncotype DX recurrence score: use of pathology-generated equations derived by linear regression analysis. Mod Pathol 2013 May;26:658−64.
[24] Farrugia DJ, Landmann A, Zhu L, Diego EJ, Johnson RR, Bonaventura M, et al. Magee Equation 3 predicts pathologic response to neoadjuvant systemic chemotherapy in estrogen receptor positive, HER2 negative/equivocal breast tumors. Mod Pathol 2017;30:1078−85.

[25] Bhargava R, Clark BZ, Dabbs DJ. Breast cancers with magee equation score of less than 18, or 18-25 and mitosis score of 1, do not require Oncotype DX testing: a value study. Am J Clin Pathol 2019; 151:316−23.

[26] Chen J, Wang Z, Lv Q, Du Z, Tan Q, Zhang D, et al. Comparison of core needle biopsy and excision specimens for the accurate evaluation of breast cancer molecular markers: a report of 1003 cases. Pathol Oncol Res 2017;23:769−75.

[27] Besusparis J, Plancoulaine B, Rasmusson A, Augulis R, Green AR, Ellis IO, et al. Impact of tissue sampling on accuracy of Ki67 immunohistochemistry evaluation in breast cancer. Diagn Pathol 2016;11:82.

[28] Boros M, Moncea D, Moldovan C, Podoleanu C, Georgescu R, Stolnicu S. Intratumoral heterogeneity for ki-67 index in invasive breast carcinoma: a study on 131 consecutive cases. Appl Immunohistochem Mol Morphol 2017;25:338−40.

[29] Zhou X, Zhang J, Yun H, Shi R, Wang Y, Wang W, et al. Alterations of biomarker profiles after neoadjuvant chemotherapy in breast cancer: tumor heterogeneity should be taken into consideration. Oncotarget 2015;6:36894−902.

[30] Millis SZ, Gatalica Z, Winkler J, Vranic S, Kimbrough J, Reddy S, et al. Predictive biomarker profiling of $>$ 6000 breast cancer patients shows heterogeneity in TNBC, with treatment implications. Clin Breast Canc 2015;15:473−81.

[31] Badve S, Gökmen-Polar Y. Tumor heterogeneity in breast cancer. Adv Anat Pathol 2015;22:294−302.

[32] Elie C, Colombet I. Méthodes d'estimation de la reproductibilité. Sang Thromb Vaiss 2011;23:138−45.

[33] Rabe K, Snir OL, Bossuyt V, Harigopal M, Celli R, Reisenbichler ES. Interobserver variability in breast carcinoma grading results in prognostic stage differences. Hum Pathol 2019;94:51−7.

[34] Aksac A, Ozyer T, Demetrick DJ, Alhajj R. CACTUS: cancer image annotating, calibrating, testing, understanding and sharing in breast cancer histopathology. BMC Res Notes 2020;13:13.

[35] Baltres A, Al Masry Z, Zemouri R, Valmary-Degano S, Arnould L, Zerhouni N, et al. Prediction of Oncotype DX recurrence score using deep multi-layer perceptrons in estrogen receptor-positive, HER2-negative breast cancer. Breast Cancer 2020;27:1007−16.

[36] Turkki R, Byckhov D, Lundin M, Isola J, Nordling S, Kovanen PE, et al. Breast cancer outcome prediction with tumour tissue images and machine learning. Breast Canc Res Treat 2019;177:41−52.

[37] Couture HD, Williams LA, Geradts J, Nyante SJ, Butler EN, Marron JS, et al. Image analysis with deep learning to predict breast cancer grade, ER status, histologic subtype, and intrinsic subtype. NPJ Breast Cancer 2018;4:30.

[38] Tran WT, Jerzak K, Lu FI, Klein J, Tabbarah S, Lagree A, et al. Personalized breast cancer treatments using artificial intelligence in radiomics and pathomics. J Med Imag Radiat Sci 2019;50:S32−41.