HHS Public Access

Author manuscript

Breast Cancer Res Treat. Author manuscript; available in PMC 2016 July 28.

Published in final edited form as:

Breast Cancer Res Treat. 2012 July ; 134(2): 751–757. doi:10.1007/s10549-012-2112-7.

TOP2A RNA Expression and Recurrence in Estrogen Receptor-Positive Breast Cancer

Joseph A. Sparano[1], **Lori J. Goldstein**[2], **Nancy E. Davidson**[3], **George W. Sledge Jr.**[4], and **Robert Gray**[5]

[1]Albert Einstein College of Medicine, Montefiore Medical Center, Bronx, NY

[2]Fox Chase Cancer Center, Philadelphia, PA

[3]University of Pittsburgh Cancer Institute, Pittsburgh, PA

[4]Indiana University School of Medicine, Indianapolis

[5]Eastern Cooperative Oncology Group, Dana Farber Cancer Institute, Harvard University, Boston, MA

Abstract

The purpose of this study is to evaluate the relationship between *TOP2A* RNA expression and recurrence in patients with operable estrogen receptor (ER) positive breast cancer. We evaluated *TOP2A* expression in a pooled analysis of 4 independent data sets with gene expression data including 752 patients with early stage, ER-positive, HER2-negative breast cancer, most of whom received either no adjuvant therapy or endocrine therapy without chemotherapy. We also used an algorithm to simulate the Onco*type* DX Recurrence Score (simRS) and the proliferation component of the Recurrence Score (simPS). Results are expressed as the hazard ratio (HR) for estimates of the effect of a one standard deviation increase in the value of the log gene expression (x + 1SD vs. x) as a continuous function. *TOP2A e*xpression was significantly associated with recurrence (HR 1.56, p<0.0001), and after adjustment for simRS (HR 1.26, p=0.003). *TOP2A* correlated somewhat with simRS (0.45), but more strongly with simPS (0.69). For those with an intermediate simRS, high *TOP2A* expression (above the median) was associated with significantly higher relapse rates at five years (HR 1.82, p=0.007). *TOP2A* expression provides prognostic information in patients with ER-positive, HER2-negative breast cancer, a population known to have low incidence of *TOP2A* gene alterations. These findings confirm prior reports indicating that *TOP2A* expression provides prognostic information in ER-positive breast cancer. *TOP2A* expression may also be useful for identifying those with an intermediate RS who are more likely to relapse, although additional validation in datasets including measured rather than simulated RS will be required.

Keywords

Breast cancer; *TOP2A* RNA expression; recurrence

Corresponding author: Joseph A. Sparano, MD, Montefiore-Einstein Cancer Center, Montefiore Medical Center-Weiler Division, 1825 Eastchester Rd. Bronx, NY 10461; Phone 718-904-2555; Fax 718-904-2892; jsparano@montefiore.org.

Presented in part on December 8 2011 at the 2011 CTRC-AACR San Antonio Breast Cancer Symposium.

Introduction

The Oncotype DX Recurrence Score (RS) is a multiparameter gene expression assay that provides prognostic and predictive information for patients with operable estrogen receptor (ER)-positive breast cancer.[1] The assay provides clinically useful information when the RS is high (>30), indicating that relative and absolute benefit from adding chemotherapy to endocrine therapy is high. It also provides clinically useful information when the RS is low (<18), indicating that risk of distant recurrence is low with endocrine therapy alone, and that benefit from chemotherapy is unlikely.[2] However, RS may be in the "intermediate" range of 18–30 in up to 50% of patients or more, where the risk of recurrence is substantial but the benefit of chemotherapy is uncertain.[3] Several studies have shown that use of the RS results in a change in treatment recommendation in about 20% of patients, usually in the direction of sparing chemotherapy when it otherwise would have been recommended because the RS is low.[4,5] Identifying other genes that provide more accurate prognostic and/or predictive information in patients with ER-positive breast cancer, especially in those who have tumors associated with an intermediate RS, may therefore have the potential for clinical utility.

Several prior reports have shown that topoisomerase 2 alpha (*TOP2A*) expression provides prognostic information in patients with ER-positive breast cancer irrespective of whether adjuvant chemotherapy was used.[6,7] *TOP2A* is not one of the 16 tumor associated genes in the Oncotype DX RS. Topoisomerase 2 alpha is a nuclear enzyme encoded by chromosome 17 that regulates topological changes in DNA by promoting transient double-strand DNA breaks, and is the primary drug target for Topoisomerase II inhibitors such as anthracyclines.[8] There is conflicting evidence whether *TOP2A* expression or gene alterations are predictive of anthracycline response or resistance.[9] We have previously reported that *TOP2A* RNA expression was significantly associated with recurrence in patients with ER-positive, HER2-negative stage I–III breast cancer who received standard doxorubicin-containing chemotherapy plus endocrine therapy in trial E2197, suggesting a link between *TOP2A* RNA expression and anthracycline resistance.[10] We also found that although *TOP2A* RNA expression correlated highly with poor tumor grade, high RS, and the five proliferation genes comprising the "proliferation group" of genes in the RS (*MKI67, BIRC5, CCNB1, MYBL2*, and *AURKA*), it nevertheless remained independently associated with recurrence when adjusted for grade and RS. In addition, high *TOP2A* expression was associated with a significantly increased risk of recurrence in patients with a low RS (hazard ratio [HR] 2.6 for *TOP2A* expression above vs. below the median, p=0.008) and intermediate RS (HR 2.0, p=0.004). Although *TOP2A* expression correlates with proliferation and the S-G2-M phases of the cell cycle, *TOP2A* is also expressed in the G0–G1 phase of the cell cycle in breast cancer, indicating its potential to provide prognostic and predictive information independent of proliferation.[11]

In the current analysis, we sought to validate the prognostic utility of *TOP2A* expression in other independent publicly available data sets with gene expression data, determine its correlation with other proliferation genes, and whether *TOP2A* expression provided prognostic information when the RS measurement was simulated using the gene expression data rather than directly measured in tumor specimens. In addition, because the RS

algorithm truncates gene expression values for the five proliferation genes in calculating the RS (assigns a value of 6.5 units for all tumors with a proliferation score up to 6.5)[1], we evaluated the effect of truncation on the prognostic information provided by *TOP2A* expression, proliferation genes in the RS, and a simulated RS.

Methods

Cohorts

Datasets were selected based on availability of gene expression data, recurrence-related endpoint, and the availability of cases with estrogen receptor (ER)-positive disease (Table 1). The datasets selected included those derived from reports by Wang et al[12] (http://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE2034, accessed on July 14, 2010), van de Vijver et al[13] (http://www.rii.com/publications/2002/nejm, accessed on July 14, 2010), Desmedt et al[14] (http://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE7390, R datasets accessed on July 23, 2010), and Loi et al[1516] (http://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE6532, R datasets accessed on July 23, 2010).

Methods for Calculating Simulated Recurrence Score and TOP2A Expression

For the Wang et al dataset[12], the data include time to distant recurrence and ER status (defined by ligand binding assay or IHC). There were 286 cases with107 having had distant relapse. Gene expression was evaluated using Affymetrix HG U133A arrays. The expression levels in the data set were log transformed (base 2) for analysis. For the van de Vijver dataset[13], the expression values given are log ratios of expression for each sample relative to pooled reference sample (the exact scale is not clear and values appear to be truncated at 2 and −2). The mapping of the oligonucleotide probes to genes provided in the data files was used. SCUBE2 was not available in the data. ER status was defined using an ESR1 log expression ratio cutoff of −0.65. The endpoint used was time to distant metastases (101 cases with distant metastases). The Desmedt et al dataset[14] included 198 case series from the TRANSBIG validation cohort who received no adjuvant therapy. Gene expression was determined using the Affymetrix HG U133A arrays, with the gene expression levels log transformed (base not specified). ER status evaluated by IHC was included in the data. The endpoint of relapse-free survival (RFS) was used here, with 91 events (56 in the first 5 years). Based on nonproportional hazards in the published analyses, follow-up was truncated at 5 years for the analyses below. For the Loi et al dataset[15,16] the full data set from GEO included 268 patients with clinical data, including the endpoint of RFS (13 of these missing distant metastasis-free survival status, apparently excluded from prior publications, were included in the RFS analysis here), of which 5 are ER-negative by ligand binding assay. The latter were excluded, leaving 263 ER-positive cases, of which 87 had RFS events. Only the probe sets from the Affymetrix HG U133A array were used, although the U133B results are also available. The gene expression levels had been log transformed (base not specified) and centered.

Expression levels or ratios as given by the authors were used without renormalization or additional quality control. Mapping of Affymetrix probe sets to genes was taken from http://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GPL96. A simulation of the Oncotype DX

RS was computed (called “simRS”) and proliferation score (called “simPS”) using the formulas from Paik et al[1], even though the scaling of the log expression levels or ratios might be different. The Her2 Axis and the Proliferation Axis also need to be truncated in the calculation. The thresholds for truncation were chosen to give the same proportion of truncated values as in our prior study using samples from the E2197 study (lower 74.1% truncated for Her2, lower 53.5% truncated for proliferation).[10] A gene missing from one of the axes was omitted and the weights adjusted to make up for the omitted gene within the group. Missing values for genes in individual samples were imputed with the mean of the nonmissing values for the gene. If multiple probe sets mapped to the same gene (which is true of for a number of genes on the U133A array), then the probe set with the largest variance was used.

Only ER or PgR positive and HER2 negative cases were included in the analysis, with positivity determined from expression of ESR1, PgR, and ERBB2, respectively (except for hormone receptors for the Loi dataset, since it was a hormone receptor positive cohort). The cutoffs were determined separately for each data set by examining histograms of the expression levels for possible bi-modal splits together with rough ideas on reasonable positive rates. The classification used is provided in Supplemental Table 1.

Statistical methods

Cox’s proportional hazards models were used to estimate the relationships between the two variables *TOP2A* expression and simRS and available time to event endpoints. Since the scaling of the variables is arbitrary, for continuous variables the estimates are expressed as the hazard ratio for a one standard deviation (SD) increase in the value of the variable (that is, for x + one SD vs. x). The same scaling is used for truncated variables as for the corresponding non-truncated, but since the truncation reduces the range, the hazard ratios are not directly comparable between the two. Models that were evaluated included continuous *TOP2A* alone, continuous simRS alone, continuous simPS alone, and joint models with *TOP2A* and simRS, both without and with truncation of *TOP2A* expression and proliferation genes of the simRS (Table 2). Models stratified on data set were also fit to estimate the average effect over the data sets and give overall tests for association. In these models, the continuous variables were standardized to have standard deviation = 1 within each data set.

Kaplan-Meier estimates of 5-year relapse free survival (RFS) or distant relapse free survival (DRFS) rates were computed for high and low *TOP2A* subsets (using the median split within each data set) within simRS risk groups formed using the proportions from the hormone receptor-positive, HER2-negative subset from our previous study (47% low, 34%, intermediate, 19% high) using the classical RS category definitions (low < 18, intermediate 18–30, high > 30) and (17% low, 54% intermediate, 29% high) using the distribution of the TAILORx risk groups (low < 11, intermediate RS 11–25, high > 25).[10] A stratified Cox proportional hazards model was also used to estimate the average *TOP2A* high vs. low hazard ratio within simRS subsets for the combined data set (using *TOP2A* and simRS groups defined separately within each data set).

Results

Characteristics of Patients Included in Datasets Used for the Pooled Analysis

The study population was derived from four previous reports summarized in Table 1 which included a total of 1042 patients, of whom 752 patients with ER-positive, HER2-negative disease were selected for inclusion in this analysis. Of the 752 patients included, 244 (32%) had a recurrence. The study population included primarily patients with node –negative disease who received either no adjuvant therapy or adjuvant endocrine therapy.

Correlation Between TOP2A Expression, Simulated Recurrence Score (simRS), and Simulated Proliferation Score (simPS)

The RS algorithm includes truncation of the PS below 6.5 units, which corresponded to 53.5% truncated for proliferation in our previous analysis of E2197 samples[10], and was used as the proportion who had *TOP2A* and PS thresholded for this analysis. *TOPA* expression without truncation correlated with the simPS without truncation (0.69) but less strongly with simRS (0.45) in which the proliferation score is truncated, and which includes 11 other genes reflecting ER-signaling, HER2 expression, invasion, and other parameters. There was stronger correlation between simPS without truncation and simRS with truncation (0.74) or without truncation (0.93), although this is expected because simPS is a component of the simRS. Also as expected, simRS with truncation correlated highly with simRS without truncation (0.90).

Relationship between Risk of Recurrence and Continuous TOP2A, Simulated Recurrence Score, and Proliferation Axis of Simulated Recurrence Score

The continuous (linear) effects of *TOP2A* expression, simulated RS (simRS), and the proliferation axis of simRS (simPS) were evaluated in models including each variable alone, or individually when evaluated in the same model, as shown in Table 2. The effect of truncation was also evaluated for simRS, simPS, and *TOP2A*. The simRS using truncated values for proliferation genes most closely approximates the actual RS algorithm used for clinical specimens.

TOP2A expression without truncation was significantly associated with recurrence in the pooled dataset (hazard ratio [HR] 1.56, 95% confidence intervals [CI] 1.37,11.78, p< 0.0001), and in each individual dataset (supplemental table 1). There was also a strong association with recurrence with truncation (HR 1.89, 95% CI 1.54, 2.32, p<0.0001).

SimRS with truncation of the PS (which replicates the RS algorithm) was also significantly associated with recurrence in the pooled dataset (HR 1.60, 95% CI 1.44, 1.78, p<0.0001), and in each individual data set (supplemental Table 1). SimRS without truncation of the PS (which does not exactly replicate the RS algorithm) was also significantly associated with recurrence in the pooled dataset (HR 1.77, 95% CI 1.57, 2.00, p<0.0001), and in each individual data set (supplemental Table 1).

SimPS with truncation (which replicates the RS algorithm) was also significantly associated with recurrence in the pooled dataset (HR 2.26, 95% CI 1.87, 2.72, p<0.0001), and appeared

to exhibit the strongest association with recurrence. The relationship was weaker but remained significant without truncation (HR 1.79, 95% CI 1.58, 2.03, p<0.0001).

We next evaluated *TOP2A* and simRS in joint models that allowed us to evaluate the prognostic information provided for each variable after adjustment for the other. When *TOP2A* without truncation was evaluated in a joint model that included simRS with truncation (which replicates the RS algorithm), *TOP2A* provided additional information to simRS (HR 1.26, 95% CI 1.08, 1.46, p=0.003), and simRS likewise provided additional information to *TOP2A* (HR 1.44, 95% CI 1.27. 1.63, p<0.0001). *TOP2A* expression level with truncation was not prognostic in other models evaluated. These models suggest that evaluation of *TOP2A* expression is most likely to provide additional prognostic information to RS when *TOP2A* expression levels are not truncated.

Relationship between Categorical TOP2A, Recurrence Score, and Recurrence

The effects of *TOP2A* expression in a pooled dataset was evaluated as a categorical variable (above vs. below the median *TOP2A* expression) using a classical and TAILORx definitions for low, intermediate, and high RS (Table 3). For an intermediate RS using the classical definition of 18–30, high *TOP2A* expression was associated with a significantly higher risk of recurrence (HR 1.82, p=0.007). There was also a strong trend using the TAILORx definition of 11–25 to define intermediate RS (HR 1.43, p=0.07).

When each dataset was evaluated individually, there was a significantly higher risk of recurrence for patients in the intermediate RS group (18–30) who had high *TOP2A* expression in two of four datasets, including Wang et al (87.1% vs. 49.5%, p=0.02) and Desmedt et al (86.3% vs. 57.9%, p=0.03), both of which included patients who received no adjuvant therapy (Supplementary Table 2). If TAILORx definitions were used, there was a significantly higher risk of recurrence for patients who had high *TOP2A* expression in one of four datasets in the intermediate RS group (Wang et al. 83.6% vs. 60.9%, p=0.04) and one of four in the low RS group (van de Vijver et al, 100% vs. 73.3%, p=0.05). (Supplementary Table 3).

Discussion

Because our previous work had shown that *TOP2A* expression exhibited the strongest association with recurrence in patients with ER-positive, HER2-negative stage I–III breast cancer treated with adjuvant doxorubicin-containing chemotherapy[10], we evaluated the relationship between *TOP2A* RNA expression and recurrence in 4 independent publicly available data sets including 752 patients with early stage, ER-positive, HER2-negative breast cancer. *TOP2A* expression was significantly associated with recurrence in the pooled dataset (HR 1.56, p<0.0001 without truncation in Table 2), and in each of the individual datasets that included patients who received no adjuvant therapy, who received tamoxifen, or who received endocrine therapy plus chemotherapy (supplemental Table 2).

Our results confirm other reports indicating that *TOP2A* expression provides prognostic information in patients with ER-positive operable breast cancer. For example, Rody et al reported that in an analysis of Affymetrix microarray data from 1,681 breast cancers, higher

TOP2A expression significantly correlated with tumor size, poor grade, HER2 expression, and positive lymph nodes, and was associated with a poorer survival in ER-positive, HER2-negative disease (p=0.001), but not ER negative disease.[6] The association between *TOP2A* gene expression and recurrence was independent of whether patients were untreated or had received adjuvant therapy, and was the single most important prognostic factor in a multivariate model (HR 2.40, 95% CI 1.68–3.43, p<0.001). Likewise, Brase et al found that *TOP2A* gene expression was significantly associated with the metastasis-free interval in 782 node-negative breast cancer patients who received no adjuvant therapy.[7]

Although other reports have found a strong association between *TOP2A* expression and recurrence, there was no information about how it compared with RS, or whether it added information to RS. We have previously shown that not only did *TOP2A* expression exhibit the strongest association with recurrence in patients with ER-positive, HER2-negative stage I–III breast cancer treated with adjuvant doxorubicin-containing chemotherapy, but it also provided prognostic information in patients with an intermediate RS that was directly measured in the same clinical specimens.[10] Although we did not have information about RS from the datasets used for the current analysis, we used an algorithm to simulate the RS (simRS) and the 5 proliferation genes (simPS) that are a component of the RS. We observed a highly significant association between recurrence and both simRS and simPS, confirming the validity of our approach in deriving rather than measuring these scores. Although *TOP2A* expression correlated modestly with simRS (0.45) and well with simPS (0.69), it provided prognostic information independent of simRS in joint models when the two variables were considered together in the pooled analysis of all 4 datasets (HR 1.26, p=0.003 without truncation in Table 2). When each dataset was evaluated individually, it was significantly or nearly significantly prognostic in 2 of the 4 dataset (supplemental Table 2).

The TAILORx trial is evaluating whether chemotherapy is beneficial in patients with ER-positive, HER2-negative, node negative breast cancer who have a ‘mid-range” RS. A “mid-range” RS was defined as a RS of 11–25 for the group randomized to chemotherapy or not in this trial rather than “intermediate” range of 18–30 as originally defined in order to reduce the chance of undertreating patients for whom chemotherapy was typically recommended according to National Comprehensive Cancer Center Network guidelines.[17] Because we found in our previous study that *TOP2A* provided complementary prognostic information in patients with a “mid-range” or “intermediate” RS treated with adjuvant doxorubicin-containing therapy[10], we evaluated whether *TOP2A* expression provided information complementary to a simulated RS, since it was not possible to directly measure RS in the clinical specimens used to derive these publicly available datasets. Similar to our previous report in patients who received adjuvant doxorubicin-containing chemotherapy, we found that high *TOP2A* expression (above the median) in patients with a mid-range RS who largely received either no adjuvant therapy or only endocrine therapy was also associated with a higher risk of recurrence (Table 3); a stronger effect seen with the classical definition of 18–30 (HR 1.82, p=0.007) than the TAILORx definition of 11–25 (HR1.43, p=0.07). The difference was driven largely by significant differences seen in two of the four datasets including patients who received either no adjuvant therapy or tamoxifen using the classical definition of 18–30 (supplemental Table 3), and one of the four datasets including patients who received no adjuvant therapy using the TAILORx definition of 11–25 (supplemental

Table 4). Irrespective of the results of the TAILORx trial when available in several years, integration *TOP2A* expression with RS offers the potential to provide more accurate prognostic information, and perhaps more accurate prediction of chemotherapy benefit should it be present in the mid-range RS group.

The RS algorithm truncates gene expression values for the five proliferation genes in calculating the RS by assigning a value of 6.5 units for all tumors with a proliferation score up to 6.5.[1] We evaluated the effect of truncation on the prognostic information provided by *TOP2A* expression, simRS, and simPS. The threshold for truncation of *TOP2A*, simRS, and simPS were chosen to give the same proportion of truncated values as in our prior study using samples from the E2197 study (lower 53.5%).[10] Compared with non-truncated values, truncation was associated with higher hazard ratios for recurrence for *TOP2A* and simPS, but not simRS. The clinical relevance and implications of this are uncertain. Non-truncated *TOP2A* expression was prognostic (HR 1.26, p=0.003 in Table 2) when evaluated in joint models including simRS with truncation (similar to the actual RS algorithm), but not in joint models including simRS without truncation (HR 1.11, p=0.21 in Table 2), suggesting that the additional prognostic information captured by *TOP2A* could also be captured by modification of the proliferation component of the RS algorithm.

There are several strengths and limitations of this analysis. Strengths include the prespecified hypothesis tested in this analysis that was derived from a prior analysis, validation of the prognostic utility of the simulated rather than actual RS and PS, robustness of the association between *TOP2A* expression and recurrence, and other data supporting the prognostic utility of *TOP2A* expression in ER-positive disease. Limitations included multiple assumptions required to derive the datasets, and the fact *TOP2A* expression was prognostic primarily in patients with an intermediate simRS who received no adjuvant endocrine therapy, which does not reflect standard practice. Moreover, our initial observation that *TOP2A* expression was prognostic in patients treated with adjuvant anthracycline-containing therapy, and that *TOP2A* is the target for anthracycline therapy, suggests that *TOP2A* may have a role in identifying patients resistant specifically to adjuvant anthracycline therapy.

In conclusion, our findings provide additional evidence that *TOP2A* expression provides prognostic information in ER-positive early stage breast cancer, and also confirmatory evidence that *TOP2A* expression may be useful for identifying patients with an intermediate RS who are at high risk for relapse. These findings require prospective validation in other prospective or prospective-retrospective trials using the actual rather than simulated RS.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

Supported in part by grants from the Department of Health and Human Services and the National Institutes of Health (CA23318 to the ECOG statistical center).

References

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. [PubMed: 15591335]
2. Paik S, Tang G, Shak S, et al. Gene Expression and Benefit of Chemotherapy in Women With Node-Negative, Estrogen Receptor-Positive Breast Cancer. J Clin Oncol. 2006
3. Sparano JA, Paik S. Development of the 21-gene assay and its application in clinical practice and clinical trials. J Clin Oncol. 2008; 26:721–728. [PubMed: 18258979]
4. Oratz R, Kim B, Chao C, et al. Physician survey of the effect of the 21-gene recurrence score assay results on treatment recommendations for patients with lymph node-positive, estrogen receptor-positive breast cancer. J Oncol Pract. 2011; 7:94–99. [PubMed: 21731516]
5. Lo SS, Mumby PB, Norton J, et al. Prospective multicenter study of the impact of the 21-gene recurrence score assay on medical oncologist and patient adjuvant breast cancer treatment selection. J Clin Oncol. 2010; 28:1671–1676. [PubMed: 20065191]
6. Rody A, Karn T, Ruckhaberle E, et al. Gene expression of topoisomerase II alpha (TOP2A) by microarray analysis is highly prognostic in estrogen receptor (ER) positive breast cancer. Breast Cancer Res Treat. 2009; 113:457–466. [PubMed: 18340528]
7. Brase JC, Schmidt M, Fischbach T, et al. ERBB2 and TOP2A in breast cancer: a comprehensive analysis of gene amplification, RNA levels, and protein expression and their influence on prognosis and prediction. Clin Cancer Res. 2010; 16:2391–2401. [PubMed: 20371687]
8. Liu LF. DNA topoisomerase poisons as antitumor drugs. Annu Rev Biochem. 1989; 58:351–375. [PubMed: 2549853]
9. Oakman C, Moretti E, Sotiriou C, Viale G, Di Leo A. Re: Topoisomerase II alpha and responsiveness of breast cancer to adjuvant chemotherapy. J Natl Cancer Inst. 2009; 101:1735–1736. author reply 6–7. [PubMed: 19893008]
10. Sparano JA, Goldstein LJ, Childs BH, et al. Relationship between Topoisomerase 2A RNA Expression and Recurrence after Adjuvant Chemotherapy for Breast Cancer. Clin Cancer Res. 2009; 15:7693–7700. [PubMed: 19996222]
11. Villman K, Stahl E, Liljegren G, Tidefelt U, Karlsson MG. Topoisomerase II-alpha expression in different cell cycle phases in fresh human breast carcinomas. Mod Pathol. 2002; 15:486–491. [PubMed: 12011253]
12. Wang Y, Klijn JG, Zhang Y, et al. Gene-expression profiles to predict distant metastasis of lymph-node-negative primary breast cancer. Lancet. 2005; 365:671–679. [PubMed: 15721472]
13. van de Vijver MJ, He YD, van't Veer LJ, et al. A gene-expression signature as a predictor of survival in breast cancer. N Engl J Med. 2002; 347:1999–2009. [PubMed: 12490681]
14. Desmedt C, Piette F, Loi S, et al. Strong time dependence of the 76-gene prognostic signature for node-negative breast cancer patients in the TRANSBIG multicenter independent validation series. Clin Cancer Res. 2007; 13:3207–3214. [PubMed: 17545524]
15. Loi S, Haibe-Kains B, Desmedt C, et al. Definition of clinically distinct molecular subtypes in estrogen receptor-positive breast carcinomas through genomic grade. J Clin Oncol. 2007; 25:1239–1246. [PubMed: 17401012]
16. Loi S, Haibe-Kains B, Desmedt C, et al. Predicting prognosis using molecular profiling in estrogen receptor-positive breast cancer treated with tamoxifen. BMC Genomics. 2008; 9:239. [PubMed: 18498629]
17. Sparano JA. TAILORx: trial assigning individualized options for treatment (Rx). Clin Breast Cancer. 2006; 7:347–350. [PubMed: 17092406]

Table 1

Characteristics of datasets

Reference	Total No. Patients (Events)	No. Patients (Events) Included in Analysis	Key characteristics	Adjuvant therapy
Wang et al	286 (107)	192 (73)	Node negative	None
Van de Vijver et al	295(101)	212 (68)	Stage 1–II, < 53 years	Chemotherapy (90), hormones (20), both (20), or none (165)
Desmedt et al	198 (91)	124 (33)	Node negative	None
Loi et al	263 (87)	224 (70)	Stage I–II, ER+	Tamoxifen

Table 2

Estimated hazard ratios and 95% confidence intervals for *TOP2A*, simulated recurrence score (simRS), and simulated proliferation score (simPS), with and without truncation.

Model Variable(s)	Variable Evaluated	Estimated Ratio	95% Confidence Intervals	P-value
TOP2A With truncation	*TOP2A* With truncation	1.89	1.54, 2.32	<0.0001
TOP2A Without truncation	*TOP2A* Without truncation	1.56	1.37, 1.78	<0.0001
simRS With truncation	simRS With truncation	1.60	1.44, 1.78	<0.0001
simRS Without truncation	simRS Without truncation	1.77	1.57, 2.00	<0.0001
SimPS With truncation	SimPS With truncation	2.26	1.87, 2.72	<0.0001
SimPS Without truncation	SimPS Without truncation	1.79	1.58, 2.03	<0.0001
TOP2A without truncation + simRS with truncation	*TOP2A* Without truncation	1.26	1.08, 1.46	0.003
	Sim RS With truncation	1.44	1.27, 1.63	<0.0001
TOP2A with truncation + simRS with truncation	*TOP2A* With Truncation	1.23	0.95, 1.58	0.11
	SimRS With truncation	1.51	1.33, 1.72	<0.0001
TOP2A with truncation + simRS without truncation	*TOP2A* With truncation	1.10	0.85, 1.42	0.48
	Sim RS Without truncation	1.73	1.50, 1.99	<0.0001
TOP2A without truncation + simRS without truncation	*TOP2A* Without truncation	1.11	0.94, 1.31	0.21
	SimRS Without truncation	1.67	1.43, 1.94	<0.0001

Table 3

Effect of categorical *TOP2A* expression (high vs. low) by simRS risk groups

	Classical Definition	**TAILORx Definition**
	Low <18 Intermediate 18–30 High > 30	Low < 11 Intermediate 11–25 High > 25
Low RS	1.25 (0.77, 2.03) p=0.37	1.04 (0.38, 2.82) p=0.94
Intermediate RS	1.82 (1.18, 2.81) p=0.007	1.43 (0.97, 2.10) p=0.07
High RS	0.71 (0.40, 1.23) p=0.22	1.21 (0.77, 1.91) p=0.41