Annals of Oncology 25: 339–345, 2014
doi:10.1093/annonc/mdt494
Published online 16 December 2013

Predicting distant recurrence in receptor-positive breast cancer patients with limited clinicopathological risk: using the PAM50 Risk of Recurrence score in 1478 postmenopausal patients of the ABCSG-8 trial treated with adjuvant endocrine therapy alone

M. Gnant[1]*, M. Filipits[2], R. Greil[3], H. Stoeger[4], M. Rudas[5], Z. Bago-Horvath[5], B. Mlineritsch[3], W. Kwasny[6], M. Knauer[7], C. Singer[8], R. Jakesz[1], P. Dubsky[1], F. Fitzal[1], R. Bartsch[9], G. Steger[9], M. Balic[4], S. Ressler[3], J. W. Cowens[10], J. Storhoff[10], S. Ferree[10], C. Schaper[11], S. Liu[12], C. Fesl[13], T. O. Nielsen[12], on behalf of the Austrian Breast and Colorectal Cancer Study Group

[1]Department of Surgery and Comprehensive Cancer Center, Medical University of Vienna, Vienna, Austria; [2]ABCSG Research, Institute of Cancer Research and Comprehensive Cancer Center, Medical University of Vienna, Vienna, Austria; [3]IIIrd Medical Department with Hematology, Medical Oncology, Hemostaseology, Rheumatology and Infectious Diseases, Oncologic Center, Paracelsus Medical University Salzburg, Salzburg; [4]Division of Oncology, Department of Internal Medicine, Medical University of Graz, Graz; [5]Clinical Institute of Pathology and Comprehensive Cancer Center, Medical University of Vienna, Vienna; [6]Department of Surgery, State Hospital Wiener. Neustadt, Wiener Neustadt; [7]Department of Surgery, Breast Cancer Center, Hospital of the Sisters of Charity Linz; [8]Department of Gynecology and Obstetrics and Comprehensive Cancer Center, Medical University of Vienna, Vienna; [9]Department of Internal Medicine I and Comprehensive Cancer Center, Medical University of Vienna, Vienna, Austria; [10]NanoString Technologies, Seattle; [11]MyRAQA, Redwood Shores, USA; [12]British Columbia Cancer Agency, Vancouver, Canada; [13]Austrian Breast and Colorectal Cancer Study Group, Vienna, Austria

Received 4 April 2013; revised 17 September 2013 & 24 September 2013; accepted 2 October 2013

Background: PAM50 is a 50-gene test that is designed to identify intrinsic breast cancer subtypes and generate a Risk of Recurrence (ROR) score. It has been developed to be carried out in qualified routine hospital pathology laboratories.

Patients and Methods: One thousand four hundred seventy-eight postmenopausal women with estrogen receptor (ER)+ early breast cancer (EBC) treated with tamoxifen or tamoxifen followed by anastrozole from the prospective randomized ABCSG-8 trial were entered into this study. Patients did not receive adjuvant chemotherapy. RNA was extracted from paraffin blocks and analyzed using the PAM50 test. Both intrinsic subtype (luminal A/B, HER2-enriched, basal-like) and ROR score were calculated. The primary analysis was designed to test whether the continuous ROR score adds prognostic value in predicting distant recurrence (DR) over and above standard clinical variables.

Results: In all tested subgroups, ROR score significantly adds prognostic information to the clinical predictor ($P < 0.0001$). PAM50 assigns an intrinsic subtype to all cases, and the luminal A cohort had a significantly lower ROR at 10 years compared with Luminal B ($P < 0.0001$). Significant and clinically relevant discrimination between low- and high-risk groups occurred also within all tested subgroups.

Conclusion(s): The results of the primary analysis, in combination with recently published results from the ATAC trial, constitute Level 1 evidence for clinical validity of the PAM50 test for predicting the risk of DR in postmenopausal women with ER+ EBC. A 10-year metastasis risk of <3.5% in the ROR low category makes it unlikely that additional chemotherapy would improve this outcome—this finding could help to avoid unwarranted overtreatment.

Clinical trial number: ABCSG 8: NCT00291759.

Key words: early breast cancer, prognosis, metastasis prediction, intrinsic subtypes, clinical prognostic factors, Risk of Recurrence (ROR)

introduction

Hormone receptor-positive breast cancer is a heterogeneous disease from a molecular and clinical perspective. In early-stage breast cancer (EBC), the relapse risk of individual patients

*Correspondence to: Dr Michael Gnant, Department of Surgery and Comprehensive Cancer Center, Medical University of Vienna, Waehringer Guertel 18-20, A-1090 Vienna, Austria. Tel: +43-1-40400-5646; Fax: +43-1-40400-6807; E-mail: michael.gnant@meduniwien.ac.at

treated with adjuvant endocrine therapy varies greatly [1, R2], although outcomes among estrogen receptor (ER)-positive, HER2-negative patients and among HER2-positive patients have significantly improved in recent decades due to the clinical benefit of adjuvant endocrine [3, R4] and trastuzumab [R5, R6] treatments in biomarker-defined subgroups.

In contrast, validated predictive markers concerning adjuvant cytotoxic treatment are still lacking in ER-positive, HER2-negative disease [R7], a group among whom the majority of patients may not individually benefit at all [8].

Because absolute benefits between treatment and control groups in trials are largest in subgroups with higher underlying risk, many clinicians base their adjuvant treatment recommendations on the overall risk of recurrence (ROR) [9], making risk assignments a high priority for clinical breast cancer research. In treatment situations where the overall relapse risk is low, absolute chemotherapy benefits will be small to the point where most physicians consider it safe to avoid adjuvant chemotherapy with all its side-effects. Avoiding unnecessary chemotherapy can also save considerable amounts of health care spending [10].

Tumor size, grade, and nodal status are currently used for risk assessment and decisions about whether adjuvant chemotherapy will be added to endocrine treatment [11]. Recently, multigene expression assays have been developed to achieve a more accurate assessment of prognosis and prediction of therapeutic benefit.

As a second-generation multigene expression assay, the PAM50 test (50 discriminator genes + 8 controls) was developed to identify intrinsic breast cancer subtypes [luminal A (LumA)/B (LumB), HER2-enriched, basal-like] [R2, 12, 13, R14, 15], which reflect the underlying biology associated with ER and HER2 pathways, and in addition includes proliferation genes and markers of the basal phenotype. The terminology of intrinsic subtypes was adopted by the 2011 St Gallen Consensus Conference to describe the paradigm for making treatment decisions in patients with EBC [9]. Luminal subtypes A and B are the most common subtypes of breast cancer in the clinically hormone receptor-positive population [10, 16, R17, R18]. LumA tumors, characterized by lower expression of genes associated with cell cycle activation and ERBB2 [R7, 15, R19] have significantly lower rates of recurrence (i.e. better prognosis) when compared with LumB, which can be quantified as a ROR score, as shown in prior PAM50 studies [12, 13, R14, 15, 16, R17–R19].

The objective of this study was to demonstrate that the PAM50-derived ROR score is a reliable prognostic indicator for distant recurrence-free survival (DRFS) in hormone receptor-positive, postmenopausal ABCSG-8 patient population treated with adjuvant systemic endocrine therapy alone [20] in the form of 5 years of tamoxifen or 2 years of tamoxifen and 3 years of anastrozole (aromatase inhibitor) as currently recommended by today's standard of care [3, 20]. We hypothesize that the ROR score provides significant additional information to classical clinicopathological parameters in hormone receptor-positive, postmenopausal breast cancer patients, and that it is possible to define and discriminate between ROR-derived risk groups based on accurate estimation of the probability of DRFS at 10 years follow-up.

methods

ABCSG-8

The study cohort consists of FFPE breast tumor tissue samples retrospectively collected and archived in the ABCSG tumor bank from patients enrolled between 1996 and 2004 in the ABCSG-8 trial. Three thousand ninety-one women with estrogen and/or progesterone receptor-positive early breast cancer (EBC) were randomized before treatment to 2 years of adjuvant tamoxifen followed by 3 years of anastrozole (Arimidex®) or 5 years of adjuvant tamoxifen [R21].

collection, consent, and procedures

In accordance with a first estimation of available tissue specimens, a database for 2255 patients of the 3901 patients randomized to ABCSG-8 trial was generated. Data for patients samples that met the eligibility criteria for the original trial were only excluded either because tissue was unavailable for the multigene expression assay to be carried out or the patient could not be re-consented. One thousand six hundred twenty patient samples could be used for this study.

re-consent process

The Austrian Data Protection Law requires consent of patients for any analysis carried out on their stored tumor samples. This general rule does not apply for already deceased patients. In line with this legal background and the respective Ethics Committee approval, patients for whom tumor samples were already collected and centralized in ABCSG's research tumor bank during the course of the ABCSG-8 study or afterward were contacted by the responsible site coordinators and informed about this study in accordance to the ICH-GCP guidelines. An informed consent form approved by the respective Ethics Committee was signed by all participating patients. Only tumor samples of patients for whom a signed and dated informed consent form was available or already deceased were used for this study. The distribution of clinicopathological parameters in the original ABCSG-8 trial cohort is shown in supplementary Table S1, available at *Annals of Oncology* online.

PAM50 assay description and ROR score calculation

PAM50 gene expression measurements were carried out on FFPE sections using the Nanostring nCounter device [R22, R23]. Methods followed prespecified and audited standard operating procedures within a CLIA-certified laboratory. Conversion of gene expression measurements into subtype and ROR scores used a fully prespecified algorithm, with risk categories (supplementary Table S2, available at *Annals of Oncology* online) based on TransATAC data [15]. Researchers generating gene expression data were blinded to clinical data. Full details are provided in supplementary Methods, available at *Annals of Oncology* online.

study end points

The primary end point was DRFS, defined as the interval from randomization until distant recurrence or death due to breast cancer. Contralateral breast cancer, secondary malignancy and death due to causes other than breast cancer were treated as censoring events. Death due to breast cancer where a recurrence was not recorded was considered an event at the date of death.

statistical analysis

All analyses were fully prespecified and defined in a written plan (details are provided in supplementary Methods, available at *Annals of Oncology* online; SAP provided as supplementary Material, available at *Annals of Oncology*

online). Those generating the ROR scores and subtypes had no access to the clinical data of the trial.

Cox proportional hazards models were used to assess the effects of individual prognostic factors, a combined linear predictor (clinical linear predictor, CLP), ROR score, ROR score-derived risk groups, and intrinsic subtypes; hazard ratios (HR) with 95% confidence intervals (95% CIs) were estimated.

A two-sided α of 0.05 was used for all tests. All analyses were carried out by two independent statisticians in parallel using SAS version 9.3 and R version 2.15.2.

results

In accordance with a first estimation of potentially available tumor tissue samples, a database for 2255 patients of 3901 patients enrolled between 1996 and 2004 in the ABCSG-8 trial was generated (supplementary Figure S1, available at *Annals of Oncology* online). Four hundred thirty-five of these patients had died during the treatment phase of ABCSG-8 trial or thereafter; for 414 of these patients, evaluable specimens were available. From 1241 patients who were still alive and signed the informed consent, 1206 tissue specimens were available. In total, 1620 (414 +1206) specimens were available for the analysis. 1478 specimens (91.2%) passed the prespecified PAM50 quality control standards and were included in the analyses. Among those which failed, 25 (1.5%) tissue blocks contained insufficient tumor on histology review, 73 (4.5%) yielded insufficient RNA for the assay, and 44 (2.7%) yielded extracted RNA that failed the quality specifications for the nCounter device.

The analysis population included 70% patients with T1-stage cancer and 71% patients with node-negative disease (supplementary Table S3, available at *Annals of Oncology* online). Median age was 63 years (range: 41–79 years). The median follow-up was 11 years.

One hundred seventy-two DRFS events (distant recurrence or death from breast cancer) were recorded in 1478 patients. Seventeen of these were censored at the time of occurrence of a secondary malignancy before a DRFS event, resulting in 155 first DRFS events: the power to meet the primary study objective was >99%. Three hundred ninety-one (26%) patients died during the period of record, 99 (25% of all deaths) due to breast cancer. A summary and listing of the characteristics of patients who developed a DRFS event is given in supplementary Table S4, available at *Annals of Oncology* online.

The prognostic value of standard clinicopathological factors was evaluated by a Cox proportional hazards regression model that showed a significant prognostic effect of tumor grade (G2/GX versus G1; HR = 2.15, 95% CI 1.21–3.80; $P = 0.009$), tumor size (T2/T3 versus T1; HR = 2.33, 95% CI 1.68–3.22; $P < 0.0001$) and nodal status (N1 versus N0: HR = 1.78, 95% CI 1.27–2.49; $P < 0.001$; N2 versus N0: HR = 2.36, 95% CI 1.22–4.59; $P = 0.011$). No significant effect of age (≥65 versus < 65 years: HR = 1.27, 95% CI 0.92–1.75; $P = .144$) or trial treatment arm (tamoxifen/anastrozole versus tamoxifen only: HR = 0.98, 95% CI 0.71–1.34; $P = 0.881$) was observed. Schoenfeld residuals did not deviate from the proportional hazard assumption in this hormone receptor positive patient population. All these variables were used for the calculation of the CLP. The optimized combination of these clinicopathological variables into the CLP resulted in a highly prognostic score (HR = 2.72, 95% CI 2.13–3.47; $P < 0.0001$), as was reported for the Clinical Treatment Score (CTS) computed in a similar fashion in the TransATAC study [15]. The inclusion of progesterone receptor status into the model had no substantial effect on the CLP (HR = 2.72, 95% CI 2.18–3.40; $P < 0.0001$).

In spite of this already highly significant prognostic effect of CLP alone on DRFS, the addition of the ROR score provides a highly significant further increase in prognostic information (CLP: HR = 2.09, 95% CI 1.62–2.71, $P < 0.0001$; ROR: HR = 1.03, 95% CI 1.02–1.04, $P < 0.0001$;log-likelihood test: $\Delta LR\chi^2 = 53.49$; $P < 0.0001$; Table 1). The HR of the ROR score corresponds to an increase in the risk of relapse of 37.5% for a 10-point increase in the ROR score. The correlation between CLP and ROR is weak to moderate (Spearman's correlation coefficient: 0.32, $P < 0.0001$).

A highly significant increase in prognostic information is also achieved by adding risk groups derived from the ROR score to the CLP (CLP: HR = 1.85, 95% CI 1.41–2.43, $P < 0.0001$; intermediate versus low risk: HR = 2.15, 95% CI 1.21–3.81, $P = 0.009$; high versus low risk: HR = 4.26, 95%CI, 2.44–7.43, $P < 0.0001$; $\Delta LR\chi^2 = 34.12$; $P < 0.0001$). The 1478 patients are approximately equally distributed across the three ROR-based risk groups. The discrimination ability between ROR-based risk groups is obvious from the DRFS curves, with the high- and the low-risk

Table 1. Additional information of ROR score and risk groups expressed as difference in log-likelihood ($\Delta LR\chi^2$) compared with CLP score alone

Group	Number of		CLP + ROR versus CLP		CLP + risk groups versus CLP	
	Patients	Events	$\Delta LR\chi^2$	*P*-value	$\Delta LR\chi^2$	*P*-value
All patients	1478	155	53.49	<0.0001	34.12	<0.0001
N0	1047	86	25.57	<0.0001	23.36	<0.0001
N+	431	69	29.61	<0.0001	18.30	0.0001
Her2-negative	1397	145	47.50	<0.0001	29.94	<0.0001
Her2-positive	77	10	5.34	0.021	4.41	0.111
N0, Her2-negative	984	79	21.69	<0.0001	20.32	<0.0001
N0, Her2-positive	59	7	2.76	0.097	3.98	0.137
N+, Her2-negative	413	66	27.65	<0.0001	17.45	0.0002
N+, Her2-positive	18	3	2.75	0.098	0.53	0.767

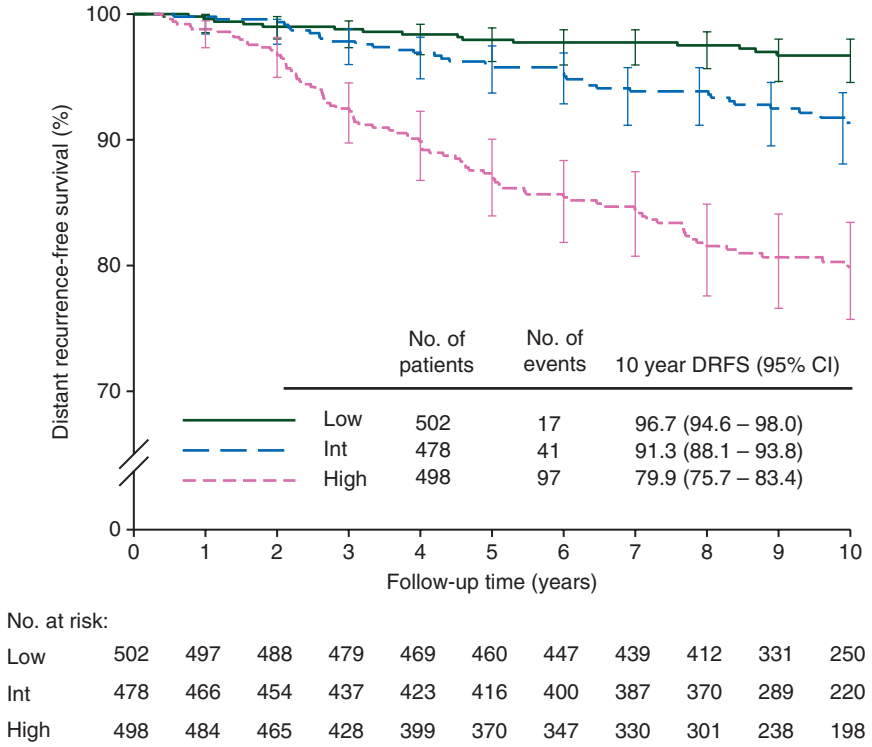

Figure 1. DRFS–Kaplan–Meier plots for the three risk groups with 95% CI.

groups already clearly separated within three years after randomization (Figure 1). The probability for 10-year DRFS is 96.7% (95% CI 94.6–98.0) for the low risk, 91.3% (88.1–93.8) for the intermediate and 79.9% (75.7–83.4) for the high-risk group.

The time-dependent ROC curve for 3-year DRFS markedly differs from a situation of no information represented by the 45° line (supplementary Figure S2, available at *Annals of Oncology* online). The C-index of 0.72 (95% CI 0.68–0.76) indicates a significant correlation between DRFS times predicted by the ROR score and observed DRFS times. The C-index values for prespecified subgroups (stratified by nodal and HER2 status) all significantly differ from the random prediction value of 0.5 with markedly higher prognostic information for nodal-positive compared with nodal-negative patients (supplementary Table S5, available at *Annals of Oncology* online). Additionally, a restricted C-index confined to patients whose ROR score differed only within 5–10 points revealed that there is prognostic information even in small ROR differences (all patients: C-index = 0.58, 95% CI 0.57–0.60; supplementary Table S5, available at *Annals of Oncology* online). C-index values obtained from CLP were always below the respective values based on the ROR score (supplementary Table S6, available at *Annals of Oncology* online).

In all tested subgroups, the ROR score and the ROR-based risk groups significantly add prognostic information to the clinical predictor (Table 1). Among node-negative patients, the prespecified criteria assigned 47% to the low-risk group, 32% to the intermediate risk group, and 21% to the high-risk group (supplementary Figure S3A, available at *Annals of Oncology* online). In contrast, 63% of the nodal-positive patients fall into the high-risk category and only 3% into the low-risk group (supplementary Figure S3B, available at *Annals of Oncology* online). In the latter, no DRFS events occurred during the follow-up period. The same distributions among risk groups observed for all patients, node-negative and node-positive patients hold for the Her2-negative subgroups.

In addition to generating a ROR score, PAM50 assigns an intrinsic subtype to all cases based on the nearest centroid. As expected in a population that was hormone receptor positive by clinical testing, most cases fall into the LumA (1004, 67.9%) or LumB (418, 28.3%) categories. However, as seen previously [13], PAM50 reclassifies a portion of cases into other subtypes, with 48 cases (3.3%) assigned as HER2E and 8 (0.5%) as basal-like (supplementary Table S1, available at *Annals of Oncology* online). DRFS was significantly higher in LumA patients compared with LumB (HR = 2.85; 95% CI 2.04–4.00; $P < 0.0001$; Figure 2). LumA/LumB subtypes add a significant amount of additional prognostic information to CLP ($\Delta LR\chi^2 = 24.42$; $P < 0.0001$; supplementary Table S7, available at *Annals of Oncology* online). This effect is also obvious in node-negative and node-positive subgroups.

The nodal-status specific ROR cutoffs used to define risk groups for node-positive and node-negative patients significantly discriminated between low- and high-risk groups at 10-year DRFS (Table 2). The ROR at 10 years was expected to be <10% for the low-risk group and >20% for the high-risk group (supplementary Table S2, available at *Annals of Oncology* online). However, although the probability of recurrence was significantly less than the predefined 10% in the low-risk group, probability of recurrence was not significantly >20% in the high-risk category for the end point of DRFS (in this relatively low-

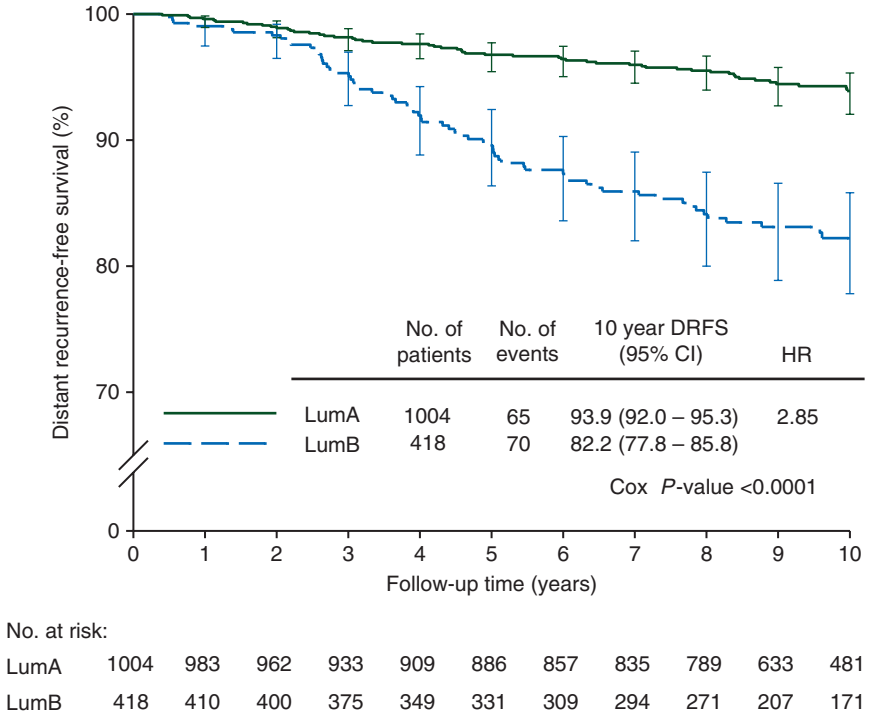

Figure 2. DRFS–Kaplan–Meier plot for luminal A and luminal B subtypes with 95% CI.

risk EBC patient population). Significant discrimination between low- and high-risk groups also occurred within all tested subgroups. Furthermore, the LumA cohort had a significantly lower ROR at 10 years compared with LumB.

discussion

In this large cohort of postmenopausal women with hormone-receptor-positive EBC and long-term follow-up, the PAM50-based ROR score accurately predicted the individual risk of distant recurrence. ROR-defined risk groups and intrinsic breast cancer subtype methods each demonstrate clinically meaningful differences with respect to 10-year risk of metastasis. This finding, in the largest clinical trial cohort where the fully prespecified PAM50 test classifier has yet been applied, is similar to what was observed on the ATAC cohort [15], fulfilling the criteria presented by Simon et al. [28] for level 1 evidence from prospective–retrospective study designs.

PAM50-defined ROR has previously been demonstrated to accurately predict ROR in tamoxifen-treated cohorts [13] as well as in the TransATAC cohort [15]. In the latter, PAM50 was shown to provide more prognostic information than the 21-gene-based RS (Oncotype-Dx 21) [1]. ABCSG-8 extends results to a somewhat different population and highly clinically relevant population, as this large adjuvant trial only included patients with low or intermediate grade tumors, none of whom received adjuvant chemotherapy. The population included represents a subpopulation of patients with luminal disease treated in the overall clinical routine. Although, due to the re-consent process, deceased patients are slightly overrepresented compared with the ABCSG-8 population, and hence DRFS may be overestimated, the study demonstrates the additional prognostic information obtained by the ROR score.

Several multigene expression assays have been developed to estimate the individual ROR of breast cancer patients, including Oncotype®, Endopredict®, and Mammaprint® [R29, R30, 31, 32, R33]. In contrast to Oncotype® and Mammaprint®, NanoString PAM50 can be carried out in any qualified pathology laboratory, eliminating the need for shipping tissue off site and consequent delays in turnaround time. The Nanostring PAM50 measures more genes than Oncotype or Endopredict®, providing subtype information as well as risk score, and does not require frozen tissue. All of these RNA-based methods share the advantage of measuring more genes than can be assessed by immunohistochemistry, in a more objective, reproducible, and robustly quantitative fashion. The comparison of validity and prognostic power between RNA-based methods and immunohistochemistry must be further investigated in the future. PAM50-based ROR adds significant prognostic information beyond classical clinicopathological disease characteristics, and this addition significantly increases the prognostic accuracy, of utmost importance in treatment situations where a patient deemed low risk may have treatments such as cytotoxic chemotherapy withheld.

One limitation of work is that we used (mainly for comparability reasons with TransATAC) a combination score of clinicopathological parameters (CLP)—such indices are always incomplete because they may not include all parameters used by physicians around the world to aid clinical decision-making. Thus, the comparative predictive assessment of multigene testing with, e.g. quantitative receptor content, progesterone receptor status, and lymphovascular invasion remains to be done. Also, we did not compare PAM50-based ROR scores with available online tools such as Adjuvant! Online or PREDICT.

The actual clinical benefit of multigene expression tests is often claimed and heavily discussed. The ultimate prospective

Table 2. Ten-year distant recurrence-free survival (DRFS) with 95% confidence interval in the risk groups and the luminal subtypes

Group	Risk group	*N*	Events	10-year DRFS (%)	
				Estimate	(95% CI)
All patients	Low	502	15	96.7	(94.6–98.0)
	Int	478	35	91.3	(88.1–93.8)
	High	498	87	79.9	(75.7–83.4)
Nodal-negative	Low	487	15	96.6	(94.4–97.9)
	Int	335	28	90.4	(86.3–93.3)
	High	225	32	84.3	(78.4–88.6)
Nodal-positive	Low	15	0	100.0	–
	Int	143	7	93.6	(86.9–97.0)
	High	273	55	76.1	(69.9–81.2)
Her2-negative	Low	489	15	96.6	(94.4–97.9)
	Int	453	34	91.1	(87.7–93.6)
	High	455	79	79.9	(75.6–83.6)
Her2-positive	Low	10	0	100.0	–
	Int	24	1	95.5	(71.9–99.3)
	High	43	8	79.4	(62.8–89.2)
Her2-negative and nodal-negative	Low	474	15	96.5	(94.3–97.9)
	Int	311	27	90.0	(85.6–93.1)
	High	199	27	84.7	(78.4–89.3)
Her2-positive and nodal-negative	Low	10	0	100.0	–
	Int	23	1	95.2	(70.7–99.3)
	High	26	5	80.8	(59.8–91.5)
Her2-negative and nodal-positive	Low	15	0	100.0	–
	Int	142	7	93.6	(86.8–96.9)
	High	256	52	76.2	(69.8–81.4)
All patients	LumA	1004	53	93.9	(92.0–95.3)
	LumB	418	65	82.2	(77.8–85.8)
Nodal-negative	LumA	725	32	95.1	(93.0–96.5)
	LumB	284	32	87.2	(82.3–90.8)
Nodal-positive	LumA	279	21	90.6	(85.9–93.9)
	LumB	134	33	71.0	(61.5–78.6)

Int, intermediate risk.

test of predictive value is currently being carried out in huge trials such as TailoRx (NCT00310180) and MINDACT (NCT00433589), but is unrealistic and inefficient to repeat for the multitude of upcoming genomic test tools. For this reason, a stringent prospective-retrospective study design, repeated in two trials, provides an alternative route to obtain Level 1 evidence [28]. If a multigene expression test—as we demonstrate for PAM50-defined ROR—is consistently able to define a low-risk group with a 10-year metastasis risk of <3.5% in multiple trial populations, it is fair to say that it is highly unlikely that additional chemotherapy would be able to improve this outcome even numerically—not to mention the unfavorable harm/benefit ratio with respect to treatment side-effects. We show that even among some node-positive breast cancers (who nowadays routinely receive adjuvant chemotherapy), we are able to identify patients who have a negligible risk of metastasis—if confirmed this would (and should) change clinical practice. Particularly in endocrine-responsive breast cancer patients with limited background risk [34, R35], this could help avoid unwarranted systematic overtreatment [8]. Avoiding overtreatment is a reasonable way to save considerable health care spending [10].

In summary, we demonstrate that the PAM50 ROR score provides additional prognostic information for DRFS over and above standard clinical variables, using all available patient samples in a large clinical trial cohort. In contrast to other assays currently available on the market, the NanoString PAM50 is designed to be carried out in local qualified hospital pathology laboratories and has been optimized to identify intrinsic breast cancer subtypes while generating a quantitative risk score. This ROR score has now been clinically validated and demonstrated to provide additional prognostic information beyond CTS and some other molecular tests [15, 36]. This added prognostic information should aid physicians in stratifying patients into distinct risk categories with different prognoses among node-negative and even node-positive patient populations, thereby aiding physicians in determining whether additional chemotherapy beyond endocrine therapy is required.

acknowledgements

We thank our patients who contributed to this and other ABCSG trials; ABCSG investigators, study nurses, and data-

management associates, both in the individual trial centers and in the ABCSG center, who provided ongoing support; particularly Hannes Fohler, Britta Klucky, and Katharina Steiner for assistance in project coordination. We also thank Doris Gao and Garrett Barry for technical assistance, and Dr David G. Huntsman, Alicja Parker, and Peggy Tsang for providing a CLIA-certified test environment at the BC Cancer Agency's Centre for Translational and Applied Genomics.

funding

The original clinical trial was in part supported by AstraZeneca. Additional follow-up, re-consent procedures, and database maintenance were supported by NanoString Technologies.

disclosure

PD is currently conducting research sponsored by Sividon and Agendia. FF reports research sponsoring by Roche, Astrazeneca, Novartis, Amgen and Pfizer and travel grants by Roche and Novartis. TN is a consultant for Bioclassifier LLC, which has licensed the PAM50 test to Nanostring Technologies. MK has received research and travel grants from Agendia BV. JWC is an employee of and stock holder in NanoString Technologies. SF is an employee of and a stockholder in NanoString Technologies, the maker of the PAM50 assay used in this study. JS is an employee of and a stockholder in NanoString Technologies, the maker of the PAM50 assay used in this study. RG reports research support and honoraria by Astrazeneca. CSc is a paid consultant of NanoString Technologies. All remaining authors have declared no conflicts of interest.

references

Additional References 'R#' can be found in Supplementary References, available at *Annals of Oncology* online.

1. Paik S, Shak S, Tang G et al. A multigene assay to predict recurrence of tamoxifen-treated, node-negative breast cancer. N Engl J Med 2004; 351: 2817–2826.
3. Early Breast Cancer Trialists' Collaborative Group (EBCTCG). Effects of chemotherapy and hormonal therapy for early breast cancer on recurrence and 15-year survival: an overview of the randomised trials. Lancet 2005; 365: 1687–1717.
8. Gnant M, Steger GG. Fighting overtreatment in adjuvant breast cancer therapy. Lancet 2009; 374: 2029–2030.
9. Gnant M, Harbeck N, Thomssen C. St. Gallen 2011: summary of the consensus discussion. Breast Care 2011; 6: 136–141.
10. Blomer JU, Rezai M et al. Using the 21-gene assay to guide adjuvant chemotherapy decision-making in early-stage breast cancer: a cost-effectiveness evaluation in the German setting. J Med Econ 2013; 16: 30–40.
11. Goldhirsch A, Wood WC, Coates AS et al. Strategies for subtypes—dealing with the diversity of breast cancer: highlights of the St Gallen international expert consensus on the primary therapy of early breast cancer 2011. Ann Oncol 2011; 22: 1736–1747.
12. Parker JS, Mullins M, Cheang MC et al. Supervised risk predictor of breast cancer based on intrinsic subtypes. J Clin Oncol 2009; 27: 1160–1167.
13. Nielsen TO, Parker JS, Leung S et al. A comparison of PAM50 intrinsic subtyping with immunohistochemistry and clinical prognostic factors in tamoxifen-treated estrogen receptor-positive breast cancer. Clin Cancer Res 2010; 16: 5222–5232.
15. Dowsett M, Sestak I, Lopez-Knowles E et al. on behalf of the ATAC and LATTE Trialists Group. Comparison of PAM50 risk of recurrence (ROR) score with oncotype DX and IHC4 for predicting risk of recurrence and distant recurrence after endocrine therapy. J Clin Oncol 2013; 31: 2783–2790.
16. Albain KS, Barlow WE, Shak S et al. Prognostic and predictive value of the 21-gene recurrence score assay in postmenopausal women with node-positive, oestrogen-receptor-positive breast cancer on chemotherapy: a retrospective analysis of a randomised trial. Lancet Oncol 2010; 11: 55–65.
20. Jonat W, Gnant M, Boccard F Effectiveness of switching from adjuvant tamoxifen to anastrozole in postmenopausal women with hormone-sensitive early-stage breast cancer: a meta-analysis. Lancet Oncol. 2006; 7: 991–996.
28. Simon RM, Paik S, Hayes DF. Use of archived specimens in evaluation of prognostic and predictive biomarkers. J Natl Cancer Inst. 2009; 101: 1446–1452.
31. Filipits M, Rudas M, Jakesz R 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: 6012–6020.
32. Dubsky P, Filipits M, Jakesz R et al. Endopredict improves the prognostic classification derived from common clinical guidelines in ER-positive, HER2-negative early breast cancer. Ann Oncol 2012; 0: 1–8.
34. Gnant M, Mlineritsch B, Schippinger W et al. Endocrine therapy plus zoledronic acid in premenopausal breast cancer. N Engl J Med 2009; 360: 679–691, 2009.
35. Jakesz R, Jonat W, Gnant M et al. ABCSG and the GABG. Switching of postmenopausal women with endocrine-responsive early breast cancer to anastrozole after 2 years' adjuvant tamoxifen: combined results of ABCSG trial 8 and ARNO 95 trial. Lancet 2005; 366: 455–462.
36. Sgroi DC, Sestak I, Cuzick J et al. Comparative performance of Breast Cancer Index (BCI) versus oncotype Dx and IHC4 in the prediction of late recurrence in hormonal receptor-positive lymph node-negative breast cancer patients: a TransATAC study. Abstract-San Antonio Breast Cancer Symposium 2012.