




BRIEF COMMUNICATION

Molecular Diagnostics

Clinically high-risk breast cancer displays markedly discordant molecular risk predictions between the MammaPrint and EndoPredict tests

Stephan Wenzel Jahn¹ , Andreas Bösl², Oleksiy Tsybrovsky^{3,4}, Christine Gruber-Rossipal³, Ruth Helfgott⁵, Florian Fitzal⁶, Michael Knauer⁷, Marija Balic⁸, Zerina Jasarevic², Felix Offner² and Farid Moifar^{1,3,4}

Inter-test concordance between the MammaPrint and the EndoPredict tests used to predict the risk of recurrence in breast cancer was evaluated in 94 oestrogen receptor-positive, HER2-negative breast cancers. We correlated histopathological data with clinical risk estimation as defined in the MINDACT trial. 42.6% (40/94) of cases were high-risk by MammaPrint, 44.7% (42/94) by EndoPredict (EPclin), and 45.7% (43/94) by clinical risk definition. Thirty-six percent of genomic risk predictions were discordant with a low inter-test correlation between EndoPredict and MammaPrint ($p = 0.012$; $\kappa = 0.27$, 95% CI [0.069, 0.46]). Clinical risk stratification did not correlate with MammaPrint ($p = 0.476$) but highly correlated with EndoPredict ($p < 0.001$). Consequently, clinically high-risk tumours ($n = 43$) were more frequently high-risk by EndoPredict than by MammaPrint (76.6% vs. 46.5%, $p = 0.004$), with 44% of cases discordantly classified and no significant association between genomic risk predictions ($p = 0.294$). Clinicians need to be aware that clinical pre-stratification can profoundly influence multigenomic test performance.

British Journal of Cancer (2020) 122:1744–1746; <https://doi.org/10.1038/s41416-020-0838-2>

INTRODUCTION

Up to 40% of women with early breast cancer receive adjuvant chemotherapy at the price of considerable overtreatment as 70–80% percent of patients are estimated to equally have survived without it.¹ Multigenomic assays can aid in the deliberation on adjuvant chemotherapy. The selection of proprietary tests is often highly individual, depending on reimbursement, geographic region and personal preferences of the oncologist. We directly compared two commonly used tests, namely MammaPrint® (MP) (Agendia, Amsterdam, Netherlands, EU), and EndoPredict® (EP) (Myriad International, Cologne, Germany, EU) as to our knowledge only two smaller studies have so far addressed the concordance between these two tests.^{2,3} After parallel analyses from identical tumours, we correlated risk-stratifications with clinical and histological data. To apply both assays within their designated specifications, we evaluated the tests on a cohort of ER+, HER2–, TNM Stage I and II breast cancers below 5 cm in diameter with up to three positive lymph nodes.

MATERIAL AND METHODS

Fifty-six cases from the Department of Pathology, Hospital of the Sisters of Charity, Linz Austria (ethics committee approval 31/09, Ordensklinikum/Hospital of the Sisters of Charity, Linz Austria) were analysed in the course of this study. MP/EP risk-prediction data from another 38 cases previously published by Bösl et al.,² selected for compliant inclusion criteria were evaluated after the authors were contacted (Schwerpunkt Krankenhaus Feldkirch Austria). Tumours from Linz had received MP testing on routine clinical requests between 2010 and 2016 and were retested with EP for study purposes. For clinical data of the combined cohort ($n = 94$) see Supplemental Table 1.

MammaPrint testing was performed centrally by Agendia, Europe. EndoPredict testing was performed at the Department of Pathology, Ordensklinikum/Hospital of the Sisters of Charity, Austria according to manufacturer's instructions. We computed EPscore and through the addition of pT and pN information, the EPclin. We used categorical "low"/"high" risk classifications for MP, EPscore, and EPclin for statistical analyses. For better readability,

¹Diagnostic and Research Institute of Pathology, Medical University of Graz, Neue Stiftingtalstraße 6, 8010 Graz, Austria; ²Institute of Pathology, Schwerpunkt Krankenhaus Feldkirch, Carinagasse 47, 6800 Feldkirch, Austria; ³Department of Pathology, Ordensklinikum/Hospital of the Sisters of Charity, Seilerstätte 4, 4010 Linz, Austria; ⁴Vincent Academy of Pathology, Seilerstätte 4, 4010 Linz, Austria; ⁵Department of Surgery, Ordensklinikum/Hospital of the Sisters of Charity, Seilerstätte 4, 4010 Linz, Austria; ⁶Department of Surgery, Comprehensive Cancer Center, Medical University of Vienna, Währinger Gürtel 18-20, 1090 Vienna, Austria; ⁷Breast Center Eastern Switzerland, Schuppisstrasse 10, 9016 St. Gallen, Switzerland and ⁸Division of Oncology, Department of Internal Medicine, Clinical Department of Oncology, Medical University of Graz, Auenbruggerplatz 15, 8036 Graz, Austria

Correspondence: Stephan Wenzel Jahn (stephan.jahn@medunigraz.at)

These authors contributed equally: Stephan Wenzel Jahn, Andreas Bösl

Received: 25 October 2019 Revised: 1 March 2020 Accepted: 26 March 2020

Published online: 27 April 2020

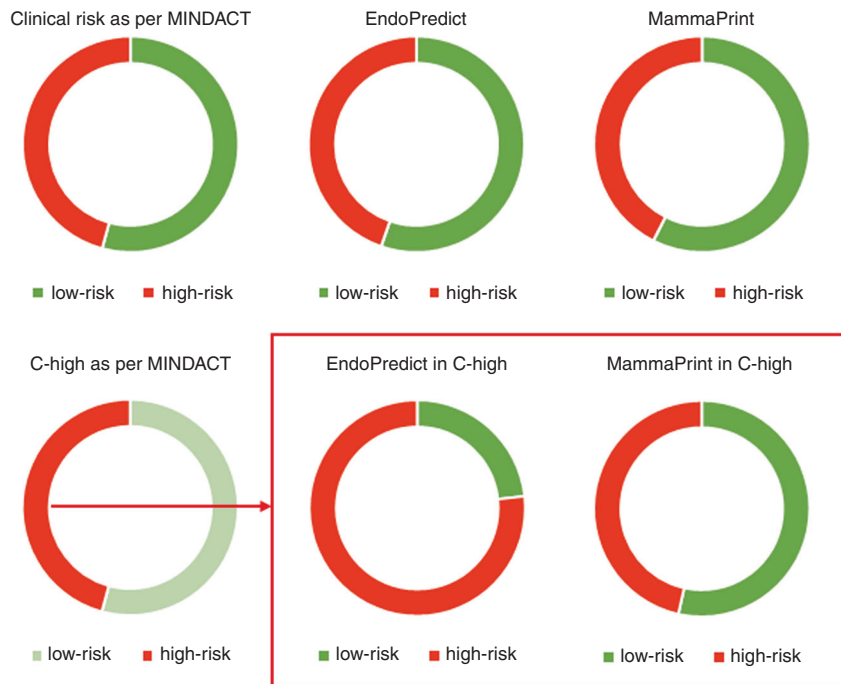


Fig. 1 Results as proportions of low and high-risk predictions. Top row displays whole cohort ($n = 94$) Left: Stratified according to clinical criteria of the MINDACT study (Cardoso et al., NEJM 2016), middle: with EndoPredict, right: with MammaPrint. Bottom row displays clinically high-risk cases. Left: Clinically high-risk cases ($n = 43$) further stratified with EndoPredict (middle) and MammaPrint (right).

mention of “EP” denotes final EPclin results. The methodology for the external cohort is described in Bösl et al.,² and proliferative activity/histological grade was evaluated according to the criteria detailed below. ER/PR positivity was defined as $\geq 1\%$ positive cells by immunohistochemistry. Risk stratifications were correlated with histological grade 1–3, three-tiered Ki-67 stratification rounded to the nearest 5% ($<10\%/10\text{--}30\%/>30\%$), progesterone receptor positivity, pathological T-stage (applicable: pT1b/T1c/T2) and nodal status (pN0/N1). We performed clinical risk assessment as described in the MINDACT trial⁴ (Supplemental appendix Table S13): For ER+/PR+, HER-2-negative tumours pathological T and N stage, as well as grade, were used for classification into “C-low” or “C-high” risk groups.

We used SPSS, V.23 (IBM, USA) and R-scripts⁵ for statistical analyses: Two-sided Chi-square test or Fisher’s exact test for associations between categorical variables, Cohen’s Kappa and Spearman’s Rho to quantify correlations, McNemar tests to compare risk-prediction frequencies. Significance was defined as $p \leq 0.05$ with all Spearman’s Rho $p \leq 0.05$ if not further specified.

RESULTS

Ninety-four cases were amenable to evaluation comprising the 56 cases with de-novo molecular testing and the external data set. 79.8% of cases were high-risk by EPscore, 44.7% by EPclin, and 42.6% by MP. Histopathological Ki-67 index was significantly associated with EPclin and MP, while pT, pN and clinical risk as per MINDACT correlated with EPclin but not MP. For multigenomic/clinical correlation, see Supplemental Table 2, for crosstabulation of test results and risk-predictions per case, see Supplemental Tables 3 and 4. Case per case MP to EPclin risk predictions were discordant in 36%. MP to EPclin risk was significantly associated ($p = 0.01$). Measure of agreement was $\kappa = 0.27$, 95% CI [0.069, 0.46] and 99% CI [0.0075, 0.52]. In clinically high-risk cases ($n = 43$), genomic high-risk predictions were 93% by EPscore, 76.7% by EPclin, and 46.5% by MP. Discordant risk predictions now increased to 44% and MP to EPclin results failed to show a

significant association ($p = 0.294$, $\kappa = 0.15$, 95% CI [−0.089, 0.39]). Clinically high-risk cases were 65% more frequently high-risk by EPclin than by MP ($p = 0.004$). Figure 1 displays the results for clinical and multigenomic risk stratification.

DISCUSSION

To our knowledge, our study is the most comprehensive EP to MP comparison to date and the first to look at inter-test performance in clinically high-risk tumours. Evaluation of 94 ER+/HER2– breast cancers demonstrated an almost equal rate of high-risk predictions for the whole cohort (42.6% vs. 44.7%) but with contradictory predictions for the same tumour in more than a third of cases (36%). This MP to EP discordance rate is in agreement with Bösl et al.² at 34%, Pelaez-Garcia et al.³ at 27.5%, and at 32% approximated in silico from microarray data.⁶ For clinically applied diagnostics a measure of agreement of $\kappa = 0.27$ is disappointing. The upper limit of the 99% CI at $\kappa = 0.52$ implies that inter-test agreement must be expected to remain unsatisfactory even in larger cohorts. The molecular signatures (MP and EPscore) are so different that no statistical association in close to a hundred cases was discernible. Only after the addition of clinical information (EPclin), a significant inter-test association was seen. Therefore clinical information in only one of the tests (EP) is not the underlying cause of discordance as previously hypothesised³ but on the contrary increases concordance. Assuming an actual 20–30% risk of recurrence,¹ both tests substantially overestimate recurrence, albeit partially in different patients.

Also, we investigated both tests in clinically high-risk cases. A recent ASCO practice guideline update⁷ for MP advises to only perform testing in tumours with high clinical risk according to the criteria of the MINDACT trial.⁴ The study combined clinical and genomic risk-prediction. Patients with clinically low-risk tumours received no benefit from chemotherapy irrespective of genomic risk, thus rendering molecular testing unnecessary. Furthermore, only patients at high clinical and high genomic risk were unequivocally advised to receive adjuvant chemotherapy. In

clinically high-risk tumours differences between MP and EP increased so that a high-risk report was now 65% more likely by EP than by MP (76.7% vs. 46.5%), and this difference was statistically significant ($p = 0.004$). Almost every other case (44%) was now discordantly classified. The clinical risk did not correlate with MP ($p = 0.481$) but weakly correlated with EPscore ($Rho = 0.303$). The correlation increased after clinical information (pT, pN) was used to derive EPclin ($Rho = 0.592$). EPclin, as well as clinical risk stratification as per MINDACT both draw on pT/pN information^{4,8} as the two most important clinical prognostic predictors. Consequently, clinical information is redundantly evaluated by EPclin in clinically high-risk tumours. The point is illustrated by a reduction of high-risk predictions from 80% (EPscore) to 45% (EPclin) for the whole cohort, compared to only a minor reduction from 93% (EPscore) to 76% (EPclin) in clinically high-risk cases. Our data compare well to the 77% genomic high-risk predictions by EPclin in the node-positive subgroup of the recent study by Sestak et al.⁹ and the 39% of clinically high-risk (HR+/HER2-) cases by MP in the MINDACT trial.⁴ Clinicians need to be aware that clinical pre-stratification can significantly impact multigenomic test performance. As the first prospective trial on EP risk-prediction is ongoing at four years follow-up,¹⁰ our results confirm the need for further comparative prospective clinical trials, with a particular focus on test performance in clinically high-risk tumours.

ACKNOWLEDGEMENTS

None

AUTHOR CONTRIBUTIONS

S.J.: conceptualisation, writing the original draft, data curation, formal analysis. A.B.: conceptualisation, project administration, formal analysis, data curation, writing—review and editing. O.T.: data curation, visualisation, software, methodology, formal analysis. C.R., R.H., F.F., M.K., M.B., Z.J.: interpretation of results as well as writing, review and editing. F.O.: interpretation of results, writing review and editing, project administration. F.M.: conceptualisation, methodology, funding acquisition, project administration, supervision

ADDITIONAL INFORMATION

Ethics approval and consent to participate The study received ethics approval by the ethics committee (approval number 31/09) of the Ordensklinikum/Hospital of the Sisters of Charity, Linz Austria. Requirement of the patients to provide informed consent in order to participate in the study was waived by the ethics board approval. The study was performed in accordance with the Declaration of Helsinki.

Data availability Data used for analysis are listed in Supplemental Table 4.

Competing interests One author (K.M.) declares competing interests through involvement in the ongoing Phase III study (SAKK 23/16): "Tailored Axillary Surgery With or Without Axillary Lymph Node Dissection Followed by Radiotherapy in

Patients With Clinically Node-positive Breast Cancer (TAXIS). A Multicenter Randomized Phase III Trial" The study receives funding from Agendia, Europe. The author does not receive any personal financial contributions or advantages due to his participation in the study.

Funding information Part of the funding for this study was provided by Sividon Diagnostics GmbH, Cologne, Germany and Myriad Service GmbH, Munich, Germany. The companies had no influence on the study design, on the interpretation or on the publication of the data of this study.

Supplementary information is available for this paper at <https://doi.org/10.1038/s41416-020-0838-2>.

Note This work is published under the standard license to publish agreement. After 12 months the work will become freely available and the license terms will switch to a Creative Commons Attribution 4.0 International (CC BY 4.0).

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

REFERENCES

1. van 't Veer, L. J., Dai, H., van de Vijver, M. J., He, Y. D., Hart, A. A., Mao, M. et al. Gene expression profiling predicts clinical outcome of breast cancer. *Nature* **415**, 530–536 (2002).
2. Bosl, A., Spitzmuller, A., Jasarevic, Z., Rauch, S., Jager, S. & Offner, F. MammaPrint versus EndoPredict: poor correlation in disease recurrence risk classification of hormone receptor positive breast cancer. *PLoS ONE* **12**, e0183458 (2017).
3. Pelaez-Garcia, A., Yebenes, L., Berjon, A., Angulo, A., Zamora, P., Sanchez-Mendez, J. I. et al. Comparison of risk classification between EndoPredict and MammaPrint in ER-positive/HER2-negative primary invasive breast cancer. *PLoS ONE* **12**, e0183452 (2017).
4. Cardoso, F., van't Veer, L. J., Bogaerts, J., Slaets, L., Viale, G., Delaloge, S. et al. 70-gene signature as an aid to treatment decisions in early-stage breast cancer. *N. Engl. J. Med.* **375**, 717–729 (2016).
5. R Core Team. A language and environment for statistical computing. Vienna, Austria (2013). <https://www.R-project.org>.
6. Zhao, X., Rodland, E. A., Sorlie, T., Volland, H. K., Russnes, H. G., Kristensen, V. N. et al. Systematic assessment of prognostic gene signatures for breast cancer shows distinct influence of time and ER status. *BMC Cancer* **14**, 211 (2014).
7. Krop, I., Ismaila, N. & Stearns, V. Use of biomarkers to guide decisions on adjuvant systemic therapy for women with early-stage invasive breast cancer: American Society of Clinical Oncology Clinical Practice Focused Update Guideline Summary. *J. Oncol. Pract.* **13**, 763–766 (2017).
8. Filipits, M., Rudas, M., Jakesz, R., Dubsy, P., Fitzal, F., Singer, C. F. 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. Off. J. Am. Assoc. Cancer Res.* **17**, 6012–6020 (2011).
9. Sestak, I., Buus, R., Cuzick, J., Dubsy, P., Kronenwett, R., Denkert, C. et al. Comparison of the performance of 6 prognostic signatures for estrogen receptor-positive breast cancer: a secondary analysis of a randomized clinical trial. *JAMA Oncol.* **4**, 545–553 (2018).
10. Ettl, J., Anders, S., Hapfelmeier, A., Noske, A., Paepke, S., Weichert, W., et al. First prospective outcome data for the clinico-molecular test Endopredict® in hormone receptor positive, HER2-negative early breast cancer in clinical routine. *San Antonio Breast Cancer Conference*. San Antonio, Texas, USA. (2018)