A new molecular predictor of distant recurrence in ER-positive, HER2-negative breast cancer adds independent information to conventional clinical risk factors

Martin Filipits¹, Margaretha Rudas², Raimund Jakesz³, Peter Dubsky³, Florian Fitzal³, Christian F. Singer⁴, Otto Dietze⁵, Richard Greil⁶, Andrea Jelen⁷, Paul Sevelda⁸, Christa Freibauer⁹, Volkmar Müller¹⁰, Fritz Jänicke¹⁰, Marcus Schmidt¹¹, Heinz Kölbl¹¹, Achim Rody¹², Manfred Kaufmann¹², Werner Schroth¹³, Hiltrud Brauch¹³, Matthias Schwab¹³, Peter Fritz¹³,¹⁴, Karsten E. Weber¹⁶, Inke S. Feder¹⁵, Guido Hennig¹⁵, Ralf Kronenwett¹⁶, Mathias Gehrmann¹⁵, Michael Gnant³, for the EP Investigators*

¹Departments of Medicine I, ²Pathology, ³Surgery, and ⁴Special Gynecology, Medical University of Vienna, Vienna; ⁵Departments of Pathology, and ⁶Internal Medicine III, Paracelsus Private Medical University, Salzburg; ⁷Department of Pathology and Microbiology, Hanusch Hospital, Vienna; ⁸Department of Gynecology, Hietzing Hospital, Vienna; ⁹Department of Pathology, Weinviertel Hospital, Mistelbach; – all in Austria; ¹⁰Department of Gynecology, University Medical Center Hamburg-Eppendorf, Hamburg; ¹¹Department of Gynecology and Obstetrics, University of Mainz, Mainz; ¹²Department of Obstetrics and Gynecology, J.W. Goethe University, Frankfurt; ¹³Dr. Margarete Fischer-Bosch Institute of Clinical Pharmacology Stuttgart and University Tübingen; ¹⁴Department of Clinical Pathology, Robert Bosch Hospital, Stuttgart; ¹⁵Siemens Healthcare Diagnostics; ¹⁶Sividon Diagnostics, Cologne; – all in Germany.

Running title: A molecular predictor of distant recurrence in breast cancer

Keywords: Breast cancer, endocrine therapy, molecular tests, gene-expression signature, prognosis

Financial support: This work has been supported by the Robert Bosch Foundation, Stuttgart, Germany, and the Federal Ministry of Education and Research (BMBF grants 01ZP0502 and 03IS2061C), Germany
Correspondence to:  Martin Filipits, PhD

Institute of Cancer Research, Department of Medicine I
Medical University of Vienna
Borschkegasse 8a, A-1090 Vienna, Austria
Phone: +43 1 4277 65237; Fax: +43 1 4277 65196

e-mail: martin.filipits@meduniwien.ac.at

*Other EP Investigators of the Austrian Breast & Colorectal Cancer Study Group (ABCSG) and the German Tamoxifen Study Group who participated in the study are listed in the Acknowledgement.
Statement of Clinical Relevance

According to current guidelines molecular tests predicting outcome of breast cancer patients can be used to assist in making treatment decisions after consideration of conventional markers. In the present study, we developed and validated a gene-expression signature predicting the likelihood of distant recurrence in patients with ER-positive, HER2-negative breast cancer treated with adjuvant endocrine therapy only. The risk score (Endopredict, EP) provided additional prognostic information to the risk of distant recurrence of breast cancer patients, independent from clinicopathological parameters, in particular Adjuvant!Online and Ki67 labeling index. By combining the EP risk score with nodal status and tumor size, we identified breast cancer patients with very low risk for distant recurrence who may be adequately and sufficiently treated with adjuvant endocrine therapy only. This new molecular test has a strong potential to assist in optimizing adjuvant therapy and thus might improve management of patients with early-stage breast cancer.
Abstract

**Purpose:** According to current guidelines molecular tests predicting outcome of breast cancer patients can be used to assist in making treatment decisions after consideration of conventional markers. We developed and validated a gene-expression signature predicting the likelihood of distant recurrence in patients with ER-positive, HER2-negative breast cancer treated with adjuvant endocrine therapy.

**Experimental Design:** RNA levels assessed by RT-qPCR in formalin-fixed paraffin-embedded (FFPE) tumor tissue were used to calculate a risk score (Endopredict, EP) consisting of 8 cancer-related and 3 reference genes. EP was combined with nodal status and tumor size into a comprehensive risk score EPclin. Both pre-specified risk scores including cutoff values to determine a risk group for each patient (low, high) were validated independently in patients from two large randomized phase III trials (ABCSG-6: n=378, ABCSG-8: n=1324).

**Results:** In both validation cohorts, continuous EP was an independent predictor of distant recurrence in multivariate analysis (ABCSG-6: P=0.010; ABCSG-8: P<0.001). Combining Adjuvant!Online, quantitative ER, Ki67, and treatment with EP yielded a prognostic power significantly superior to the clinicopathological factors alone (c-indices: 0.764 vs. 0.750, P=0.024 [ABCSG-6]; 0.726 vs. 0.701, P=0.003 [ABCSG-8]). EPclin had c-indices of 0.788 and 0.732 and resulted in 10-year distant recurrence rates of 4% and 4% in EPclin-low-risk and 28% and 22% in EPclin-high-risk patients in ABCSG-6 (P<0.001) and ABCSG-8 (P<0.001), respectively.

**Conclusions:** The multigene EP risk score provided additional prognostic information to the risk of distant recurrence of breast cancer patients, independent from clinicopathological parameters. The EPclin score outperformed all conventional clinicopathological risk factors.
Introduction

Current guidelines for the treatment of early-stage breast cancer recommend adjuvant chemotherapy for most patients with estrogen receptor (ER)-negative or HER2-positive tumors (1-2). In contrast, selecting chemotherapy for patients with ER-positive, HER2-negative disease is more challenging as these patients represent a wide spectrum of different risk profiles: Women who derive little benefit from chemotherapy in addition to endocrine therapy and women with high-risk disease where chemotherapy is very appropriate. Current guidelines recommend the use of validated multigene tests for the decision whether to add chemotherapy or not after consideration of conventional markers (1). Recently, several molecular tests for prediction of breast cancer prognosis have been developed in order to aid clinical decision making (3-5). Some tests have been developed in heterogeneous patient cohorts (6-7). Therefore, a clear answer about their prognostic or predictive value in ER-positive and HER2-negative patients remains elusive, because all published validation results are based on cohorts that include HER2-positive patients. Moreover, none of the tests published to date have been shown to outperform a combination of hormone receptors, HER2 and Ki67 assayed by immunohistochemistry. Here, we present the validation of Endopredict (EP), a new RNA-based multigene score predicting the likelihood of distant recurrence in patients with early stage, ER-positive and HER2-negative breast cancer treated with adjuvant endocrine therapy only. The EP risk score and its combination with the clinical risk factors tumor size and nodal status (EPclin) were generated in a large training set of 964 tumor samples. The pre-specified risk scores and cut-off values were then validated in two independent external validation cohorts of 1702 formalin-fixed paraffin-embedded (FFPE) samples. The patients who donated the samples had been enrolled in two large randomized trials of the Austrian Breast and Colorectal Cancer Study Group (ABCSG) and were treated with adjuvant endocrine therapy only. The aim of the study was to show that the EP score can predict distant recurrence in both cohorts and can significantly add information to standard clinicopathological factors.
Methods

EP and EPclin risk scores

A detailed description of the training process of EP and Epclin risk scores including sample cohorts, selection of the final set of genes, algorithm generation, and threshold definition is shown in the supplementary appendix.

In brief, the final training set for the development of the EP and Epclin risk score consisted of 964 ER-positive, HER2-negative tumors from patients treated with adjuvant tamoxifen only. Because the immunohistochemistry-based hormone receptor status was not available for all samples, selection was based on the ESR1/ERBB2 gene expression using pre-specified cut-off levels (supplementary appendix, section 2.4). In a top-down approach, we developed and defined the EP score consisting of 8 cancer-related genes of interest (GOI: BIRC5, UBE2C, DHCR7, RBBP8, IL6ST, AZGP1, MGP, STC2) and 3 normalization genes (CALM2, OAZ1, RPL37A) (Supplementary Tables S1, S4).

The design and calculation of the final pre-specified EP score as used for validation is described as follows: Mathematically, EP is arranged as a linear combination to predict distant recurrence (Supplementary Figure S4). Relative expression of each GOI was assessed as delta cycle threshold (ΔCt) values based on normalization on the average of three reference genes (CALM2, OAZ1 and RPL37A):

\[ \Delta C t(GOI) = 20 - C t(GOI) + (C t(CALM2) + C t(OAZ1) + C t(RPL37A)) / 3 \] (equation 1)

The ΔCt values were combined into the predictive unscaled risk score \( s_u \):

\[ s_u = 0.41 \cdot \Delta C t(BIRC5) - 0.35 \cdot \Delta C t(RBBP8) + 0.39 \cdot \Delta C t(UBE2C) - 0.31 \cdot \Delta C t(IL6ST) - 0.26 \cdot \Delta C t(AZGP1) + 0.39 \cdot \Delta C t(DHCR7) - 0.18 \cdot \Delta C t(MGP) - 0.15 \cdot \Delta C t(STC2) - 2.63 \] (equation 2)

To avoid negative score values, we defined the final, rescaled EP risk score (s): 0, if \( 1.5 \cdot s_u + 18.95 < 0 \)
\[
\begin{align*}
    s &= 15, & \text{if } 1.5 \cdot s_u + 18.95 > 15 \\
    &= 1.5 \cdot s_u + 18.95, & \text{otherwise}
\end{align*}
\] (equation 3)

The EP risk score ranges from 0 to 15; higher values indicate a higher risk of recurrence.

\(s_{\text{clin}}\), a combined score consisting of the EP risk score and clinical parameters, was constructed from the training set:

\[
    s_{\text{clin}} = 0.35 \cdot t + 0.64 \cdot n + 0.28 \cdot s
\] (equation 4)

where \(t\) codes the tumor size (1: \(\leq\) 1 cm, 2: \(>\) 1 cm to \(\leq\) 2 cm, 3: \(>\) 2 cm to \(\leq\) 5 cm, 4: \(>\) 5 cm), and \(n\) the nodal status (1: negative, 2: 1 to 3 positive nodes, 3: 4 to 10 positive nodes, 4: \(>\) 10 positive nodes).

Thresholds for EP and \(s_{\text{clin}}\) to discriminate patients into low risk and high risk of distant recurrence were developed in the training set at 5 and 3.3, respectively, and pre-specified accordingly for validation.

**Patients and tumor samples of the validation cohorts**

The present investigation is part of the ABCSG translational research program (abcsg.research). Women included in the validation sets had participated in ABCSG-6 (tamoxifen-only arm) or ABCSG-8 trial and received either tamoxifen for 5 years or tamoxifen for 2 years followed by anastrozole for 3 years. Inclusion criteria and the main clinical results of these trials were reported previously (8-9). Breast cancer patients with ER-negative and/or HER2-positive tumors were excluded from the analyses. Baseline clinical data of the validation cohorts are given in Supplementary Table S7. No samples from the validation cohorts were used for training. The clinical characteristics according to EP risk groups of the analyzed 378 patients of ABCSG-6 and the 1324 patients of ABCSG-8 included in this study are listed in Supplementary Table S8. Out of the successfully analyzed samples, 46 (12%, ABCSG-6) and 91 (7%, ABCSG-8) patients had distant recurrences at some point in time after surgery; the mean follow-up time was 97.4 (ABCSG-6) and 72.3 (ABCSG-8) months.
ABCSG-6 was a randomized phase III trial comparing tamoxifen alone for 5 years with tamoxifen in combination with aminoglutethimide for the first 2 years of treatment in postmenopausal women. In ABCSG-8, postmenopausal breast cancer patients were randomly assigned to receive either tamoxifen for 5 years or tamoxifen for 2 years followed by anastrozole for 3 years.

All FFPE tumor blocks were collected at the time of surgery prior to adjuvant therapy and were stored at room temperature. Approval was obtained from institutional review boards. Tumor sections of 4-10 µm were cut. One section was stained by hematoxilin/eosin to confirm the presence of invasive carcinoma, and further sections were used for molecular analyses. Pathologists from participating ABCSG centers sent a representative FFPE tumor block from each woman to the central lab of abcsg.research at the Medical University of Vienna. Overall, twenty-four centers contributed samples (see Acknowledgements).

HER2 and Ki67 were assayed by immunohistochemistry and evaluated centrally in the abcsg.research lab at the Medical University of Vienna according to standard criteria (10-11). For Ki67 samples a cutoff value 11% was used for dichotomization into low or high expression (12). In case of HER2 2+ scores, FISH analyses were used to determine ERBB2 gene amplification.

**RNA extraction and gene expression analysis with quantitative RT-PCR**

Total RNA was extracted from one 5 or 10 µm whole FFPE tissue section with a silica bead-based, fully automated isolation method (Tissue Preparation System; VERSANT Tissue Preparation Reagents; Siemens Healthcare Diagnostics, Tarrytown, US) as described in supplementary appendix (13). All samples were analyzed with quantitative one-step reverse transcription PCR (RT-qPCR) on an ABI PRISM® 7900HT (Applied Biosystems, Darmstadt, Germany) using SuperScript® III Platinum® One-Step Quantitative RT-PCR System with ROX (Invitrogen, Karlsruhe, Germany). All PCR assays were performed in triplicate (details see supplementary appendix). Normalized expression of GOI, EP and EPclin were calculated as described above (equations 1-4).
Clinicopathological risk classification

AdjuvantOnline\(^1\) (software version 8.0) was used to calculate the clinicopathological risk based on patient’s age, tumor size, tumor grade, ER status and number of positive lymph nodes. Patients were considered as having low clinical risk when the 10-year risk as calculated by AdjuvantOnline was below 9\% \((14)\).

Validation and statistical analyses

The two validation studies were performed using pre-specified objectives, assay methods, calculation procedures of scores, and cutoff values. Sample processing, RT-qPCR analyses and calculations of scores were performed completely blinded to clinical and outcome data by the laboratory scientists in Cologne. The scores and risk groups for each patient were then transferred to the ABCSG for statistical analysis in Vienna. The primary endpoint of the statistical analysis was distant recurrence. Statistical analysis of the validation study as well as coding of variables used in Cox regression, c-index calculations and associated P-values is described in more detail in section 3 of the supplementary appendix. Recurrence rates were estimated using the Kaplan–Meier method. All reported P-values are results of two-sided tests unless stated differently. P-values less than 5\% were considered statistically significant.

As primary measure of the prognostic performance the c-index was used (The Design Library: C-index\(^2\) \((15)\).

\(^1\)www.adjuvantonline.com

\(^2\)http://lib.stat.cmu.edu/S/Harrell/Design.html

Its relation to the time-dependent area under the receiver operator characteristic curve is shown in Figure S8 in supplementary appendix. The c-index estimate of a set of variables (e.g. the combination of the AdjuvantOnline score and the EP score) was calculated as described using 10,000 randomizations \((15)\).
P-values to assess whether the EP score contains additional information on distant recurrence with respect to a fixed set of variables X are one-sided and based on the permutation test using 10,000 random permutations with null hypothesis “EP is not related to outcome” and the c-index as test statistic.

Statistical analyses were performed using SPSS software, version 15.0 (SPSS Inc., Chicago, USA) and MATLAB software, version R2009b (The MathWorks, Inc., Natick, USA).

Results

Validation of EP and EPclin in the ABCSG-6 and ABCSG-8 cohorts

The EP score was developed using gene expression data of 964 ER-positive, HER2-negative tumors (supplementary Figure S4; supplementary appendix, section 2). Using the pre-specified threshold for discriminating samples into low or high risk of distant metastasis, the EP risk score was independently validated in two large randomized phase III trials. As in the training set, all women had ER-positive, HER2-negative tumors and received adjuvant tamoxifen or tamoxifen-anastrozole treatment (supplementary Figure S70). RT-qPCR was successful in 378 of 395 (95.7%) and 1324 of 1330 (99.5%) tumors of the two validation cohorts. The EP risk score as continuous predictor estimates the risk of distant recurrence in individual patients at 5 or 10 years (Figure 1A)

The relation between the EP risk score and distant recurrence was assessed by Cox models adjusted for age, tumor size, lymph node status, tumor grade, ER-, PR-, Ki67-expression (IHC), and type of adjuvant endocrine therapy. In these multivariate analyses, the EP risk score was an independent predictor of distant recurrence in ABCSG-6 and ABCSG-8 (Table 1). Subsequently, patients were dichotomized into EP-low and EP-high risk groups according to the cutoff value pre-specified in the training set.

In Kaplan-Meier analyses, we observed significant differences in distant recurrence between EP-low-risk and EP-high-risk patients in both trials (Figure 2A, 2C). At 10 years, the distant recurrence rates for patients with EP-low and EP-high were 8% (3%-13%) and 22% (15%-
29% in ABCSG-6 (P<0.001) and 6% (2%-9%) and 15% (11%-20%) in ABCSG-8 (P<0.001), respectively.

Additionally, the pre-specified Epclin score (combination of EP and the two clinical risk factors nodal status and tumor size) was also validated in the ABCSG-6 and ABCSG-8 co-horts. Epclin is a continuous predictor of distant recurrence at 5 or 10 years (Figure 1B). Distant recurrence rates of patients with EPclin-low and EPclin-high were 4% (1%-8%) and 28% (20%-36%) in ABCSG-6 (P<0.001) and 4% (2%-5%) and 22% (15%-29%) in ABCSG-8 (P<0.001), respectively (Figure 2B, 2D).

Performance of EP in subgroups

In subgroup analyses, we assessed potential heterogeneities of the prognostic performance of the continuous EP risk score by Cox regression analyses adjusted for the clinical risk as calculated by Adjuvant!Online. As shown in Figure 3, a higher EP risk score was associated with a significantly increased risk of distant recurrence in all analyzed subgroups including tumor size, nodal status, grading, Ki67 and ER expression. In addition, there was no heterogeneity among the tamoxifen-only arm of ABCSG-6 and both treatment arms of ABCSG-8.

Exploratory subgroup analyses were performed using Kaplan-Meier analyses with dichotomized EP-low and EP-high risk groups. We found significant differences in distant recurrence between EP-low-risk and EP-high-risk patients in small (T1) and large (T2/T3) tumors, in node-negative and node positive patients, in grade 1 and 2, in ER low (<80%) and high (>80%), in Ki67 low (<11%, luminal A) tumors as well as in Adjuvant!Online low (<9%) and high (≥9%) patients (Supplementary Figure S9).

EP score adds statistically significant information to established clinicopathological risk factors

To assess the contribution of different clinicopathological and molecular variables to the prediction of distant recurrence, we calculated unbiased estimates of the c-index for several combinations of variables (Figure 4). Combination of the established prognostic markers
(nodal status, tumor size, age, grade, quantitative ER (IHC), Ki67 and treatment [ABCSG-8 only]) resulted in c-indices of 0.705 in ABCSG-6 and 0.700 in ABCSG-8. However, addition of the molecular EP score to these clinicopathological markers resulted in a significant improvement of the prognostic power (c-indices: 0.727 in ABCSG-6, 0.728 in ABCSG-8).

Similar results were obtained when using the Adjuvant!Online score as predefined multicomponent clinicopathological risk assessment: C-indices significantly increased from 0.749 to 0.785 in ABCSG-6 and from 0.690 to 0.733 in ABCSG-8 by addition of the molecular EP score. Since quantitative ER (IHC, within ER-positive samples) and Ki67 are not included in the Adjuvant!Online score, we examined the combination of these variables with Adjuvant!Online and assessed the additional prognostic information of EP. We could show that EP significantly added prognostic power also to these clinicopathological variables (Adjuvant!Online, quantitative ER, Ki67, and treatment [ABCSG-8 only]) in both cohorts (0.750 vs. 0.764 in ABCSG-6, 0.701 vs. 0.726 in ABCSG-8). Most interestingly, the EPclin score had higher c-indices than all combinations of clinicopathological variables reported above (0.788 in ABCSG-6, 0.732 in ABCSG-8).

Finally, in an exploratory analysis, we investigated the distant recurrence rate of patients with tumors classified discordantly between Adjuvant!Online and EPclin. We found 467 (27.4%) samples with discordant classification: Adjuvant!Online high risk and EPclin low risk (n=345) or Adjuvant!Online low risk and EPclin high risk (n=122). Within the subgroup of discordant samples outcome follows the EPclin classification rather than the Adjuvant!Online classification (P = 0.0496) (Figure 5).

Discussion

In this study, we developed and validated a new RNA-based molecular test which predicts distant recurrence in ER-positive, HER2-negative early-stage breast cancer patients. The EP score significantly adds prognostic information to established clinicopathological factors such as Adjuvant!Online in combination with quantitative ER expression and Ki67. This is a differentiating feature not shown by other breast cancer prognosis tests so far. Moreover, we
generated and validated a combined molecular/clinicopathological score (EPclin) which outperformed the conventional risk factors alone. EPclin might be helpful in therapeutic decision making when the use of therapy in addition to endocrine therapy is uncertain after consideration of conventional markers. With our results we reach level I evidence according to Simon et al, because our study consisted of two category B studies using archived samples from two similar but separate prospective clinical trials (16).

The EP risk score has several advantages compared with currently available or published prognostic multigene assays. Identification of the candidate genes was based on three predefined requirements. Firstly, we used a large discovery cohort with tumor samples of nearly 1000 patients. Secondly, we analyzed a homogeneous breast cancer subgroup consisting of ER-positive, HER2-negative patients only. Finally, we exclusively included patients treated homogeneously with adjuvant tamoxifen in the training set. Using this approach, we minimized confounding effects from other tumor subtypes or anticancer drugs unrelated to the primary research question. Furthermore, we have chosen a conservative approach for algorithm design aiming at selecting the most robust assembly of prognostic genes, which resulted in a set of eight genes of interest and 3 reference genes, all measured in triplicate. The requirement of both, platform- (HG-U133A arrays and RT-qPCR) and sample-type- (fresh frozen and FFPE), transferability, assessed in a large cohort of matched tumor samples is expected to contribute to assay robustness, an important feature when the test will be used in clinical routine. Moreover, for the performance of the test no special preparation, storage or shipment of tissue is necessary, and assessment of gene expression was possible in 99% of the samples. This is why the test is expected to be compatible to a standard clinical workflow of breast cancer patient care using easily available FFPE tissue.

The EP risk score has not only prognostic value independent of established clinicopathological risk factors in each of the two validation studies ABCSG-6 and ABCSG-8, but provides also significant additional prognostic information to Adjuvant!Online score, quantitative ER immunohistochemistry and Ki67 labeling index. The combination of the molecular EP risk score with the clinicopathological risk factors nodal status and tumor size resulted in the EPclin score identifying risk groups with striking differences in the 10-year
distant recurrence rates. EPclin-low-risk patients had a 10-year risk of distant recurrence of 4% and had, therefore, adequately been treated with adjuvant endocrine therapy only. In contrast, for EPclin-high-risk patients with a 10-year distant recurrence risk of 28% (ABCSG-6) and 22% (ABCSG-8) adjuvant endocrine therapy alone may not be sufficient and additional or other adjuvant treatment may be indicated. A limitation of our study is that the optimal adjuvant therapy in the EPclin-high-risk patient population remains unknown for now and needs to be determined in well-designed prospective clinical trials. A further limitation is that the validation studies have a lower number of grade 3 tumors than what would be expected in a population of postmenopausal breast cancer. However, EP-based risk classification was of borderline significance in this small subgroup of patients.

In summary, using FFPE tumor tissue the EP risk score significantly predicted distant recurrence of breast cancer patients. The EP risk score provided prognostic information independent of - and in addition to - clinicopathological variables, in particular to Adjuvant!Online and Ki67 labeling index. By combining the EP risk score with clinicopathological risk factors, we identified breast cancer patients with low risk for distant recurrence who may be adequately and sufficiently treated with adjuvant endocrine therapy only. Using this new easy-to-perform multigene tool in clinical practice has a strong potential to assist in optimizing adjuvant therapy by reducing both undertreatment and overtreatment and thus might improve management of patients with early-stage breast cancer.

References


Acknowledgements

The following investigators also participated in this study:

Austria:


P. Mayer, M. Moik, C. Menzel, B. Milneritsch, C. Rass, R. Reitsamer, G. Russ (Salzburg Hospital, Salzburg);

K. Haider, W. Kwasny, E. Markis, W. Stiglbauer (Wiener Neustadt Hospital, Wiener Neustadt);

R. Grill, A. Nader, M. Stierer (Haususch Hospital, Vienna);

M. Medl, S. Redtenbacher, W. Ulrich (Hietzing Hospital, Vienna);

G. Reiner (Weinviertel Hospital, Mistelbach);

F. Hofbauer, M. Lang (Oberpullendorf Hospital, Oberpullendorf);

W. Horvath (Güssing Hospital, Güssing);

P. Kier, A. Reiner-Concin, K. Renner (Sozialmedizinisches Zentrum Ost Hospital, Vienna);

P. Balcke, G. Salem, R. Sedivy (St. Pölten Hospital, St. Pölten);

E. Nessler, W. Neunteufel (Dornbirn Hospital, Dornbirn);

A. Haid, A. Lang, Z. Jasarevic, F. Offner (Academic Teaching Hospital Feldkirch, Feldkirch);

C. Gruber, F. Kugler, G. Michlmayer, S. Pöstlberger, R. Schildberger, W. Sega (Sisters of Mercy Hospital, Linz);

F. Fridrik, R. Greul, G. Hochreiner, G. Syré, C. Tausch, G. Wahl (Linz Hospital, Linz);

A. Chott, H. Hartleb, H. Ludwig, H. Salzer (Wilhelminenspital, Vienna);

R. Kocher, G. Leitner, F. Stangl, R. Stering, H. Rabl (Leoben Hospital, Leoben);

E. Forsthuber, J. Omann, A. Urbania, K. Holzmüller, H. Rogatsch (Klagenfurt Hospital, Klagenfurt);

H. Feichtinger, M. Rottmann, J. Schüller (Rudolfstiftung Hospital, Vienna);

F. Burger, O. Braun (Waldviertel Hospital, Horn);
G. Bayer-Tiefenbach, G. Böhm, K. Mach (Oberwart Hospital, Oberwart);
B. Gebhard, W. Rausch (Vöcklabruck Hospital, Vöcklabruck);
R. Punzengruber, W. Öhlinger (Amstetten Hospital, Amstetten);
W. Döller, E. Melbinger (Wolfsberg Hospital, Wolfsberg);
H. Brustmann, S. Naudé, P. Riss (Mödling Hospital, Mödling);

Germany:
M. Eichelbaum (Dr Margarete Fischer-Bosch Institute of Clinical Pharmacology, Stuttgart and University Tübingen)
W. Simon (Robert Bosch Hospital, Stuttgart)
C. Poremba (Zentrum für Histologie, Zytologie und Molekulare Diagnostik Trier, Trier)

We thank T. Ropers, S. Scharff, C. Windbergs, R. Köhler, T. Acht, A. Frehn, C. Roth for excellent technical assistance.
Figure legend

Figure 1. Estimated probability of distant recurrence as continuous functions of the EP risk score (A) and the EPclin risk score (B).

The continuous relation between the respective score and the probability of developing a distant recurrence within the first 5 and 10 years after surgery is described by an independent model for each score generated from all ABCSG-6 and ABCSG-8 data (n=1702). The model is based on fitting the logarithm of the baseline cumulative hazard function by a cubic spline being a function of the logarithm of (event) time as proposed elsewhere (17). For both scores independently, the model describing the data best was found as having two degrees of freedom for the splines, the score as linear covariate, and no interaction between time and score.

The dashed curves indicate the 95% confidence interval. The gray histogram in the background shows the distribution of scores for the patients.

Figure 2. Kaplan-Meier plot of distant recurrence by EP and EPclin risk groups.

Distant recurrence according to EP risk groups (A, C) and EPclin risk groups (B, D) in patients from the two validation cohorts (ABCSG-6 and ABCSG-8). Cut-points for EP were pre-specified at 5 (3.3 for EPclin) in the training set. Numbers in brackets indicate the 95% CI of the hazard ratio.

Figure 3. Forest plot of the adjusted EP risk score hazard ratios.

The forest plot shows the adjusted EP risk score (as continuous variable) unit hazard ratios (squares) and 95% confidence intervals (horizontal lines) according to several clinical variables in the combined ABCSG-6 and ABCSG-8 cohort. Hazard ratios were adjusted for AdjuvantOnline. Sizes of the squares are proportional to the number of patients.

Figure 4. C-Index indicating the performance of combinations of different prognostic factors in the ABCSG 6 and 8 studies.
n = nodal status, t = tumor size, g = grade, ER = estrogen receptor, EP = Endopredict;
The values on the x axis are unbiased estimates of the c-index of the linear combination of
one or more variables by Cox regression. Statistical tests indicate whether the c-index
increases significantly by addition of EP to a fixed set of clinicopathological variables, (see
section 3.3 of the supplementary appendix).

Figure 5. Comparison of the Adjuvant!Online score and EPclin (ABCSG-6 and ABCSG-8).
The subgroups of samples concordant or discordant between Adjuvant!Online and EPclin
are shown. Cut-points were 9% for Adjuvant!Online and 3.3 for EPclin.
Table 1. Multivariate Cox proportional hazard models for estimating the contribution of variables to predict distant recurrence.

<table>
<thead>
<tr>
<th>Cox model</th>
<th>Variable</th>
<th>ABCSG-6</th>
<th>ABCSG-8</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Unit HR (95% CI)*</td>
<td>P-value</td>
</tr>
<tr>
<td>Multivariate Cox model</td>
<td>Age</td>
<td>1.00 (0.96-1.04)</td>
<td>0.864</td>
</tr>
<tr>
<td></td>
<td>Tumor size</td>
<td>1.09 (0.70-1.71)</td>
<td>0.704</td>
</tr>
<tr>
<td></td>
<td>Nodal status</td>
<td>2.47 (1.75-3.48)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td></td>
<td>Grade</td>
<td>0.81 (0.48-1.37)</td>
<td>0.435</td>
</tr>
<tr>
<td></td>
<td>ER (IHC)</td>
<td>0.90 (0.58-1.40)</td>
<td>0.650</td>
</tr>
<tr>
<td></td>
<td>PGR (IHC)</td>
<td>0.83 (0.63-1.10)</td>
<td>0.199</td>
</tr>
<tr>
<td></td>
<td>Ki67</td>
<td>1.03 (1.00-1.06)</td>
<td>0.086</td>
</tr>
<tr>
<td></td>
<td>Treatment arm</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>EP score</td>
<td>1.19 (1.04-1.36)</td>
<td>0.010</td>
</tr>
<tr>
<td>Bivariate Cox model</td>
<td>Adjuvant! score</td>
<td>1.03 (1.02-1.04)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td></td>
<td>EP score</td>
<td>1.19 (1.06-1.32)</td>
<td>0.002</td>
</tr>
</tbody>
</table>

*HR = hazard ratio;

The table shows four different Cox regression models with different sets of variables in both validation cohorts. Each model is shown in two columns: Column "Unit HR" contains the 95% confidence interval for the unit hazard ratio of the respective variable while column “P-value” contains the P-value for the probability of the regression coefficient to be zero.

Variables/units were coded according to table S6 in the supplementary appendix.
Figure 1

A

B
Figure 2

(A) ABCSG-6

(B) ABCSG-6

(C) ABCSG-8

(D) ABCSG-8

Numbers at risk:

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Numbers at risk</td>
<td>191 182 178 118 109 104 77</td>
<td>187 174 157 98 87 81 71</td>
<td>641 628 594 422 179 90 45</td>
<td>683 663 602 417 197 109 58</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>EPclin low</th>
<th>EPclin high</th>
</tr>
</thead>
<tbody>
<tr>
<td>Numbers at risk</td>
<td>208 200 194 129 120 114 91</td>
<td>170 156 141 87 76 71 57</td>
</tr>
<tr>
<td>Variable</td>
<td>Hazard Ratio</td>
<td>Events / Patients</td>
</tr>
<tr>
<td>----------------------------------</td>
<td>--------------</td>
<td>-------------------</td>
</tr>
<tr>
<td>Age</td>
<td></td>
<td></td>
</tr>
<tr>
<td>≤60 years</td>
<td></td>
<td>46/579</td>
</tr>
<tr>
<td>&gt;60 years</td>
<td></td>
<td>91/1123</td>
</tr>
<tr>
<td>Tumor size</td>
<td></td>
<td></td>
</tr>
<tr>
<td>≤2 cm</td>
<td></td>
<td>62/1136</td>
</tr>
<tr>
<td>&gt;2 cm</td>
<td></td>
<td>75/566</td>
</tr>
<tr>
<td>Nodal status</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Negative</td>
<td></td>
<td>59/1165</td>
</tr>
<tr>
<td>1 to 3 positive nodes</td>
<td></td>
<td>48/454</td>
</tr>
<tr>
<td>&gt;4 positive nodes</td>
<td></td>
<td>30/83</td>
</tr>
<tr>
<td>Tumor grade (n=1700)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>G1</td>
<td></td>
<td>18/379</td>
</tr>
<tr>
<td>G2</td>
<td></td>
<td>105/1252</td>
</tr>
<tr>
<td>G3</td>
<td></td>
<td>14/69</td>
</tr>
<tr>
<td>Estrogen receptor</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Low</td>
<td></td>
<td>19/177</td>
</tr>
<tr>
<td>Medium</td>
<td></td>
<td>44/553</td>
</tr>
<tr>
<td>High</td>
<td></td>
<td>74/972</td>
</tr>
<tr>
<td>Progesterone receptor</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Negative</td>
<td></td>
<td>39/353</td>
</tr>
<tr>
<td>Low</td>
<td></td>
<td>30/295</td>
</tr>
<tr>
<td>Medium</td>
<td></td>
<td>31/562</td>
</tr>
<tr>
<td>High</td>
<td></td>
<td>37/492</td>
</tr>
<tr>
<td>Ki67 (n=1638)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>≤11</td>
<td></td>
<td>81/1271</td>
</tr>
<tr>
<td>&gt;11</td>
<td></td>
<td>53/367</td>
</tr>
<tr>
<td>Adjuvant therapy</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Low risk</td>
<td></td>
<td>32/843</td>
</tr>
<tr>
<td>High risk</td>
<td></td>
<td>105/859</td>
</tr>
<tr>
<td>Tamoxifen</td>
<td></td>
<td>98/1029</td>
</tr>
<tr>
<td>Tamoxifen+Anastrozole</td>
<td></td>
<td>39/673</td>
</tr>
<tr>
<td>ABCSG-6 Tamoxifen</td>
<td></td>
<td>46/378</td>
</tr>
<tr>
<td>ABCSG-8 Tamoxifen</td>
<td></td>
<td>52/651</td>
</tr>
<tr>
<td>ABCSG-8 Tamoxifen + Anastrozole</td>
<td></td>
<td>39/673</td>
</tr>
<tr>
<td>Total</td>
<td></td>
<td>137/1702</td>
</tr>
</tbody>
</table>
Figure 5

Absence of distant recurrence over time for different groups:
- AdjuvantOnline low, EPclin low
- AdjuvantOnline high, EPclin low
- AdjuvantOnline low, EPclin high
- AdjuvantOnline high, EPclin high

Numbers at risk:

<table>
<thead>
<tr>
<th></th>
<th>0</th>
<th>20</th>
<th>40</th>
<th>60</th>
<th>80</th>
<th>100</th>
<th>120</th>
</tr>
</thead>
<tbody>
<tr>
<td>721</td>
<td>710</td>
<td>678</td>
<td>474</td>
<td>237</td>
<td>153</td>
<td>95</td>
<td></td>
</tr>
<tr>
<td>345</td>
<td>335</td>
<td>317</td>
<td>208</td>
<td>122</td>
<td>84</td>
<td>55</td>
<td></td>
</tr>
<tr>
<td>122</td>
<td>121</td>
<td>117</td>
<td>85</td>
<td>48</td>
<td>33</td>
<td>19</td>
<td></td>
</tr>
<tr>
<td>514</td>
<td>481</td>
<td>419</td>
<td>288</td>
<td>165</td>
<td>114</td>
<td>82</td>
<td></td>
</tr>
</tbody>
</table>
A new molecular predictor of distant recurrence in ER-positive, HER2-negative breast cancer adds independent information to conventional clinical risk factors

Martin Filipits, Margaretha Rudas, Raimund Jakesz, et al.

*Clin Cancer Res* Published OnlineFirst August 1, 2011.