

RESEARCH ARTICLE

# MammaPrint versus EndoPredict: Poor correlation in disease recurrence risk classification of hormone receptor positive breast cancer

Andreas Bösl<sup>1\*</sup>, Andreas Spitzmüller<sup>2</sup>, Zerina Jasarevic<sup>1</sup>, Stefanie Rauch<sup>3</sup>, Silke Jäger<sup>4</sup>, Felix Offner<sup>1</sup>

**1** Department of Molecular Diagnostics, Institute of Pathology, Academic Teaching Hospital Feldkirch, Feldkirch, Austria, **2** Department of Genome Oriented Bioinformatics, Technical University of Munich, Freising, Germany, **3** Department of Surgery, Academic Teaching Hospital Feldkirch, Feldkirch, Austria, **4** FH Campus Vienna, University of Applied Sciences, Vienna, Austria

☯ These authors contributed equally to this work.

\* [Andreas.Boesl@lkhf.at](mailto:Andreas.Boesl@lkhf.at)



OPEN ACCESS

**Citation:** Bösl A, Spitzmüller A, Jasarevic Z, Rauch S, Jäger S, Offner F (2017) MammaPrint versus EndoPredict: Poor correlation in disease recurrence risk classification of hormone receptor positive breast cancer. PLoS ONE 12(8): e0183458. <https://doi.org/10.1371/journal.pone.0183458>

**Editor:** William B. Coleman, University of North Carolina at Chapel Hill School of Medicine, UNITED STATES

**Received:** April 28, 2017

**Accepted:** August 6, 2017

**Published:** August 29, 2017

**Copyright:** © 2017 Bösl et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

**Data Availability Statement:** All relevant data are within the paper and its Supporting Information files.

**Funding:** The analyses for the EndoPredict test were carried out blinded to the outcome data and the MammaPrint results in the lab of the manufacturer Sividon Diagnostics in Germany. All reagents for the EndoPredict analysis were provided by Sividon Diagnostics. In addition to that the authors received no specific funding for this

## Abstract

### Introduction

Correct risk assessment of disease recurrence in patients with early breast cancer is critically important to detect patients who may be spared adjuvant chemotherapy. In clinical practice this is increasingly done based on the results of gene expression assays. In the present study we compared the concordance of the 70-gene signature MammaPrint (MP) with the 12 gene assay EndoPredict (EP).

### Methods

Representative tissue of 48 primary tumours was analysed with the MP during routine diagnostic purposes. Corresponding formalin-fixed, paraffin-embedded tissue was thereafter analysed by the EP test. Risk categories of both tests were compared.

### Results

41 of 48 tumours could be directly compared by both tests. Of the 17 MP low risk cases, only 9 were considered low risk by EP (53% agreement) and of the 24 MP high risk cases, 18 were high risk by EP (75% agreement). Discrepancies occurred in 14 of 41 cases (34.1%). There was only a weak and non-significant correlation between the MP and EP test with an overall concordance of only 66%. The original therapeutic recommendation was based on the MP and would have been changed in 38% of the patients following EP test results. 4 patients developed distant metastases. The respective tumours of these patients were all classified as high risk by the EP, but only 3 were classified as high risk by the MP.

work. Sividon Diagnostics had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

**Competing interests:** All reagents for the EndoPredict analysis were provided by Sividon Diagnostics. In addition to that the authors received no specific funding for this work. Sividon Diagnostics had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript. This does not alter our adherence to PLOS ONE policies on sharing data and materials.

**Abbreviations:** ABCSG, Austrian Breast & Colorectal Cancer Study Group; AEC, 3-amino-9-ethylcarbazole; ASCO, American Society of Clinical Oncology; AZGP1, Alpha-2-glycoprotein-1, zinc-binding; BIRC5, Baculovirus IAPRepeat-containing 5; CALM2, Calmodulin 2; CEP17, centromere enumerator probe for chromosome 17; Ct, cycle threshold; DHCR7, 7-Dehydrocholesterol reductase; DIG, digoxigenin; DNA, deoxyribonucleic acid; ER, estrogen receptor; FFPE, formalin-fixed, paraffin-embedded (tissue); H&E, hematoxylin and eosin; HBB, hemoglobin, beta; HER2, human epidermal growth factor receptor 2; IDC, invasive ductal carcinoma; IHC, immunohistochemistry; IL6ST, interleukin 6 signal transducer; ILC, invasive lobular cancer; MGP, Matrix GLA protein; NST, no special type; OAZ1, Ornithine decarboxylase antizyme 1; PR, progesterone receptor; RBBP8, Retinoblastoma binding protein 8; RNA, ribonucleic acid; ROX, 6-Carboxyl-X-Rhodamine; RPL37A, Ribosomal protein L37a; RT-qPCR, quantitative reverse transcription polymerase chain reaction; SISH, silver in situ hybridization; STC2, Stanniocalcin 2; UBE2C, Ubiquitin-conjugating enzyme E2C.

## Conclusion

Both tests resulted in different treatment recommendations for a significant proportion of patients and cannot be used interchangeably. The results underscore the urgent need for further comparative analyses of multi-genomic tests to avoid misclassification of disease recurrence risk in breast cancer patients.

## Introduction

Every year approximately one and a half million new cases of breast cancer are registered worldwide [1]. While breast cancer mortality has substantially dropped over the past decades due to earlier detection and adjuvant therapies it is still difficult to tailor adjuvant chemotherapy for individual patients. The risk of recurrence and indication to administer chemotherapy is largely dependent on the classical clinic-pathological prognostic factors of tumour size, involvement of lymph nodes, histologic grade, expression of hormone receptors, and human epidermal growth factor receptor 2 (HER2) amplification [2,3,4,5]. While these factors aid in the determination of disease recurrence risk it remains a particular problem to identify those patients unlikely to benefit from adjuvant chemotherapy. It has been estimated that in early-stage breast cancer only 15% of tumours will recur suggesting that based on current treatment guidelines many patients will receive chemotherapy without benefit [2,3,4,6]. Given this dilemma there is great interest in the diagnostic application of prognostic and predictive multi-gene expression assays. In clinical practice, these multi-gene classifiers are currently mostly used for estrogen (ER)-positive/HER2-negative breast cancer. The tests primarily rely on quantification of ER- and proliferation-related genes and combine these into multivariate prediction models.

MammaPrint is one of the first-generation prognostic signatures. It is a micro-array based 70 gene assay, which was launched in 2002 [7]. In validation studies the MammaPrint assay was shown to successfully predict metastasis free survival and overall survival and it has been suggested that it might be useful to select patients not requiring adjuvant chemotherapy [7,8,9]. The most recently published results of the large prospective study of the MammaPrint assay (MINDACT) showed that approximately 46% of women with early-stage breast cancer with a high clinico-pathological risk and low genomic risk might not require chemotherapy [10].

EndoPredict is a second generation multigene test launched in 2011 [11]. It is a 12 gene reverse transcription-quantitative real-time PCR (RT-qPCR)-based assay, including disease-relevant as well as control and reference genes. It selectively targets the group of ER-positive/HER2-negative breast cancers [11,12,13,14,15,16]. Based on the relative expression of the genes-of-interest a 12-gene molecular score (EP score) is calculated. Subsequently tumour size and lymph node status are taken into account to calculate the clinico-pathological EPclin score, the final test result. The EndoPredict test was validated independently in three large randomized phase III trials with  $n = 378$  (ABCSG-6),  $n = 1324$  (ABCSG-8), and  $n = 928$  (trans-ATAC) patients resulting in a level of evidence of 1 [11,12,13]. In all three validation studies the EndoPredict was an independent predictor of distant recurrence. In the same three cohorts the EndoPredict test was especially investigated for the prediction of early and late distant metastasis and reliably identifies the subgroups of patients with an excellent long-term prognosis after 5 years of only endocrine therapy [17].

The concordance of the test results of the various gene expression assays on the market is currently studied by several groups [18,19,20,21,22,23]. To the best of our knowledge no group has yet compared the MammaPrint and EndoPredict test.

In the present study, we examined the concordance of the risk classification of the MammaPrint and EndoPredict test in 48 ER+/HER-negative breast cancer samples.

## Material and methods

### Patient's characteristics

48 tissue samples from patients of the Breast Centre Vorarlberg with primary invasive breast cancer which had been analysed with the multi-gene assay MammaPrint during routine patient care between 2007 and 2015 were included in the present study. The median age of the patients was 55 years. The clinico-pathological characteristics of the patients are shown in Table 1. 24 of 48 (50%) of the tumours were invasive ductal carcinomas (IDC), 23/48 (48%) invasive lobular carcinomas (ILC) and 1 tumour (2.1%) was an invasive ductal carcinoma with squamous components. Invasive lobular cancers constitute approximately 15% of all breast cancers. The number of invasive lobular carcinomas in our cohort is obviously higher than that percentage. The decision to conduct the MammaPrint tests (and subsequently the comparative EndoPredict tests) was purely based on the clinical needs. To address this bias the results of the MammaPrint and EndoPredict test were analyzed both for the whole cohort and for ILC and IDC separately. 60% (29/48) of the patients were node-negative and 40%

**Table 1. Characteristics of the breast cancer cohort.**

Variables	N (%)	Variables	N (%)
<b>Age</b>		<b>Chemotherapy</b>	
<40	3 (6.3)	Yes	15 (31.2)
>40	45 (93.7)	No	33 (68.8)
<b>Menopause</b>		<b>Estrogen receptor</b>	
premenopausal	18 (37.5)	Positive	46 (95.8)
postmenopausal	26 (54.2)	Negative	2 (4.2)
Unknown	4 (8.3)		
<b>Tumour size</b>		<b>Progesterone receptor</b>	
≤2cm	31 (64.6)	Positive	39 (81.2)
>2cm—≤ 5cm	14 (29.2)	Negative	9 (18.7)
Unknown	3 (6.2)		
<b>Nodal status</b>		<b>HER2 status</b>	
Negative	29 (60.4)	Negative	46 (95.8)
Positive (1–3)	17 (35.4)	Positive	2 (4.2)
Positive (>3)	2 (4.2)		
<b>Grading</b>			
G1	2 (4.2)		
G2	22 (45.8)		
G3	24 (50)		
<b>Histology</b>			
IDC/NST	24 (50)		
ILC	23 (47.9)		
IDC with squam. components	1 (2.1)		

<https://doi.org/10.1371/journal.pone.0183458.t001>

(19/48) node-positive, respectively. The majority of the tumours were positive for ER (96% [46/48]) and PR (81% [39/48]) and negative for HER2 (96% [46/48]). 92% (44/48) of the tumours were positive for ER and negative for HER2 which represent the target group for gene expression profiling and the EndoPredict test as well. The MammaPrint test can be performed on fresh or freshly frozen breast cancer tissue or FFPE tissue [24]. 22 freshly frozen (46%) and 26 FFPE (54%) tumour samples were analysed. MammaPrint test results had been considered together with other clinical and pathological factors by the interdisciplinary tumour board for clinical decision making regarding administration of adjuvant chemotherapy.

The study was approved by the Research Ethics Committee Vorarlberg (No. 2014-4/1).

### Determination of the ER, PR, HER2 and Ki-67 status using immunohistochemistry and dual silver in situ hybridization

The assessment of hormone receptors (ER, PR), HER2 and the cellular proliferation marker Ki-67 was carried out on the assembled tissue microarray representing the areas used for total RNA isolation for EndoPredict testing. The immunohistochemical staining and the Dual silver in situ hybridization (SISH) were performed on the Ventana BenchMark ULTRA (Ventana® Medical Systems Inc., Tucson, USA) using commercial antibodies and DNA probes following manufacturers' recommendations. For the detection of the primary antibodies for ER, PR and HER2 the BASIC AEC Detection Kit (Ventana®) was used and for Ki-67 the ultraView Universal Alkaline Phosphatase Red Detection Kit (Ventana®). The following primary antibody clones were used: Estrogen receptor (SP1, Ventana®), progesterone receptor (1E2, Ventana®), HER2 (c-erb2, IDLabs, Birmingham, England) and Ki-67 (30-9, Ventana®). All procedures for the SISH analysis were carried out following the recommended protocol of the manufacturer using the INFORM HER2 Dual SISH DNA Probe Cocktail (Ventana®). For the detection of the HER2 probe the ultraView SISH DNP Detection Kit (Ventana®) and for the CEP17 probe the ultraView Red ISH DIG Detection Kit were used, respectively. The cut off for positive ER and PR was set to 1% of positive stained nuclei. HER2 immunohistochemistry was scored as described in the current ASCO guidelines [25]. Ki-67 expression analysis was divided into three categories, in which  $\leq 15\%$  proliferation activity was scored as low proliferation, 16–25% as intermediate and  $> 25\%$  as positive or high proliferation. The used cut-offs were defined in-house. SISH testing was evaluated in reference to the ASCO guidelines. 20 tumour cell nuclei were assessed and the cut-off for positive HER2 status was set as HER2/CEP17 ratio  $\geq 2$ .

### MammaPrint test

Upon the recommendation of the breast cancer tumour board, the samples were sent to the reference laboratory of Agendia (Agendia, Amsterdam, Netherlands) for MammaPrint testing according to their standard operating procedure (SOP) guidelines. For this assay, fresh tissue as well as FFPE tissue can be used [24]. Sections of fresh tissue containing the largest amount of invasive tumour cells were chosen by an experienced pathologist, and sent to Agendia laboratories according to the recommendations of the company. Concerning FFPE material, one representative paraffin block was chosen, containing the largest amount of invasive tumour cells on the corresponding hematoxylin & eosin (H&E) slides. Additionally, 10 unstained serial sections of 5 micrometer thickness per tumour block containing approximately 30% tumour cells were cut from the paraffin blocks, mounted on glass slides and shipped. The test determines the expression of 70 candidate genes and hundreds of reference genes which are measured 6 times each [26]. The multi-gene assay covers the biological motives related to the six hallmarks of cancer including apoptosis, self-sufficiency in growth signals, limitless replicative potential, tissue invasion and metastasis, in-sensitivity to anti-growth signals and sustained

angiogenesis [26,27]. By means of this expression profile a risk classification to develop distant metastasis is developed [8,26,28,29,30].

## EndoPredict test

EndoPredict was performed retrospectively using the same archived paraffin blocks assessed by MammaPrint previously. For the 22 fresh tissue samples analyzed with MammaPrint, corresponding FFPE tumour blocks were chosen. First, a tissue section was cut from each sample for H&E staining, the area of invasive tumour cells was identified and marked under the light microscope for tumour cell enrichment and sampling. As recommended by the manufacturer the area of the invasive tumour was at least 30% of the whole tissue section [29]. One adjacent unstained tissue section (thickness of 10 $\mu$ m) was transferred into a 2ml, RNase-free sample tube for total RNA extraction. Total RNA extraction was performed using a manual, silica-coated magnetic bead-based method (Versant<sup>®</sup> Tissue Preparation Reagent Kit, Siemens Healthcare Diagnostics GmbH, Erlangen, Germany) following the manufacturer's recommendations. RNA was eluted in 100 $\mu$ l elution buffer. Subsequent quantitative reverse-transcription real-Time PCR (RT-qPCR) analysis was performed to determine the yield of total RNA and to check for DNA contamination. For the determination of the amount of RNA and of residual genomic DNA qPCR assays for the RNA of the 60S ribosomal protein L37a coding gene (RPL37A) as well as for DNA of the  $\beta$ -globin (HBB) gene were used according to manufacturer's instructions. If needed, an additional DNase digestion was performed to recover genomic DNA-free total RNA.

The multigene assay EndoPredict (Sividon Diagnostics GmbH, Cologne, Germany / Myriad Genetics, Zurich, Switzerland) uses the quantitative reverse-transcription real-Time PCR (RT-qPCR) technique to detect the expression of eight genes-of-interest (AZGP1, BIRC5, DHCR7, IL6ST, MGP, RBBP8, STC2, UBE2C) and three reference genes (CALM2, OAZ1, RPL37A) applied for normalization as well as of the amount of residual genomic DNA (HBB) using the SuperScript III Platinum One-Step Quantitative RT-PCR Kit with ROX (Invitrogen, Carlsbad, USA) in a VERSANT<sup>®</sup> kPCR System (Siemens Healthcare Diagnostics GmbH) [14,15]. The corresponding sequences of primer and FAM/TAMRA-labelled probes were published previously [11]. After upload of PCR data, tumour size and nodal status the EP and EPclin scores were calculated with the web-based EndoPredict Report Generator. The EP score is purely based on the gene-expression data. The EPclin score, the final result of the EndoPredict test which is used for clinical decision making, considers tumour size and nodal status in addition to the molecular EP score [11]. Using predefined cut off values patients were classified by the software into low or high risk of distant metastasis. The cut off value for the EP score was 5.0 and for the EPclin score 3.32867 [11].

For all 48 study samples the extraction of a sufficient amount of total RNA and generation of valid results was possible. Scores and risk groups were generated for each patient and compared to the risk classification of the MammaPrint test. The analyses for the EndoPredict test were carried out blinded to the outcome data and the MammaPrint results in the lab of the manufacturer Sividon Diagnostics in Germany. All reagents for the EndoPredict analysis were provided by Sividon Diagnostics. In addition to that the author(s) received no specific funding for this work. Sividon Diagnostics had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript. This does not alter our adherence to PLOS ONE policies on sharing data and materials.

## Interdisciplinary tumour board

An interdisciplinary tumour board is conducted weekly at the Breast Center Vorarlberg for the management of breast cancer patients including medical oncologists, surgeons, radiation

therapists, pathologists, radiologists, and other specialists for treatment planning. The results of the different departments are assembled and interpreted. Since 2007 the results of the gene expression profiling test MammaPrint are taken into account additionally. In the framework of a retrospective and virtual tumour board the results of the MammaPrint risk classifications were replaced by the EndoPredict test results. During a decision impact analysis the identity of the patients and treatment recommendations were blinded and reassessed taking EndoPredict risk classifications into account.

## Statistics

All statistical analyses were performed using R version 3.0.2 [31]. Correlation between continuous variables was evaluated using Pearson's correlation coefficient ( $r$ ) or Spearman's rank correlation coefficient ( $\rho$ ) using *cor.test()* function. In case of categorical values contingency tables were generated. Concordance was then calculated as the number of matching cases divided by the total number of observations and evaluated using Cohen's kappa ( $\kappa$ ), calculated as

$$\kappa = \frac{p_o - p_c}{1 - p_c},$$

with  $p_o$  being the observed concordance and  $p_c$  the proportion of cases for which agreement is expected by chance. Additionally, two-sided Fisher's exact test as implemented in the *fisher.test()* function was used to evaluate the observed concordance. For all statistical tests a level of significance of 95% was applied.

## Results

### MammaPrint test results

A valid risk classification by the MammaPrint test was possible in 45 of the 48 cases (94%). In 3 of 48 tumours (6%) the RNA quality was too low for a proper risk assessment. 17 of these 45 tumours (38%) were classified as low risk and 28 (62%) as high risk. When comparing the tumour grade with the risk classification we found no statistically significant correlation with the MammaPrint test results ( $p = 0.413$ , S1 Fig).

### EndoPredict test results

The EndoPredict test provided valid test results for all 48 tumours including an 12-gene molecular EP score and EPclin risk score. For statistical analysis 4 of 48 (8%) of the patients had to be excluded, because their tumours were estrogen receptor negative or HER2 positive and did not meet the inclusion criteria of the EndoPredict test (S1 Fig).

According to the molecular EP score, 18% (8/44) of the patients were classified as low risk and 82% (36/44) of the patients as high risk. The EPclin score re-classified 20% (9/36) EP high risk patients into the low risk group resulting in 39% (17/44) of the patients with a low risk classification and in 61% (27/44) of the patients with a high risk of disease recurrence. A strong concordance of 80% was detected between the EP score and the EPclin score-based risk classifications ( $\kappa = 0.522$ ,  $p = 0.137 \times 10^{-3}$ ). This was to be expected since the EPclin score is directly depending on EP score. When comparing the tumour grade with the risk classification we found no statistically significant correlation with the EP score ( $p = 0.184$ ) and the EPclin score ( $p = 0.062$ ), respectively.

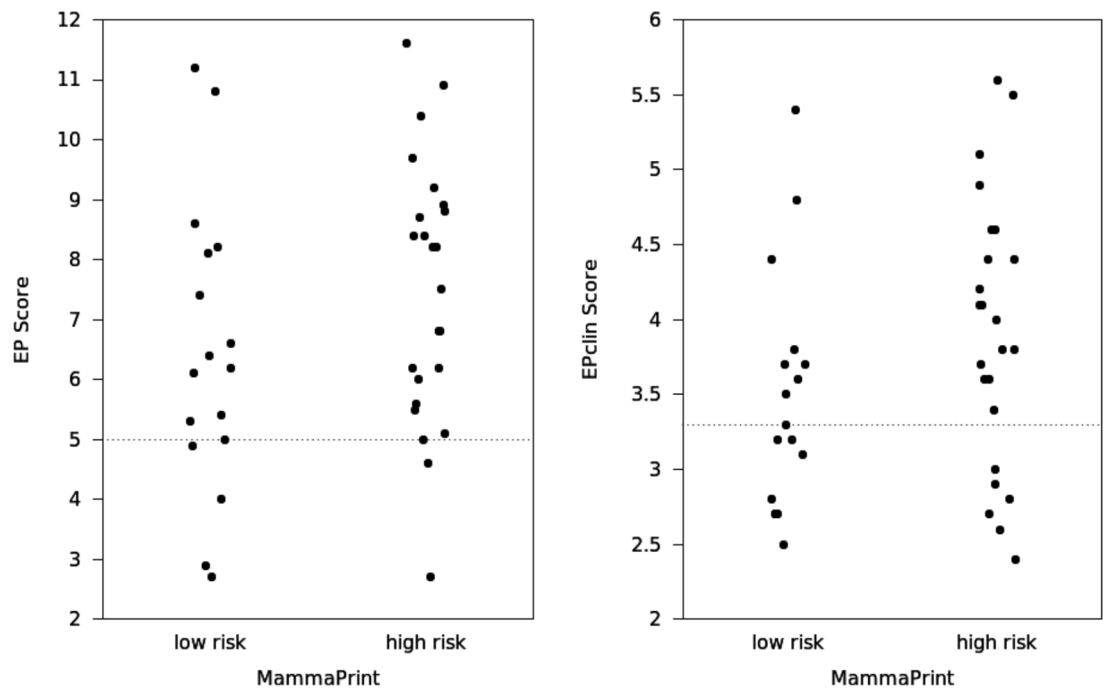


### MammaPrint versus EP Score-based risk classification

Comparing the MammaPrint and EP score-based risk classifications, an overall concordance of 63.4% (26/41) was found. Both tests classified 10% (4/41) and 54% (22/41) of the comparable cases as low risk and high risk, respectively. However, 32% (13/41) of the cases classified as low risk by MammaPrint were classified as high risk by the EP score. 5% (2/41) were classified as high risk by the MammaPrint and low risk by the EP score. As reflected by Cohen’s Kappa coefficient  $\kappa = 0.168$ , there was only a weak agreement between the two molecular classifications (Fig 1; Table 2). Fisher’s exact test did not confirm a statistically significant agreement between the two classification methods beyond of what would have been expected by chance ( $p = 0.212$ ). Separate analysis of ILC and IDC revealed a tendency of a better correlation of the MammaPrint and EP score in ILC. However, this was not statistically significant (Fisher’s exact test) (S1 and S2 Tables).

### MammaPrint versus EPclin Score-based risk classification

Only a poor and non-significant overall concordance of 66% (27/41) was found between MammaPrint and EPclin score-based risk classifications ( $\kappa = 0.284$ ,  $p = 0.102$ ) (Fig 1; Table 2). For 22% (9/41) of the patients both tests classified the tumours as low risk and for 44% (18/41) as high risk. However, of the 17 MammaPrint low risk cases, only 9 were low risk based on the EPclin score (53% agreement). A better agreement was seen in the high risk group. Of the 24 MammaPrint high risk cases 18 were also high risk based on the EPclin score (75% agreement). Separate analysis of the ILC and IDC subsets showed a tendency of a better correlation of the MammaPrint and EPclin test result in ILC than in IDC. Discrepancies occurred in 4 of 20 cases of ILC and in 10 of 21 of IDC. These differences were however, not statistically significant (Fisher’s exact test). (S1 and S2 Tables).



**Fig 1. Correlation between the MammaPrint risk classification and the EP score (left) and EPclin score (right).** Threshold values for EP- and EPclin-based risk classification are indicated as dashed lines.

<https://doi.org/10.1371/journal.pone.0183458.g001>

### Correlation of the Ki-67 index with the MammaPrint, the EP and EPclin score

When comparing the Ki-67 proliferation indices with the molecular risk classifiers we found a statistically significant correlation with the MammaPrint test results ( $p = 0.499 \times 10^{-2}$  using Fisher’s Exact test; Fig 2A) and the molecular EP scores, ( $p = 0.241 \times 10^{-2}$ ,  $r = 0.44$ ; Fig 2B and S1 Table). By contrast, there was no significant relation between the Ki-67 index and the EPclin score ( $p = 0.099$ ,  $r = 0.25$ ; Fig 2C and S1 Table).

### Decision impact analysis

In the framework of interdisciplinary tumour board meetings the effect of the MammaPrint and EndoPredict test results on the choice of therapy were compared.

During the interdisciplinary tumour boards of the Breast Centre Vorarlberg in the years 2007–2015, in an initial step the choice of therapy was considered, taking only classical clinico-pathological parameters into account. These included tumour size, nodal status, hormone receptor and HER2 status and the Ki-67 proliferation index. Sufficient information was available in 47 patients. Based on these data for 45% (21/47) of the patients an adjuvant chemotherapy would have been the choice of therapy, whereas in 55% (26/47) the clinicians would have refrained from chemotherapy recommendation (Fig 3A). Thereafter, the MammaPrint test results were included leading to the actual therapeutic strategy chosen for each patient. MammaPrint results were available for 44 patients and led to a reduction of patients regarded to need adjuvant chemotherapy. Thus, only 36% (16/44) of the patients actually received chemotherapy (Fig 3B). In tumour board meetings of the year 2015, all historic patients were presented again and the MammaPrint test results were replaced by the EndoPredict test results as a hypothetical construct. In this context the recommendation for chemotherapy would have increased to 61% (26/44) of the patients (Fig 3C). 9% (4/44) of the patients were excluded because of a negative hormone receptor (ER negative) or positive HER2 status.

### Follow up data

In the framework of this study follow up data of all patients and in 69% (33/48) risk classifications of both gene expression profiling tests (MammaPrint and EndoPredict) were available.

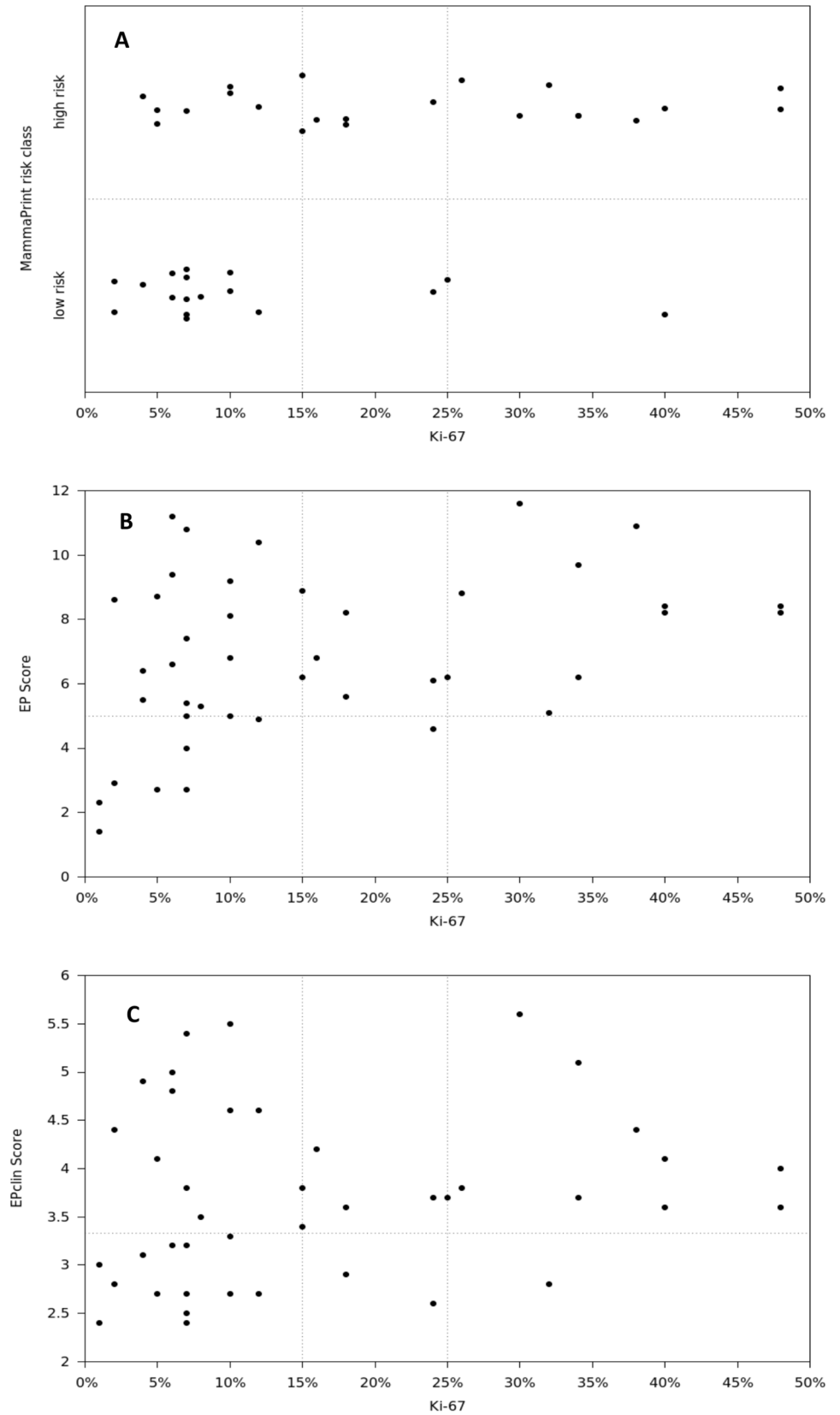
**Table 2. Correlation between EP/ EPclin score and MammaPrint.**

n = 41		MP low risk	MP high risk	Overall concordance	Cohen’s κ	Fisher’s Exact test
Σ	EP low risk	4 (9.8%)	2 (4.9%)	63.4	0.168	0.212
	EP high risk	13 (31.7%)	22 (53.7%)			
G1+2	EP low risk	3 (15%)	1 (5%)	60.0	0.200	0.582
	EP high risk	7 (35%)	9 (45%)			
G3	EP low risk	1 (4.8%)	1 (4.8%)	66.7	0.087	1
	EP high risk	6 (28.5%)	13 (61.9%)			
Σ	EPclin low risk	9 (22.0%)	6 (14.6%)	65.9	0.284	0.102
	EPclin high risk	8 (19.5%)	18 (43.9%)			
G1+2	EPclin low risk	6 (30%)	4 (20%)	60.0	0.200	0.656
	EPclin high risk	4 (20%)	6 (30%)			
G3	EPclin low risk	3 (14.3%)	2 (9.5%)	71.4	0.308	0.280
	EPclin high risk	4 (19.0%)	12 (57.1%)			

Σ: whole cohort; G1+2: only tumours of grade 1+2; G3: only tumours of grade 3.

<https://doi.org/10.1371/journal.pone.0183458.t002>





**Fig 2.** Correlation of the Ki-67 proliferation indices and MammaPrint (A), EP (B) or EPclin (C) score risk classification (threshold values of Ki-67 are registered as vertical dashed lines. Tumour cells with a Ki-67 proliferation index <15% were scored as negative, 16%-25% as intermediate and >25% as positive).

<https://doi.org/10.1371/journal.pone.0183458.g002>

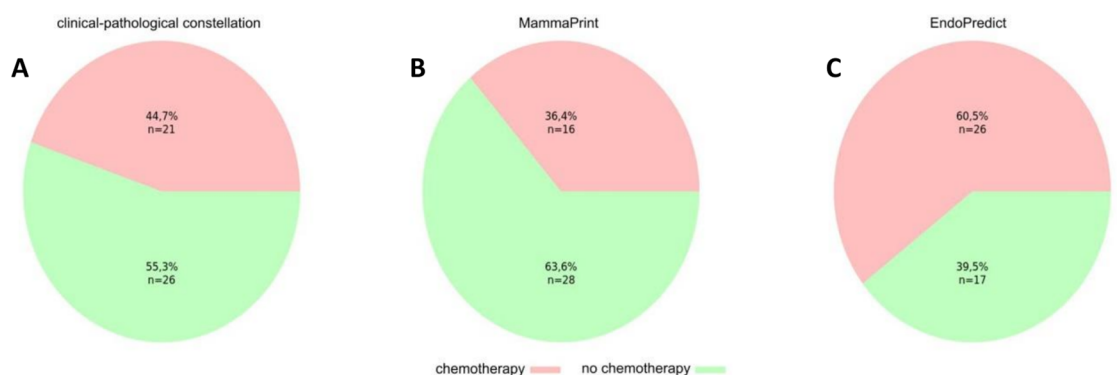
4/33 (12.1%) patients developed distant metastasis during median observation time of 53 months (range 20–96 months). The patients developed distant metastasis after 25, 27, 29 and 31 months, respectively (Table 3). The primary tumours of all of these 4 patients were classified as high risk tumours by the EndoPredict test. The MammaPrint classified 1 of the 4 metastasizing tumours as low risk.

### Discussion

Treatment decisions based on clinico-pathological factors only result in overtreatment of a significant fraction of early breast cancer patients. It has been estimated that approximately 85% of the patients do not benefit from adjuvant chemotherapy regarding 10-year breast cancer-specific survival [2,3,4,6]. This dilemma led to the introduction of various diagnostic multigene expression assays such as MammaPrint, Oncotype DX, Breast Cancer Index, PAM50 and the EndoPredict test. However, the value of these tests to predict prognosis and individualize adjuvant chemotherapy is a matter of an ongoing debate and only a few studies have attempted to compare these assays in a head to head fashion [18,20,21,22,32,33,34,35,36,37,38,39].

To our knowledge, the present study is the first to compare the first generation test MammaPrint and the second generation test EndoPredict. We found a discrepant risk classification in 14 of the 41 (34%) directly comparable tumours which would have resulted in a different treatment in 38% of our patients. Compared to the MammaPrint the EndoPredict test results would have led to significantly higher percentage of patients put on adjuvant chemotherapy (61% instead of 36%). Our results are provocative and support previous observations that multigene expression assay are neither easily interchangeable nor comparable [21,22,34].

MammaPrint was one of the first gene expression assays on the market. It is a 70-gene classifier and was developed to distinguish patients with a high probability of metastasis free survival from patients with high risk to develop distant metastases within 5 years of diagnosis [8,30]. The genes analysed are related to various characteristics of cancer cells including self-sufficiency in growth signals, insensitivity to anti-growth signals, deregulation of apoptosis, limitless replication, tissue invasion and metastasis, and sustained angiogenesis [26]. The test



**Fig 3. Decision impact analysis.** A) Initial choice of therapy based on taking only the clinico-pathological parameters tumour size, nodal status and hormone receptor status as well as HER2 status into account B) Actual choice of therapy based on the inclusion of the MammaPrint test result C) Hypothetical choice of therapy based on the inclusion of the EPclin Score test result instead of the MammaPrint results.

<https://doi.org/10.1371/journal.pone.0183458.g003>

**Table 3. Overview of the risk classification of patients with distant metastasis during observation time.**

Sample ID	Metastasis (Months)	MammaPrint	EP score	EP score value	EPclin score	EPclin score value	Ki-67 (%)	Recurrence site
17	25	high risk	high risk	11.4	high risk	4.5	40	Brain <sup>1</sup>
28	27	low risk	high risk	11.2	high risk	4.8	24	Bone <sup>2</sup>
38	29	high risk	high risk	10.9	high risk	4.4	60	Bone <sup>3</sup>
44	31	high risk	high risk	9.7	high risk	5.1	34	Bone <sup>3</sup>

<sup>1</sup> chemotherapy recommended but refused by patient

<sup>2</sup> chemotherapy omitted based on clinical parameters and MammaPrint test result

<sup>3</sup> chemotherapy recommended and applied

<https://doi.org/10.1371/journal.pone.0183458.t003>

was analysed in various validation studies and proved to be of prognostic value independent of tumour stage and histopathologic factors such as tumour grade and the Nottingham prognostic index [7,40,41,42,43]. A common observation of these studies was that the MammaPrint works better in the prognostication of early recurrences. In a meta-analysis it was also shown that the test may also provide predictive information when applying chemotherapy in high risk patients [44]. The potential benefit of the MammaPrint test was recently highlighted in the phase-3 study of 6993 early-stage breast cancer patients (MINDACT). In this comprehensive study it was shown that 46% of women with a high clinical but low genetic risk do not require chemotherapy. Additionally, it was demonstrated that the 5-year overall survival of patients with estrogen-positive, HER2-positive or negative and either node-negative or node-positive breast cancer was similar [10]. A re-evaluation of the patients after an extended time period of 10 years will be of great interest.

However, recently so called second generation assays have entered the market. They address the main target group of estrogen receptor positive tumours more specifically and are thought to predict not only early but also late recurrences within a time frame of 10 years. Furthermore, they do also incorporate classical clinico-pathologic risk factors [22,34]. The first of the second generation tests launched is the EndoPredict assay. It was introduced as a novel multigene classifier to assess the prognosis in ER+/HER2-negative breast cancer patients. It is a 12 gene expression test including genes related to proliferation and to hormone receptor activity [11]. In contrast to the MammaPrint, EndoPredict test is designed to be performed decentralized which—by many pathologists—is regarded a major advantage [14]. The robustness of the EndoPredict assay in a decentralized test setting using formalin-fixed, paraffin-embedded tissue (FFPE) was successfully proven in seven European pathology institutions with 100% concordance of test results between the different sites [14]. The EndoPredict was validated in three phase III trials involving patients on endocrine therapy only (ABCSG6, ABCSG8, transATAC) and proved to be a significant risk factor classifier for distant disease recurrence [12;45]. A remarkable result of these studies was that approximately 50% of the patients were determined to be “high risk” and would have been subjected to treatment with chemotherapy, if the molecular EP score would have been taken as the basis for therapeutic decision [11,19,23]. A similar, but even more pronounced observation was made in the present study. More than 80% of our patients were classified as high risk by the EP score. This might reflect a selection bias since a relatively high percentage (50%) of our patients presented with grade 3 tumours based on histopathologic grading. However, the final test result of the EndoPredict test combines the molecular EP score with tumour size and nodal status. Based on the resulting EPclin scores, 20% (9/36) of our EP high risk patients were re-classified into the low risk group resulting in 61% (27/44) of our patients with a high risk of disease recurrence. A

similar observation of a “risk-reduction” by the inclusion of traditional clinical parameters was made in the above-mentioned clinical validation trials of EndoPredict. Overall, the percentage of high risk patients in our patient group seemed to correlate when comparing the EPclin score and the MammaPrint. However, of the 24 MammaPrint high risk cases only 18 were high risk by EPclin score resulting in only a weak correlation (75% agreement). Similarly, of the 17 MammaPrint low risk cases, only 9 were low risk by EPclin score (53% agreement). Thus, discrepancies between the EPclin score and the MammaPrint occurred in more than a third of our patients (14 of 41 patients). Based on the purely clinical decision to conduct gene expression tests in our cohort the percentage of lobular cancers was higher than expected. Interestingly, there was a tendency of a better correlation between the MammaPrint and EPclin test result in ILC than in IDC. However, these differences were not statistically significant.

The reason for a higher proportion of EPclin high risk patients in our cohort compared to most of the previous reports might be a large proportion of patients with node positive disease (40%) and/or grade 3 tumours (50%) reflecting a clinically higher risk cohort.

A similar correlation of only 75% was described between the EndoPredict test and the Oncotype Dx Recurrence Scores. [20]. The Oncotype Dx is a RT-qPCR based 21-gene assay, including 16 genes related to proliferation and 5 housekeeping-genes for normalization [46,47].

Recently, the EPclin score was directly compared to purely clinico-pathologic risk classifications and was found to be superior to these classifiers [17]. Furthermore, the EPclin score was used in a recent decision impact study and led to a change of therapy in 37.7% of patients with 25.4% patients placed on endocrine therapy alone and 12.3% on additional chemotherapy [19]. Compared to the clinico-pathologic parameters assessed for our study group the application of the EPclin score would have placed an additional 15.8% on adjuvant chemotherapy, which can be explained by the high risk cohort (40% node positive; 50% grade 3 tumours). By contrast, the MammaPrint led to a reduction of chemotherapies by 8.3% in our study.

Due to the low number of samples, short follow-up data and the low event rate an assessment of the diagnostic potential of both gene expression assays was not possible. Only descriptive data analysis was performed. Our study cannot claim which gene expression assay gave the more accurate risk assessment but raises the important question if all of the patients received the right treatment. 4 patients developed distant metastasis during observation time. EndoPredict identified all of them as high risk while the MammaPrint misclassified one of the metastasizing tumours as low risk. In this case chemotherapy was omitted based on clinical parameters and the MammaPrint test result. In total our study underlines the necessity of further comparative studies of the currently used gene expression assays.

## Supporting information

**S1 Table. Statistical evaluation of the correlation between EP- /EPclin score and MammaPrint in the ILC subset.**  $\Sigma$ : ILC subset; G1+2: only tumours of grade 1+2; G3: only tumours of grade 3.  
(DOCX)

**S2 Table. Statistical evaluation of the correlation between EP- /EPclin score and MammaPrint in the IDC subset.**  $\Sigma$ : IDC; G1+2: only tumours of grade 1+2; G3: only tumours of grade 3.  
(DOCX)

**S3 Table. Concordance of the Ki-67 proliferation indices and risk classification of the gene expression assays.**  
(DOCX)

**S1 Fig. CONSORT diagram of the study cohort.**  
(TIF)

## Acknowledgments

For critical review of the manuscript we are grateful to Drs. Angelika Reiner, Christian Marth, Michael Gnant and William Sterlacci.

## Author Contributions

**Conceptualization:** Andreas Bösl.

**Data curation:** Andreas Bösl, Andreas Spitzmüller.

**Formal analysis:** Andreas Bösl, Andreas Spitzmüller.

**Investigation:** Andreas Bösl.

**Methodology:** Andreas Bösl.

**Project administration:** Andreas Bösl.

**Software:** Andreas Spitzmüller.

**Supervision:** Andreas Bösl, Felix Offner.

**Validation:** Andreas Bösl.

**Visualization:** Andreas Bösl.

**Writing – original draft:** Andreas Bösl.

**Writing – review & editing:** Andreas Bösl, Andreas Spitzmüller, Zerina Jasarevic, Stefanie Rauch, Silke Jäger, Felix Offner.

## References

1. Ferlay J, Soerjomataram I, Dikshit R, Eser S, Mathers C, Rebelo M et al. Cancer Incidence and Mortality Worldwide in GLOBOCAN 2012. *Int J Cancer* 2015 Mar 1; 136:E359–86. <https://doi.org/10.1002/ijc.29210> PMID: 25220842
2. Morris SR, Carey LA. Molecular profiling in breast cancer. *Rev Endocr Metab Disord* 2007;185–98. <https://doi.org/10.1007/s11154-007-9035-3> PMID: 17464566
3. Mehta R, Jain RK, Badve S. Personalized medicine: the road ahead. *Clin Breast Cancer* 2011, 11:20–6. <https://doi.org/10.3816/CBC.2011.n.004> PMID: 21421518
4. Weigelt B, Reis-Filho JS, Swanton C. Genomic analyses to select patients for adjuvant chemotherapy: trials and tribulations. *Ann Oncol* 2012; 23 Suppl 10.
5. Senkus E, Kyriakides S, Ohno S, Penault-Llorca F, Poortmans P, Rutgers E et al. Primary Breast Cancer: ESMO Clinical Practice Guidelines. *Ann Oncol* 2015; 26.
6. Peto R, Davies C, Godwin J, Gray R, Pan HC, Clarke M et al. Comparisons between different polychemotherapy regimens for early breast cancer: meta-analyses of long-term outcome among 100 000 women in 123 randomised trials. *The Lancet* 2012; Volume 379: 432–444.
7. van de Vijver MJ, He Y, van't Veer LJ, Dai H, Hart A, Voskuil D et al. A gene-expression signature as a predictor of survival in breast cancer. *N Engl J Med* 2002; 347:1999–2009. <https://doi.org/10.1056/NEJMoa021967> PMID: 12490681
8. van't Veer LJ, Dai H, van der Vijver MJ, He YD, Hart AA, Mao M et al. Gene expression profiling predicts clinical outcome of breast cancer. *Nature* 2002; 415:530–6. <https://doi.org/10.1038/415530a> PMID: 11823860
9. Drukker CA, Bueno-de-Mesquita JM, Retèl VP, van Harten WH, van Tinteren H et al. A prospective evaluation of a breast cancer prognosis signature in the observational RASTER study. *Int J Cancer* 2013; 133:929–36. <https://doi.org/10.1002/ijc.28082> PMID: 23371464

10. Cardoso F, van't Veer LJ, Bogaerts J, Slaets L, Viale G et al. 70-Gene Signature as an Aid to Treatment Decisions in Early-Stage Breast Cancer. *N Engl. J Med* 2016; 375:717–729. <https://doi.org/10.1056/NEJMoa1602253> PMID: 27557300
11. Filipits M, Rudas M, Jakesz R, Dubsy P, Fitzal F, Singer CF et al. A new molecular predictor of distant recurrence in ER-positive, HER2-negative breast cancer adds independent information to conventional clinical risk factors. *Clin Cancer Res* 2011; 17:6012–20. <https://doi.org/10.1158/1078-0432.CCR-11-0926> PMID: 21807638
12. Buus R, Ivana S, Kronenwett R, Denkert C, Dubsy P, Krappmann U, Scheer M, Petry C, Cuzick J, Dowsett M. *J Natl Cancer Inst* 2016; 108 (11).
13. Cobain EF, Hayes DF. Indications for prognostic gene expression profiling in early breast cancer. *Curr Treat Options Oncol.* (2015); 16:23. <https://doi.org/10.1007/s11864-015-0340-x> PMID: 25929335
14. Denkert C, Kronenwett R, Schlake W, Bohmann K, Penzel R, Weber K et al. Decentral gene expression analysis for ER+/Her2- breast cancer: results of a proficiency testing program for the EndoPredict assay. *Virchows Arch* 2012; 460:251–9. <https://doi.org/10.1007/s00428-012-1204-4> PMID: 22371223
15. Kronenwett R, Bohmann K, Prinzler J, Sinn BV, Haufe F, Roth C et al. Decentral gene expression analysis: analytical validation of the Endopredict genomic multianalyte breast cancer prognosis test. *BMC Cancer* 2012; 12:456. <https://doi.org/10.1186/1471-2407-12-456> PMID: 23039280
16. Martin M, Brase JC, Ruiz A, Prat A, Kronenwett R, Calvo L et al. Prognostic ability of EndoPredict compared to research-based versions of the PAM50 risk of recurrence (ROR) scores in node-positive, estrogen receptor-positive, and HER2-negative breast cancer. A GEICAM/9906 sub-study. *Cancer Res* 2012; 72: P210–11.
17. Dubsy P, Brase JC, Jakesz R, Rudas M, Singer CF, Greil R et al. The EndoPredict score provides prognostic information on late distant metastases in ER+/HER2- breast cancer patients. *Br J Canc* 2013; 109: 2959–2964.
18. Martin M, Brase JC, Ruiz A, Prat A, Kronenwett R, Calvo L et al. Prognostic ability of EndoPredict compared to research-based versions of the PAM50 risk of recurrence (ROR) scores in node-positive, estrogen receptor-positive, and HER2-negative breast cancer. A GEICAM/9906 sub-study. *Breast Canc Res Treat* 2016; 156: 81–89.
19. Müller B, Keil E, Lehmann A, Winzer KJ, Richter-Ehrenstein C, Prinzler J et al. The EndoPredict Gene-Expression Assay in Clinical Practice—Performance and Impact on Clinical Decisions. *PLoS One* 2013; 8:e68252. <https://doi.org/10.1371/journal.pone.0068252> PMID: 23826382
20. Varga Z, Sinn P, Fritsche F, von Hochstetter A, Noske A, Schraml P et al. Comparison of EndoPredict and Oncotype DX Test Results in Hormone Receptor Positive Invasive Breast Cancer. *PLoS One* 2013; 8:e58483. <https://doi.org/10.1371/journal.pone.0058483> PMID: 23505515
21. Alvarado MD, Prasad C, Rothney M, Cherbavaz DB, Sing AP, Baehner F et al. A Prospective Comparison of the 21-Gene Recurrence Score and the PAM50-Based Prosigna in Estrogen Receptor-Positive Early-Stage Breast Cancer. *Adv Ther* 2015; 32: 1237–1247. <https://doi.org/10.1007/s12325-015-0269-2> PMID: 26610383
22. Györfy B, Hatzis C, Sanft T, Hofstatter E, Aktas B, Pusztai L. Multigene prognostic tests in breast cancer: past, present, future. *Breast Canc Res* 2015; 17:11.
23. Coates AS, Winer EP, Goldhirsch A, Gelber RD, Gnant M, Piccart-Gebhart M et al. Tailoring therapies-improving the management of early breast cancer: St. Gallen International Expert Consensus on the Primary Therapy of Early Breast Cancer 2015. *Ann Oncol*: 1–14.
24. Sapino A, Roepman P, Linn SC, Snel MH, Delahaye LJ, van den Akker J et al. MammaPrint molecular diagnostics on formalin-fixed, paraffin-embedded tissue. *J Mol Diagn* 2014; 16:190–7. <https://doi.org/10.1016/j.jmoldx.2013.10.008> PMID: 24378251
25. Wolff AC, Hammond ME, Hicks DG, Dowsett M, McShane LM, Allisson KH et al. Recommendations for Human Epidermal Growth Factor Receptor 2 Testing in Breast Cancer: American Society of Clinical Oncology/College of American Pathologists Clinical Practice Guideline Update In: American Society of Clinical Oncology Guidelines *J Clin Oncol* 2014; Vol 31:3997–4013.
26. Tian S, Roepman P, van't Veer L, Bernards R, de Snoo F, Glas AM. Biological Functions of the Genes in the MammaPrint Breast Cancer Profile Reflect the Hallmarks of Cancer. *Biomark Insights* 2010; 5: 129–138. <https://doi.org/10.4137/BMI.S6184> PMID: 21151591
27. Hanahan D, Weinberg RA. The hallmarks of cancer. *Cell* 2000; 100:57–70. PMID: 10647931
28. Buyse M, Loi S, van't Veer L, Viale G, Delorenzi M, Glas AM et al. Validation and clinical utility of a 70-gene prognostic signature for women with node-negative breast cancer. *J Natl Cancer Inst* 2006; 98:1183–92. <https://doi.org/10.1093/jnci/djj329> PMID: 16954471
29. Glück S, de Snoo F, Peeters J, Stork-Sloots L, Somlo G. Molecular subtyping of early-stage breast cancer identifies a group of patients who do not benefit from neoadjuvant chemo-therapy. *Breast Canc Res Treat* 2013; 139: 759–767.



30. Glas AM, Floore AD, Delahaye LJ, Witteveen AT, Pover RC, Bakx N et al. Converting a breast cancer microarray signature into a high-throughput diagnostic test. *BMC Genomics* 2006; 7:278. <https://doi.org/10.1186/1471-2164-7-278> PMID: 17074082
31. R Core Team. R: A Language and Environment for Statistical Computing. R Foundation for statistical Computing, Vienna, Austria 2013.
32. Sinn P, Aulmann S, Wirtz R, Schott S, Marmé F, Varga Z et al. Multigene Assays for Classification, Prognosis, and Prediction in Breast Cancer: a Critical Review on the Background and Clinical Utility. *Geburtshilfe Frauenheilkunde* 2013; 73:932–940.
33. Zhao X, Rodland EA, Sorlie T, Moen Vollaun HK, Russnes HG, Kristensen H et al. Systematic assessment of prognostic gene signatures for breast cancer shows distinct influence of time and ER status. *BMC Cancer* 2014. 14:211. <https://doi.org/10.1186/1471-2407-14-211> PMID: 24645668
34. Kos Z, Nielsen TO. Developing a new generation of breast cancer clinical gene expression tests. *Breast Canc Res* 2014; 16:103.
35. Dabbs D, Brufsky A, Jankowitz RC, Puhalla S, Lee A, Oesterreich S et al. Comparison of test results and clinical outcomes of patients assessed with both MammaPrint and Oncotype DX with pathologic variables: An independent study. *J Clin Oncol* 2014; 32:5s.
36. Kelly CM, Bernard PS, Krishnamurthy S et al. Agreement in risk prediction between the 21-gene recurrence score assay (Oncotype DX®) and the PAM50 breast cancer intrinsic Classifier in early stage estrogen receptor-positive breast cancer. *Oncologist* 2012; 17:492–498. <https://doi.org/10.1634/theoncologist.2012-0007> PMID: 22418568
37. Maroun R, Saleh R, Asselah J, Omeroglu A, Gulbeyaz A, Meterissian S et al. A head-to-head comparison of MammaPrint and Oncotype Dx: A McGill University Health Center Expertise. *J Clin Oncol* 2015; Vol 33, No 15.
38. Bombard Y, Rozmovits L, Trudeau M, Leighl NB., Deal K, Marshall DA. The value of personalizing medicine: medical oncologists' view on gene expression profiling in breast cancer treatment. *Oncologist* 2015; 20:351–6. <https://doi.org/10.1634/theoncologist.2014-0268> PMID: 25746345
39. Dowsett M, Sestak I, Lopez-Knowles E et al. Comparison of PAM50 Risk of Recurrence Score With Oncotype DX and IHC4 for Predicting Risk of Distant Recurrence After Endocrine Therapy. *J Clin Oncol* 2013; 31:2783–2790. <https://doi.org/10.1200/JCO.2012.46.1558> PMID: 23816962
40. Wittner BS, Sgroi DC, Ryan PD, Brunisma TJ, Glas AM, Male A et al. Analysis of the MammaPrint breast cancer assay in a predominantly postmenopausal cohort. *Clin Cancer Res* 2008; 14:2988–2993. <https://doi.org/10.1158/1078-0432.CCR-07-4723> PMID: 18483364
41. Bueno-de-Mesquita JM, Linn SC, Keijzer R, Wesseling J, Nuyten DS, van Krimpen C et al. Validation of 70-gene prognosis signature in node-negative breast cancer. *Breast Cancer Res Treat* 2009; 117:483–495. <https://doi.org/10.1007/s10549-008-0191-2> PMID: 18819002
42. Mook S, Schmidt MK, Viale G, Pruner G, Eekhout I, Floore A et al. The 70-gene prognosis-signature predicts disease outcome in breast cancer patients with 1–3 positive lymph nodes in an independent validation study. *Breast Canc Res Treat* 2009; 166:295–302.
43. Mook S, Schmidt MK, Weigelt B, Kreike B, Eekhout I, van de Vijver MJ et al. The 70-gene prognosis signature predicts early metastasis in breast cancer patients between 55 and 70 years of age. *Ann Oncol* 2010; 21:717–22. <https://doi.org/10.1093/annonc/mdp388> PMID: 19825882
44. Knauer M, Mook S, Rutgers EJ, Bender RA, Hauptmann M, van de Vijver MJ et al. The predictive value of the 70-gene signature for adjuvant chemotherapy in early breast cancer. *Breast Canc Res Treat* 2010; 120:655–661.
45. Dubsky P, Filipits M, Jakesz R, rudas M, Singer CF, Greil R et al. EndoPredict improves the prognostic classification derived from common clinical guidelines in ER-positive, HER2-negative early breast cancer. *Ann Oncol* 2012; 24:640–647. <https://doi.org/10.1093/annonc/mds334> PMID: 23035151
46. Paik S, Shak S, Baker J, Cronin M. A multigene assay to predict recurrence of tamoxifen-treated, node-negative breast cancer. *N Engl J Med*. 2004; 351:2817–26. <https://doi.org/10.1056/NEJMoa041588> PMID: 15591335
47. Paik S. Is gene array testing to be considered routine now? *Breast* 2011; 20 Suppl 3:S87–91.