The value of multigene predictors of clinical outcome in breast cancer: an analysis of the evidence

Amalia M Issa\textsuperscript{ab}, Vivek S Chaudhari\textsuperscript{ab} & Gary E Marchant\textsuperscript{c}

\textsuperscript{a} Program in Personalized Medicine and Targeted Therapeutics, University of the Sciences, 600 South 43rd Street, Philadelphia, PA 19104, USA

\textsuperscript{b} Department of Health Policy and Public Health, University of the Sciences, 600 South 43rd Street, Philadelphia, PA 19104, USA

\textsuperscript{c} Sandra Day O’Connor College of Law, Arizona State University, PO Box 877906, Tempe, AZ 85827-7906, USA

Published online: 03 Jun 2015.

To cite this article: Amalia M Issa, Vivek S Chaudhari & Gary E Marchant (2015) The value of multigene predictors of clinical outcome in breast cancer: an analysis of the evidence, Expert Review of Molecular Diagnostics, 15:2, 277-286

To link to this article: http://dx.doi.org/10.1586/14737159.2015.983476
The value of multigene predictors of clinical outcome in breast cancer: an analysis of the evidence

**Amalia M Issa**1,2, Vivek S Chaudhari1,2 and Gary E Marchant3

1Program in Personalized Medicine and Targeted Therapeutics, University of the Sciences, 600 South 43rd Street, Philadelphia, PA 19104, USA
2Department of Health Policy and Public Health, University of the Sciences, 600 South 43rd Street, Philadelphia, PA 19104, USA
3Sandra Day O’Connor College of Law, Arizona State University, PO Box 877906, Tempe, AZ 85287-7906, USA

*Author for correspondence: a.issa@usciences.edu*

**Objective:** Multigene predictors are being used increasingly in early-stage breast cancer patients for prediction and prognosis. However, one consequence of the increased use of multigene predictors, and the heightened efforts toward their incorporation into routine clinical practice, is the potential for future malpractice litigation. It is, therefore, important to ascertain the strength of the evidence for using the different commercially available multigene predictor assays clinically. We evaluated the literature for evidence of clinical validity of four currently available gene signatures and to assess the influence of the 21-gene-expression assay on changes in treatment recommendations. Methods: A systematic search of the peer-reviewed literature from January 2002 to March 2014 for multigene predictor assays was carried out, and a meta-analysis was conducted. Results: The adjusted Cox hazard ratio average for studies that met the eligibility criteria was 3.538 (95% CI: 1.513–8.469). The 21-gene signature showed the highest stability in the estimation of likelihood of distant risk of recurrence. Using the recurrence scores resulted in changes in treatment recommendations in 31.8% of all patients in the studies. Conclusion: This study may provide insight about the use of multigene predictors in clinical practice for prediction and prognosis of breast cancer.

**KEYWORDS:** breast cancer • Cox hazard ratios • gene-expression profiling • meta-analysis • multigene predictors • recurrence score

Although advances in molecular diagnostics and therapeutic interventions have led to decline in breast cancer mortality, particularly over the last 15 years, breast cancer remains a frequently diagnosed cancer among women [1]. It has been estimated that a total of 232,670 new cases will be diagnosed in the US in 2014 [1].

A key challenge for physicians and their patients in the management of breast cancer in making decisions about the treatment options relates to the use of adjuvant chemotherapy (CHT) [2,3], following surgical resection of the tumor. Adjuvant CHT has been shown to increase recurrence-free and overall survival in early-stage patients as a whole [4], but it is difficult to identify the patients who would not have experienced recurrence in the first place and do not need CHT. The effectiveness of these interventions is based on risk stratification of patients and has involved the use of validated computer-based models using algorithms designed to calculate the rates for recurrence of the disease, such as the Adjuvant! Online [5] developed using Surveillance, Epidemiology, and End Results registry data; the Nottingham Prognostic Index [6]; and/or the utilization of panel recommendations or clinical guidelines issued by scientific or professional organizations such as St. Gallen Consensus [7] and the National Comprehensive Cancer Network [8].

In more recent years, a number of molecular multigene predictor gene-expression profiling (GEP) assays have been developed and increasingly used to identify early-stage breast cancer patients who are most likely to benefit from adjuvant CHT [9].

As a consequence of the increased use of multigene predictors and other molecular companion diagnostics, clinicians, policy...
markers and other decision makers are working toward developing and managing healthcare systems that incorporate personalized genomic medicine into routine clinical practice. This, increased use of multigene predictors and other molecular diagnostics has the potential to affect future malpractice litigation risks [10,11].

It is, therefore, important to ascertain the strength of the evidence for using the different commercially available GEP assays based upon whether these tests provide clinically valid and useful information and the influence of the 21-gene-expression assay, commercially marketed as Oncotype DX® (Genomic Health Inc., Redwood City, CA, USA), on physician treatment recommendations, since this is the most frequently used multigene predictor GEP assay in the US.

Several studies have been published evaluating the use of the Oncotype DX recurrence score (RS) on physician recommendations for treatment options (adjuvant CHT and/or hormonal therapy) [12] or whether the gene signatures of interest were comparable in terms of their clinically validity of the estimations for risk of recurrence. However, these studies primarily evaluated how knowledge of the RS affects management of treatment for eligible early-stage breast cancer patients and did not analyze gene-signature adjusted multivariate Cox hazard ratios to allow more direct comparisons of the prognostic value of the gene signature for the likelihood of 10-year relapse-free survival and overall survival of breast cancer between published reports.

We systematically reviewed the peer-reviewed published literature and conducted a meta-analysis in order to determine whether sufficient evidence exists in the literature to substantiate the clinical validity of GEP to predict recurrence and assess the influence of the 21-gene-expression assay on changes in physician treatment recommendations. Given the increasing need for clinicians and patients, and potentially representatives of the legal community, to interpret findings from the literature to make informed decisions about the use of GEP assays in breast cancer care, we conducted a systematic review of the literature and meta-analysis aimed at summarizing and synthesizing the current evidence.

**Methods**

**Study identification & selection**

We systematically searched the peer-reviewed literature from January 2002 through March 2014 for multigene predictor assays. Searches were performed using several databases including PubMed Medline, Ovid (including Medline and Journals@Ovid Full Text) and EMBASE for studies of multigene predictors of clinical outcome in breast cancer. We independently reviewed a list of titles and abstracts for relevance to identify the articles that describe either a validation of any GEP assay for early-stage breast cancer and/or the use of the RS to make or change treatment recommendations in breast cancer. Articles reviewed were limited to those published in the English language, and were excluded if they lacked relevance to GEP for breast cancer or did not deal with human populations. The full text of any article that passed this initial screening was retrieved for further review. Full articles were also retrieved for any article for which it was not possible to determine relevance from the abstract, in order to ensure a comprehensive search. The references of relevant published articles were also manually searched to identify other potentially relevant articles.

Studies were included for further review if they met the following criteria: original research studies published in peer-reviewed journals; studies focusing on validation of any studies on any of the specific gene-signatures; studies of clinical utilization of the 21-gene expression assay (Oncotype DX®; Genomic Health, Redwood City, CA) recurrence score to make treatment recommendations for patients. Reviews of the literature or meta-analyses that did not include original research were excluded.

**Data extraction & analysis**

Data from each of the articles that were selected for further review were abstracted onto a structured abstraction form. Abstractive data included: authors’, year of publication, total number of study participants, gene signature adjusted multivariate Cox model hazard ratio and the respective 95% CI, total number of patients with changes using RS for treatment recommendations, total number of patients changing from a combination of CHT plus hormone therapy (HT) to HT alone or from HT alone to a combination of CHT. Data were also extracted about the type of study (prospective or retrospective) and which clinical practice guidelines (if this information was available) and/or algorithm-based estimators of risk (such as Adjuvant! Online and/or the Nottingham Prognostic Index) were used for initial treatment recommendations.

The information available for adjusted multivariate Cox hazard ratios and their 95% CI for different gene signatures was pooled, averaged by gene signature and weighted (using a random effects model). Statistical heterogeneity was calculated using the χ² test for heterogeneity and the I² percentage. A probability level for the χ² statistic less than or equal to 10% (p ≤ 0.10) and/or an I² greater than 50% would be considered indicative of statistical heterogeneity.

Data about differences between treatment recommendations before and after the availability of RSs were taken directly from the reviewed articles, pooled and averaged to ascertain how the RSs were used to change initial treatment recommendations. Effect sizes for differences in treatment recommendations were calculated using Cohen’s d and the two-tailed correlation coefficient, r, using means and standard deviations of two independent groups of equal size. Data were analyzed using SPSS® Statistics Release 21 (IBM®, Armonk, NY, USA).

**Results**

Our initial broad search of the literature yielded 590 articles (FIGURE 1). Of the remaining 52 studies that met the eligibility criteria, 25 studies were eligible for a meta-analysis of the gene signature associated standardized multivariate Cox hazard ratios. Table 1 summarizes the characteristics of the included studies. These 25 studies used four gene signatures: MammaPrint®
(44%; n = 11), Mammostrat® (8.0%; n = 2), OncoType DX (24.0%; n = 6) and the Rotterdam signature (24.0%; n = 6). The sample size of these studies ranged between 86 and 4598 subjects [13,14] with a mean of 702 patients enrolled. We contrasted the prognostic value of the likelihood of 10-year relapse-free survival and overall survival of breast cancer by gene signatures included in this meta-analysis by comparing the reports of adjusted multivariate Cox hazards ratios. The adjusted Cox hazard ratio average for the series of 25 studies was 3.538 (95% CI: 1.513–8.469, Table 1). Point estimates and their 95% CIs showed some degree of overlap.

In our analysis, the 21-gene signature OncoType DX showed the highest stability in the estimation of likelihood (mean: 2.70; 95% CI: 1.56–4.72). Around this range of variation in the adjusted multivariate Cox hazards ratios, the closest one to OncoType DX was for MammaPrint (mean: 3.43; 95% CI: 1.43–8.64) followed by Mammostrat (mean: 1.61; 95% CI: 1.30–2.02). The Rotterdam 76-gene signature showed the farthest variation overall (mean: 5.21; 95% CI: 18.50) (Figure 2).

Our search of the literature for evidence of clinical utility of GEP in breast cancer care ultimately yielded 27 studies, all focused on the 21-GEP assay, OncoType DX. We limited our analysis to studies that reported changes in treatment recommendations from initial recommendations based on using the OncoType DX RS. Twenty-seven studies reporting changes in initial treatment recommendations following the use of RS met the eligibility criteria (Table 2) and included 6655 patients with a mean of 246.5 participants per study (standard deviation: 377.2; 95% CI: 104.2–388.8). We found wide variation in the number of participants, from the largest sample of 1822 patients [15] to the smallest sample of 22 patients [16]. More than two-thirds of the studies (70.4%; n = 19) were published between 2011 and 2014. Three-fourths of the studies were retrospective (74.1%; n = 20). Studies reported that initial treatment recommendations for adjuvant CHT were formulated following clinical practice guidelines endorsed by professional associations, mainly the National Comprehensive Cancer Network (35.0%; n = 14), St. Gallen Consensus (15.0%; n = 6) or the American Society of Clinical Oncology (12.5%; n = 5). Several studies reported using algorithms for risk estimation, such as Adjuvant! Online (17.5%; n = 7) and the Nottingham Prognostic Index (5.0%; n = 2).

Using the RSs resulted in changes in treatment recommendations from initial treatment recommendations in 31.8% (n = 2113) of all patients in the studies. A reduction in the administration of CHT but maintenance of hormonal therapy was observed in the majority of study patients (83.3%; n = 1761). However, addition of CHT to hormonal therapy was found in less than one-fifth of the pooled study patients (16.4%; n = 346). These changes from the initial treatment recommendations were statistically significant (paired t(26) = 5.82, p < 0.001, two-tailed; Cohen’s d = 2.0; effect size r = 0.715; coefficient of determination r² = 51.1%), suggesting that treatment recommendations based on use of the OncoType DX RSs resulted in less administration of CHT. Reported percentages in changes from initial treatment recommendations about systemic adjuvant CHT varied from less than 20% up to 74% of all pooled study patients [17–19].

Discussion
We evaluated the evidence in the peer-reviewed literature for whether the currently available gene signatures for early-stage breast cancer are associated with the likelihood of distant recurrence of breast cancer based upon reported adjusted...
multivariate Cox hazard ratios and the strength of evidence for influence of the Oncotype DX RS from the 21-GEP assay on changes in physician treatment recommendations from initial recommendations.

The four gene signatures evaluated appear to be associated with predicting distant recurrence and point estimates and their 95% CIs showed some degree of overlap. This supports the suggestion by some authors that perhaps combining one or more of the already developed multigene predictors will likely provide more precise prognostic and predictive information about recurrence risk with concomitant therapeutic outcomes, and one such attempt has been reported [20–22].

In our analysis, the 21-gene signature Onco

Table 1. Weighted adjusted multivariate Cox hazard ratio and 95% CI.

<table>
<thead>
<tr>
<th>Study (year)</th>
<th>Patient enrollment</th>
<th>Cox HR</th>
<th>Standard error</th>
<th>Weighted CI lower</th>
<th>Weighted CI upper</th>
<th>Ref.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mook et al. (2010)</td>
<td>964</td>
<td>3.450</td>
<td>0.050</td>
<td>3.333</td>
<td>3.567</td>
<td>[43]</td>
</tr>
<tr>
<td>Knauer et al. (2010)</td>
<td>541</td>
<td>0.210</td>
<td>0.020</td>
<td>0.171</td>
<td>0.249</td>
<td>[44]</td>
</tr>
<tr>
<td>Kok et al. (2009)</td>
<td>364</td>
<td>2.210</td>
<td>0.078</td>
<td>2.057</td>
<td>2.363</td>
<td>[32]</td>
</tr>
<tr>
<td>Buyse et al. (2006)</td>
<td>307</td>
<td>1.500</td>
<td>0.067</td>
<td>1.363</td>
<td>1.637</td>
<td>[45]</td>
</tr>
<tr>
<td>van de Vijer et al. (2002)</td>
<td>295</td>
<td>4.600</td>
<td>0.125</td>
<td>4.355</td>
<td>4.845</td>
<td>[46]</td>
</tr>
<tr>
<td>Bueno-de-Mesquita et al. (2009)</td>
<td>274</td>
<td>5.400</td>
<td>0.142</td>
<td>5.125</td>
<td>5.675</td>
<td>[47]</td>
</tr>
<tr>
<td>Mook et al. (2009)</td>
<td>241</td>
<td>2.990</td>
<td>0.111</td>
<td>2.772</td>
<td>3.208</td>
<td>[48]</td>
</tr>
<tr>
<td>Saghatchian et al. (2013)</td>
<td>173</td>
<td>2.698</td>
<td>0.124</td>
<td>2.453</td>
<td>2.943</td>
<td>[49]</td>
</tr>
<tr>
<td>Knauer et al. (2010)</td>
<td>168</td>
<td>5.780</td>
<td>0.185</td>
<td>5.416</td>
<td>6.144</td>
<td>[50]</td>
</tr>
<tr>
<td>Glas et al. (2006)</td>
<td>162</td>
<td>5.600</td>
<td>0.186</td>
<td>5.236</td>
<td>5.964</td>
<td>[51]</td>
</tr>
<tr>
<td>Mook et al. (2010)</td>
<td>148</td>
<td>4.400</td>
<td>0.172</td>
<td>4.062</td>
<td>4.738</td>
<td>[52]</td>
</tr>
<tr>
<td>Avg. MammaPrint®</td>
<td>330.6</td>
<td>3.531</td>
<td>0.115</td>
<td>3.304</td>
<td>3.757</td>
<td></td>
</tr>
<tr>
<td>Bartlett et al. (2012)</td>
<td>4598</td>
<td>1.910</td>
<td>0.020</td>
<td>1.870</td>
<td>1.950</td>
<td>[14]</td>
</tr>
<tr>
<td>Ross et al. (2008)</td>
<td>1294</td>
<td>1.400</td>
<td>0.032</td>
<td>1.336</td>
<td>1.464</td>
<td>[53]</td>
</tr>
<tr>
<td>Avg. Mammostrat</td>
<td>2946.0</td>
<td>1.655</td>
<td>0.026</td>
<td>1.603</td>
<td>1.707</td>
<td></td>
</tr>
<tr>
<td>Partin et al. (2011)</td>
<td>1674</td>
<td>2.160</td>
<td>0.036</td>
<td>2.090</td>
<td>2.230</td>
<td>[19]</td>
</tr>
<tr>
<td>Gray et al. (2011)</td>
<td>1436</td>
<td>1.430</td>
<td>0.062</td>
<td>1.368</td>
<td>1.492</td>
<td>[55]</td>
</tr>
<tr>
<td>Dowsett et al. (2010)</td>
<td>1372</td>
<td>4.360</td>
<td>0.042</td>
<td>4.250</td>
<td>4.470</td>
<td>[56]</td>
</tr>
<tr>
<td>Tang et al. (2011)</td>
<td>1319</td>
<td>2.370</td>
<td>0.052</td>
<td>2.287</td>
<td>2.453</td>
<td>[22]</td>
</tr>
<tr>
<td>Paik et al. (2006)</td>
<td>651</td>
<td>1.780</td>
<td>0.086</td>
<td>1.678</td>
<td>1.882</td>
<td>[57]</td>
</tr>
<tr>
<td>Albain et al. (2010)</td>
<td>367</td>
<td>2.640</td>
<td>0.151</td>
<td>2.474</td>
<td>2.806</td>
<td>[40]</td>
</tr>
<tr>
<td>Avg. Oncotype DX</td>
<td>1136.5</td>
<td>2.457</td>
<td>0.071</td>
<td>2.358</td>
<td>2.556</td>
<td></td>
</tr>
<tr>
<td>Zhang et al. (2009)</td>
<td>300</td>
<td>6.130</td>
<td>0.143</td>
<td>5.850</td>
<td>6.410</td>
<td>[54]</td>
</tr>
<tr>
<td>Wang et al. (2005)</td>
<td>286</td>
<td>5.550</td>
<td>0.139</td>
<td>5.277</td>
<td>5.823</td>
<td>[58]</td>
</tr>
<tr>
<td>Desmedt et al. (2008)</td>
<td>198</td>
<td>5.110</td>
<td>0.161</td>
<td>4.795</td>
<td>5.425</td>
<td>[21]</td>
</tr>
<tr>
<td>Desmedt et al. (2007)</td>
<td>198</td>
<td>2.800</td>
<td>0.119</td>
<td>2.567</td>
<td>3.033</td>
<td>[59]</td>
</tr>
<tr>
<td>Foekens et al. (2006)</td>
<td>180</td>
<td>11.360</td>
<td>0.251</td>
<td>2.670</td>
<td>48.400</td>
<td>[60]</td>
</tr>
<tr>
<td>Prat et al. (2012)</td>
<td>1380</td>
<td>1.490</td>
<td>0.015</td>
<td>1.426</td>
<td>1.554</td>
<td>[61]</td>
</tr>
<tr>
<td>Avg. Rotterdam 76</td>
<td>423.7</td>
<td>5.407</td>
<td>0.138</td>
<td>3.764</td>
<td>11.774</td>
<td></td>
</tr>
<tr>
<td>Effect summary</td>
<td>1209.2</td>
<td>3.262</td>
<td>0.088</td>
<td>2.757</td>
<td>4.949</td>
<td></td>
</tr>
</tbody>
</table>

HR: Hazard ratio.
including tumor genomics (e.g., Ki67, human epidermal growth factor receptor 2 biomarker status). It further suggests the clinical validity of Onco\textsuperscript{type} DX as a predictive and prognostic diagnostic for making treatment decisions for breast cancer\cite{23}. Although other authors have identified these gene signatures as predicting clinical outcome\cite{23}, to the best of our knowledge, this study is the first to pool together and statistically analyze standardized adjusted Cox hazard ratios, which are estimations of the probability of 10-year recurrence.

Our findings that treatment recommendations were changed based on using the RS in more than 30% of patients in the pooled analysis of studies are consistent with other individual studies which found treatment recommendations changed in some 24–30% of women with node-negative cancer following RS use\cite{24–27}. Our analysis demonstrated that changes in treatment recommendations mostly resulted in proportionally reduced administration of CHT in the majority of cases and less frequently in the addition of CHT. This result is not surprising, since CHT is an accepted standard therapeutic intervention in breast cancer\cite{28}. However, prior individual studies had only found changes in treatment recommendations varied from 9.6 to 71 with a mean of 29.8\cite{29,30}. Our finding may have possible policy-relevant consequences for clinical utilization, testing costs and potential future liability cases.

![Forest plots showing disease-free survival. Squares: Reported hazard ratios from each study; Diamonds: Each gene signature’s average and the overall average of the hazard ratio.](image)
Table 2. Proportional changes to treatment recommendations following utilization of the 21-gene signature recurrence score.

<table>
<thead>
<tr>
<th>Study (year)</th>
<th>Type of study</th>
<th>Country of publication</th>
<th>Number of cases/study participants</th>
<th>Guidelines or decision algorithm used pre-RS</th>
<th>Receipt of adjuvant therapy</th>
<th>% Change from CHT to HT</th>
<th>% Change from HT to CHT</th>
<th>% Change total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Oratz et al. (2007)</td>
<td>Retrospective</td>
<td>USA</td>
<td>68</td>
<td>AO</td>
<td></td>
<td>25.0</td>
<td>25.0</td>
<td>25.0</td>
</tr>
<tr>
<td>Erb et al. (2007)</td>
<td>Retrospective</td>
<td>USA</td>
<td>124</td>
<td>NCCN</td>
<td></td>
<td>30.0</td>
<td>30.0</td>
<td>30.0</td>
</tr>
<tr>
<td>Lo et al. (2010)</td>
<td>Prospective</td>
<td>USA</td>
<td>89</td>
<td>CPC</td>
<td></td>
<td>22.5</td>
<td>3.4</td>
<td>25.9</td>
</tr>
<tr>
<td>Asad et al. (2008)</td>
<td>Retrospective</td>
<td>USA</td>
<td>85</td>
<td>NCCN</td>
<td></td>
<td>38.8</td>
<td>4.7</td>
<td>43.5</td>
</tr>
<tr>
<td>Henry et al. (2009)</td>
<td>Retrospective</td>
<td>USA</td>
<td>29</td>
<td>AO</td>
<td></td>
<td>25.9</td>
<td>7.4</td>
<td>33.3</td>
</tr>
<tr>
<td>Geffen et al. (2009)</td>
<td>Retrospective</td>
<td>Israel</td>
<td>25</td>
<td>AO</td>
<td></td>
<td>24.0</td>
<td>12.0</td>
<td>36.0</td>
</tr>
<tr>
<td>Klang et al. (2010)</td>
<td>Retrospective</td>
<td>Israel</td>
<td>313</td>
<td>NCCN and ASCO</td>
<td></td>
<td>33.5</td>
<td>6.4</td>
<td>39.9</td>
</tr>
<tr>
<td>Ademuyiwa et al. (2011)</td>
<td>Retrospective</td>
<td>USA</td>
<td>276</td>
<td>AO and NPI</td>
<td></td>
<td>31.9</td>
<td>6.2</td>
<td>38.0</td>
</tr>
<tr>
<td>Partin et al. (2011)</td>
<td>Retrospective</td>
<td>USA</td>
<td>169</td>
<td>St Gallen; NCCN and AO</td>
<td></td>
<td>71.0</td>
<td>3.0</td>
<td>74.0</td>
</tr>
<tr>
<td>Oratz et al. (2011)</td>
<td>Prospective</td>
<td>USA</td>
<td>89</td>
<td>NCCN and ASCO</td>
<td></td>
<td>35.0</td>
<td>CPC</td>
<td>35.0</td>
</tr>
<tr>
<td>Joh et al. (2011)</td>
<td>Retrospective</td>
<td>USA</td>
<td>154</td>
<td>NCCN and ASCO</td>
<td></td>
<td>24.9</td>
<td>CPC</td>
<td>24.9</td>
</tr>
<tr>
<td>Haas et al. (2011)</td>
<td>Retrospective</td>
<td>USA</td>
<td>364</td>
<td>NCCN</td>
<td></td>
<td>61.8</td>
<td>0.3</td>
<td>62.1</td>
</tr>
<tr>
<td>Hassett et al. (2012)</td>
<td>Retrospective</td>
<td>USA</td>
<td>1111</td>
<td>St Gallen and NCCN</td>
<td></td>
<td>9.6</td>
<td>8.0</td>
<td>17.6</td>
</tr>
<tr>
<td>Eiermann et al. (2012)</td>
<td>Prospective</td>
<td>Germany</td>
<td>366</td>
<td>CPC</td>
<td></td>
<td>21.6</td>
<td>10.6</td>
<td>33.0</td>
</tr>
<tr>
<td>Albanell et al. (2012)</td>
<td>Retrospective</td>
<td>Spain</td>
<td>107</td>
<td>St Gallen and ESMO</td>
<td></td>
<td>20.6</td>
<td>11.2</td>
<td>31.8</td>
</tr>
<tr>
<td>Kamal et al. (2011)</td>
<td>Retrospective</td>
<td>USA</td>
<td>186</td>
<td>St Gallen and NCCN</td>
<td></td>
<td>10.8</td>
<td>8.1</td>
<td>18.9</td>
</tr>
<tr>
<td>Schneider and Khalil (2012)</td>
<td>Retrospective</td>
<td>USA</td>
<td>89</td>
<td>NCCN; ASCO and AO</td>
<td></td>
<td>39.3</td>
<td>4.5</td>
<td>43.8</td>
</tr>
<tr>
<td>Malo et al. (2012)</td>
<td>Retrospective</td>
<td>USA</td>
<td>118</td>
<td>NCCN</td>
<td></td>
<td>42.4</td>
<td>0.8</td>
<td>43.2</td>
</tr>
<tr>
<td>deBoer et al. (2013)</td>
<td>Prospective</td>
<td>Australia</td>
<td>151</td>
<td>St Gallen</td>
<td></td>
<td>35.8</td>
<td>15.5</td>
<td>24.5</td>
</tr>
<tr>
<td>Chen et al. (2013)</td>
<td>Retrospective</td>
<td>USA</td>
<td>1822</td>
<td>NCCN and ASCO</td>
<td></td>
<td>23.3</td>
<td>3.1</td>
<td>26.4</td>
</tr>
<tr>
<td>Holt et al. (2013)</td>
<td>Prospective</td>
<td>UK</td>
<td>142</td>
<td>AO and NPI</td>
<td></td>
<td>18.3</td>
<td>8.5</td>
<td>26.8</td>
</tr>
<tr>
<td>Stemmer et al. (2013)</td>
<td>Retrospective</td>
<td>Israel</td>
<td>282</td>
<td>St Gallen</td>
<td>CPC</td>
<td>45.6</td>
<td>CPC</td>
<td>45.6</td>
</tr>
<tr>
<td>Davidson et al. (2013)</td>
<td>Prospective</td>
<td>Canada</td>
<td>150</td>
<td>CPC</td>
<td></td>
<td>20.0</td>
<td>15.0</td>
<td>35.0</td>
</tr>
</tbody>
</table>

Although several reviews have recently evaluated the literature related to multigene predictors of clinical outcome in breast cancer [31–33], these did not specifically statistically pool and analyze the proportional changes in treatment recommendations that were made as a result of using RSs. Our study suggests that GEP tests, particularly the use of the 21-gene assay, Oncotype DX, are increasingly influencing how breast cancer patients are treated.

Although studies that directly show an association with outcomes following the use of multigene predictors are limited, evidence is accumulating that tests such as the 21-gene signature (Oncotype DX) and the 70-gene signature (MammaPrint) contribute additional prognostic value beyond conventional clinical indicators [27,34–36]. Thus, treatment recommendations for adjuvant CHT should include GEP as an integral part of any decision algorithm [20,37].

It is important for clinicians to be aware of how these multigene predictive assays are increasingly being used in practice [38] and their limitations. In addition to factors such as risk of distant recurrence that affect treatment, clinicians should also consider potential risks as well as patient’s quality of life. As this field and what constitutes standard of care appears to be continually evolving, it is equally important for clinicians to avail themselves of the opportunities to learn about the latest clinical practice guidelines and accepted standards in order to avoid any potential litigation [39], even though it may be premature to consider litigation at this time.

Since the standard of care for multigene predictors is still fraught with uncertainty, this presents with a ‘damned if you do, damned if you don’t’ dilemma for physicians [39]. A physician who fails to utilize such a test and recommends against CHT on other grounds may face a lawsuit by a patient with recurring breast cancer who argues after the fact that the physician was negligent in not relying on a multigene predictor test. Alternatively, if the physician recommends a multigene predictor assay and no CHT based on the results of that test, the physician may face a lawsuit alleging premature reliance on a test that was not yet subsequently validated [39]. In either case, with the advantage of perfect hindsight, the physician could be portrayed by a skilled trial lawyer as having acted too slowly or too quickly to rely on multigene predictor assays.

While uncertainty and rapid change in medical technologies create an inherent liability risk for physicians, the best way to protect against such liability is for the physician to document in the patient’s medical record at the time of decision that the physician was aware of the available and uncertain evidence on the clinical utility of GEP assays and made an informed professional judgment in the face of such uncertainty.

We recognize that policy regarding the use of multigene predictors should ideally be governed by evidence from prospective randomized clinical trials [23,40]. Until the results of prospective clinical trials such as the TAILOR-X [41] and the MIND- ACT [42] become available, systematic assessments and meta-analyses, such as this study, ought to be useful for providing insight to clinicians and policymakers and for better understanding potential liability in this burgeoning field.

### Expert commentary
 Multigene predictors, including GEP, are being used increasingly for risk stratification to identify early-stage breast cancer patients who are most likely to benefit from adjuvant CHT. During the past decade, a number of studies have been published that focused on the analytical and clinical validity, and the clinical utility of at least four multigene predictors, of which the most frequently used in the US populations is the 21-gene signature Oncotype DX. Although incorporation of multigene predictor assays into routine clinical breast cancer practice has not yet been fully realized, GEP for prediction and prognosis is increasingly being used in clinical settings, with a consequent interest on the part of clinicians and policymakers toward developing and managing healthcare systems that incorporate personalized genomic medicine as a standard of medical practice. This has the potential to affect future malpractice litigation risks. This study suggests that multigene predictor assays, particularly use of the 21-gene assay, are increasingly being used in practice. It appears that this trend
toward adoption of multigene predictors and other molecular genomic technologies will continue as more advances in technologies for genomic profiling are made. As what constitutes standard of care continues to evolve, it is important for physicians to document in the patient’s medical record at the time of decision that the physician was aware of the available and uncertain evidence on the clinical utility of multigene predictor assays and made an informed professional judgment in the face of such uncertainty.

Five-year view

We believe that as multigene predictors and other genomic technologies continue to evolve, it is likely that the potential for malpractice litigation will increase as well. Historically, there are precedents for increased litigation with increased use and adoption of novel technologies, primarily due to increased patient expectations. Thus, we can expect that in the next 5–10 years, as genomic technologies such as multigene predictor assays become more routinely integrated into clinical practice and health systems, physicians will likely be dealing with the first wave of malpractice litigation.

Key issues

- Multigene predictors, including gene expression profiling, are being used increasingly for risk stratification to identify early-stage breast cancer patients who are most likely to benefit from adjuvant chemotherapy.
- One consequence of the increased use of gene expression profiling and other molecular companion diagnostics, and the heightened efforts toward their incorporation into routine clinical practice, is the potential for future malpractice litigation.
- A systematic search of the literature from January 2002 to March 2014 and meta-analyses were performed.
- The analysis of Cox hazard ratios demonstrated that gene signatures appear to be associated with predicting distant recurrence.
- Point estimates and their 95% CIs showed some degree of overlap between the four gene signatures assessed; however, the 21-gene signature, Oncotype DX®, showed the highest stability in the estimation of likelihood of distant risk of recurrence.
- Using recurrence scores resulted in changes in treatment recommendations from initial treatment recommendations in 31.8% of all patients in the studies.
- The results of this study may provide insight about the use of multigene predictors in clinical practice for prediction and prognosis in breast cancer, and improved understanding of potential liability issues in the field of personalized genomic medicine.

References

42. Viale G, Slaets L, Bogaerts J, et al. High concordance of protein (by IHC), gene (by FISH; HER2 only), and microarray readout (by TargetPrint) of ER, PgR, and HER2: results from the EORTC 10041/BIG 03-04 MINDACT trial. Ann Oncol 2014;25(4):816-23


74. Fried G, Moskovitz M. Treatment decisions in estrogen receptor-positive early breast cancer patients with intermediate oncopotype DX recurrence score results. Springerplus 2014;3:71