Methodological quality assessment of studies investigating the MammaPrint test

Study feature	Qualities sought	Bueno-de-Mesquita *et al.* (2009)[93]	Gevensleben *et al.* (2010)[94]	Ishitobi *et al.* (2010)[96]	Kok *et al.* (2010)[99]	Kunz *et al.* (2011)[92]	Mook *et al.* (2010)[75]	Na *et al.* (2011)[100]
Sample of patients	Inclusion criteria defined	Y	Y	Y	Y	Y	Y	Y
	Sample selection explained	Y	Y	Y	Y	Y	Y	Y
	Adequate description of diagnostic criteria	Y	Y	?	Y	Y	Y	Y
	Clinical and demographic characteristics fully described	Y	Y	Y	Y	Y	Y	Y
	Representative (selected by random selection or as consecutive cases)	Y (consecutive)	Y (consecutive)	U	U	U	Y (consecutive)	U
	Assembled at a common (usually early) point in the course of their disease	Y (pT1–2, N0)	?	Y (all N0)	?	Y (T1-T3, N0–3)	Y (T1–2, N0)	Y (T1–2, N0, M0)
	Complete (all eligible patients were included)	U	N	N	U	N	N	N
Follow-up of patients	Sufficiently long	Y (median 5.8 years)	U	Y (median 7.1 years)	Y (9.6–11.1 years)	U	Y (median 11.6 years)	U
Outcome	Objective	Y	Y	Y	Y	Y	Y	Y
	Unbiased (e.g. assessment blinded to prognostic information)	Y	?	?	?	?	Y	?
	Fully defined	Y	? (risk classification)	N	Y	? (risk classification)	Y	? (risk classification)
	Appropriate	Y (OS, DMFS)	? (risk classification)	Y (DMFS, risk classification)	Y (BCSS)	? (risk classification)	Y (DMFS, BCSS, risk classification)	? (risk classification)
	Known for all or a high proportion of patients	Y	Y	Y	Y	Y	Y	Y

Study feature	Qualities sought	Bueno-de-Mesquita *et al.* (2009)[93]	Gevensleben *et al.* (2010)[94]	Ishitobi *et al.* (2010)[96]	Kok *et al.* (2010)[99]	Kunz *et al.* (2011)[92]	Mook *et al.* (2010)[75]	Na *et al.* (2011)[100]
Prognostic variable	Fully defined, including details of method of measurement if relevant	Y (MammaPrint, St Gallen, NPI, Adjuvant! Online)	Y (MammaPrint, St Gallen, Adjuvant! Online)	Y (MammaPrint, St Gallen)	Y (MammaPrint, endocrine response category)	Y (MammaPrint, St Gallen, Adjuvant! Online)	Y (MammaPrint, Adjuvant! Online)	Y (MammaPrint, St Gallen, NIH, Adjuvant! Online)
	Precisely measured	Y	Y	Y	Y	?	?	Y
	Available for all or a high proportion of patients	Y	Y	Y	Y	Y	Y	Y
	If relevant, cut-point(s) defined and justified	Y (reference provided)	Y (reference provided)	Y	Y (reference provided)	N	Y (reference provided)	Y
Analysis	Continuous predictor variable analysed appropriately	Y	N	N	Y	?	Y	N
	Statistical adjustment for all important prognostic factors	Y	N	N	Y	?	Y	N
Intervention subsequent to inclusion in cohort	Fully described	N	N	N	N	?	N	N
	Intervention standardised or randomised	N	N (96% had chemotherapy/ endocrine therapy; no further details provided)	N (retrospective study: 73% had adjuvant hormonal therapy, 28% adjuvant chemotherapy; no further details provided)	N [retrospective study: 100% tamoxifen treated (no further details provided) and no neoadjuvant therapy; consecutive series: 100% tamoxifen naive]	? (prospective study)	N [retrospective study: 18% had adjuvant endocrine (tamoxifen) therapy; no further details provided]	N (retrospective study: 73% had chemotherapy; no further details provided)

N, no; U, unclear/not reported; Y, yes.

Summary of results: MammaPrint test (new data)

<table>
<tr><th>Study</th><th>Outcomes/end points</th><th>Results</th><th>Authors' conclusions</th><th>Comments</th></tr>
<tr>
<td>Bueno-de-Mesquita et al. (2009)[93]</td>
<td>Time from surgery to distant metastasis as first event (counted as failures)
OS (defined as time from surgery to death)</td>
<td>
Classification and disease outcome
Univariate analysis
OS at 5 years
<table>
<tr><th>Method</th><th>HR</th><th>95% CI</th><th>p-value</th></tr>
<tr><td>MammaPrint</td><td>3.4</td><td>1.2 to 9.6</td><td>0.021[a]</td></tr>
<tr><td>Adjuvant! Online</td><td>2.5</td><td>0.59 to 11</td><td>0.22</td></tr>
<tr><td>NPI</td><td>2.8</td><td>0.99 to 7.8</td><td>0.053</td></tr>
<tr><td>CBO[102]</td><td>2.3</td><td>0.84 to 6.6</td><td>0.11</td></tr>
<tr><td>St Gallen</td><td>3.0</td><td>0.4 to 22</td><td>0.29</td></tr>
</table>
a The probability of OS (as first event) was 97% (±SE 2%) for good and 82% (±SE 5%) for poor prognosis signature patients (p-value not reported).

Distant metastasis (as first event) at 5 years
<table>
<tr><th>Method</th><th>HR</th><th>95% CI</th><th>p-value</th></tr>
<tr><td>MammaPrint</td><td>5.7</td><td>1.6 to 20</td><td><0.007[a]</td></tr>
<tr><td>Adjuvant! Online</td><td>4.6</td><td>0.61 to 35</td><td>0.14</td></tr>
<tr><td>NPI</td><td>2.2</td><td>0.78 to 6.5</td><td>0.14</td></tr>
<tr><td>CBO[102]</td><td>1.8</td><td>0.64 to 5.3</td><td>0.26</td></tr>
<tr><td>St Gallen</td><td>2.5</td><td>0.34 to 19</td><td>0.37</td></tr>
</table>
a The probability of remaining free of distant metastasis (as first event) was 98% (±SE 2%) for good and 78% (±SE 6%) for poor prognosis signature patients (p-value not reported).
</td>
<td>The MammaPrint test is also an independent prognostic factor in node-negative breast cancer patients for women diagnosed in recent years</td>
<td>Data on NPI, OS, distant metastasis as first event. Additional data on updated follow-up results (median 10.2 years) reported by van de Vijver et al.[64]</td>
</tr>
</table>

Study	Outcomes/end points	Results	Authors' conclusions	Comments

Multivariate analysis with clinical risk indicators

OS, adjusted for performance of MammaPrint test (poor/high vs. good/low)

Method	HR	95% CI	*p*-value
Adjuvant! Online	3.0	1.0 to 8.9	0.04
NPI	2.7	0.87 to 8.1	0.09
CBO[102]	2.9	0.98 to 8.6	0.06
St Gallen	3.1	1.0 to 9.2	0.04

Distant metastasis (as first event), adjusted for performance of MammaPrint test (poor/high vs. good/low)

Method	HR	95% CI	*p*-value
Adjuvant! Online	4.8	1.3 to 17	0.018
NPI	5.4	1.4 to 21	0.015
CBO[102]	5.6	0.98 to 8.6	0.011
St Gallen	5.8	1.5 to 22	0.010

Gene expression profiling comparison between MammaPrint test and risk assessment based on (A) Adjuvant Online!, (B) St Gallen guidelines, (C) NPI and (D) CBO guidelines 2004[105]

A.

Clinical risk (Adjuvant! Online)	MammaPrint, *n* (%)			Discordant finding
	Good	Poor	Total	
Low	23 (19)	6 (5)	29 (24)	38% (47/123), 95% CI 30% to 47%, kappa: 0.252
High	41 (33)	53 (43)	94 (76)	
Total	64 (52)	59 (48)	123 (100)	

B.

Clinical risk (St Gallen guidelines)	MammaPrint, *n* (%)			Discordant finding
	Good	Poor	Total	
Low	15 (12)	1 (1)	16 (13)	41% (50/123), 95% CI 32% to 49%, kappa: 0.211
Intermediate/high	49 (40)	58 (47)	107 (87)	
Total	64 (52)	59 (48)	123 (100)	

Study	Outcomes/end points	Results	Authors' conclusions	Comments

C.

Clinical risk (NPI)	MammaPrint, *n* (%)			Discordant finding
	Good	Poor	Total	
Low	46 (37)	14 (11)	60 (49)	26% (32/123), 95% CI 18% to 34%, kappa: 0.480
Moderate/high	18 (15)	45 (37)	63 (51)	
Total	64 (52)	59 (48)	123 (100)	

D.

Clinical risk (CBO guidelines[102])	MammaPrint, *n* (%)			Discordant finding
	Good	Poor	Total	
Low	41 (33)	14 (11)	55 (45)	30% (37/123), 95% CI 22% to 38%, kappa: 0.401
High	23 (19)	45 (37)	68 (55)	
Total	64 (52)	59 (48)	123 (100)	

Study	Outcomes/end points	Results	Authors' conclusions	Comments
Gevensleben *et al.* (2010)[94]	Comparison of risk prediction by the MammaPrint test with that using St Gallen criteria and Adjuvant! Online	**Gene expression profiling compared with current risk classifications**	The MammaPrint test provides improved prediction of recurrence risk compared with currently used guidelines	

	MammaPrint ($n=140$)		
	Low risk ($n=78$)	High risk ($n=62$)	Discordant finding
St Gallen criteria[a]			
Low ($n=7$)	6[b]	1	–
Intermediate ($n=123$)	70	53	
High ($n=10$)	2	8[c]	
Adjuvant! Online[c]			
Low ($n=45$)	33[b]	12	57 (41%)
High ($n=95$)	45	50[b]	

a St Gallen risk classification according to Goldhirsch *et al.*[95]
b These values were summed to obtain concordant findings.
c Adjuvant! Online risk classification according to Ravdin *et al.*[107]

Treatment advice

For 59/62 patients with a poor prognosis signature identified by the MammaPrint test, the clinical treatment was recorded. In total, 19 (32%) of these patients did not receive adjuvant systemic treatment other than endocrine therapy and were potentially undertreated. In contrast, 35/77 patients who were classified as having a good prognosis by the MammaPrint test, and for whom treatment was known, received chemotherapy and were potentially overtreated. As a result, the MammaPrint test would have resulted in altered treatment advice for 40% of patients

Study	Outcomes/end points	Results	Authors' conclusions	Comments
Ishitobi *et al.* (2010)[96]	DMFS (not defined) Correlation between the MammaPrint test risk category and clinicopathological parameters (St Gallen criteria)	**Classification and disease outcome** *DMFS* At 5 years (probability): G1 (low risk): 100% vs. G2 (high risk): 94%, HR not reported **Risk classification and distant metastasis** Among the 102 patients, 20 were classified as low risk and 82 as high risk. Based on the 1998 St Gallen criteria[103] all patients were classified as intermediate or high risk. The 2009 St Gallen criteria[97] use more refined criteria to define the low-risk group and classify seven patients as low risk. This is still lower (7/100) than the number identified by the MammaPrint test (20/102) ($p=0.009$). See table for further details	The MammaPrint test accurately identified Japanese breast cancer patients at low risk of developing recurrences. In fact, 100% of the individuals in the low-risk group remained metastasis free for the duration of the observation period	

Group	Patients ($n=102$)	DM	DF	*p*-value	Predictive value (%)	
					PPV	NPV
MammaPrint						
High risk	82	8	74	0.143	9.8	100
Low risk	20	0	20			
St Gallen 1998[103] version[a]						
Intermediate or high risk	100	8	92	NA	8.0	100
Low risk	0	0	0			
St Gallen 2009[97] version[a]						
Intermediate or high risk	93	8	85	0.419	8.6	100
Low risk	7	0	7			

DF, disease free; DM, distant metastasis; NA, not available.
a Criteria with missing data.

Study	Outcomes/end points	Results	Authors' conclusions	Comments
Kok *et al.* (2010)[99]	BCSS (defined as time from surgery to breast cancer-related death)	**Classification and disease outcome** BCSS in patients treated with adjuvant tamoxifen according to the MammaPrint test At 5 years: G1 (low risk): 96.2% (±SE 2.2%) vs. G2 (high risk): 72.5% (±SE 7.4%); p=NR; univariate HR: NR At 10 years: G1 (low risk): 80.6% (±SE 4.7%) vs. G2 (high risk): 63.4% (±SE 8.2%); p=NR; univariate HR: 2.78 (95% CI 1.30 to 5.94; p=0.008) BCSS in patients treated with adjuvant tamoxifen according to endocrine response categories (St Gallen consensus: highly endocrine responsive: ER and PR ≥50%; incompletely endocrine responsive: ER and/or PR low or with either one absent) At 5 years: G1 (high response): 98.0% (±SE 2.0%) vs. G2 (incomplete response): 82.1% (±SE 4.7%); p=NR; univariate HR: NR At 10 years: G1 (high response): 93.1% (±SE 3.9%) vs. G2 (incomplete response): 61.3% (±SE 6.3%); p= NR; univariate HR: 7.22 (95% CI 2.17 to 24.00; p=0.001) BCSS in patients treated with no adjuvant systemic treatment according to MammaPrint test At 5 years: G1 (low risk): 97.6% (±SE 1.6%) vs. G2 (high risk): 80.9% (±SE 5.0%); p=NR; univariate HR: NR At 10 years: G1 (low risk): 90.2% (±SE 3.3%) vs. G2 (high risk): 63.3% (±SE 6.3%); p=NR; univariate HR: 4.52 (95% CI 2.01 to 10.2; p<0.001) BCSS in patients treated with no adjuvant systemic treatment according to endocrine response categories At 5 years: G1 (high response): 92.9% (±SE 2.6%) vs. G2 (incomplete response): 85.7% (±SE 5.0%); p=NR; univariate HR: NR At 10 years: G1 (high response): 82.0% (±SE 4.0%) vs. G2 (incomplete response): 72.6% (±SE 6.5%); p=NR; univariate HR: 1.78 (95% CI 0.86 to 3.66; p=0.118)		Data also available for combined use of the MammaPrint test and endocrine response categories but not extracted

Study	Outcomes/end points	Results	Authors' conclusions	Comments
Kunz *et al.* (2011)[92]	Comparison of risk prediction by the MammaPrint test with that by the St Gallen guidelines 2007/9 and Adjuvant! Online	**GEP compared with current risk classifications** Method / Low risk (*n*) / Intermediate risk (*n*) / High risk (*n*) MammaPrint: 29 / – / 15 St Gallen criteria[a]: 4 / 34[b] / 6 Adjuvant! Online[c]: 19 / – / 25 a St Gallen risk classification according to Goldhirsch *et al.*[97] b In the group of women with intermediate risk, the MammaPrint test assigned 23 patients to low risk and 11 to high risk. c Patients were classified as having low clinical risk when the 10-year OS rate as predicted by Adjuvant! Online was >88% for ER+ tumours and >92% for ER– tumours. **Risk assessment by the MammaPrint test according to nodal status in women with early breast cancer** Low risk (*n*) / High risk (*n*) LN+ disease: 13 / 5 LN– disease: 19 / 9 **Comparison of the clinicopathological features with those of previous validation studies** Data reported but not extracted	Using gene expression analysis as an additional tool, patients with an intermediate clinical risk can be accurately separated into low- and high-risk groups. The gene expression analysis provides more accurate information on recurrence risk than conventional clinicopathological criteria	

<table>
<tr><th>Study</th><th>Outcomes/end points</th><th>Results</th><th>Authors' conclusions</th><th>Comments</th></tr>
<tr><td>Mook et al. (2010)[75]</td><td>DMFS (defined as time from surgery to distant metastasis as first event; counted as failures)
BCSS (defined as time from surgery to breast cancer-related death)
Comparison of risk prediction by the MammaPrint test with that by Adjuvant! Online</td><td>Classification and disease outcome
DMFS
At 5 years: G1 (low risk): 93% (±SE 3%) vs. G2 (high risk): 72% (±SE 6%); $p=0.07$; univariate HR: 4.6 (95% CI 1.8 to 12.0; $p=0.001$)
At 10 years: G1: 80% (±SE 5%) vs. G2: 67% (±SE 7%); $p=$NR; univariate HR: NR
Over entire follow-up period: univariate HR: 1.8 (95% CI 0.9 to 3.5; $p=0.07$)
BCSS
At 5 years: G1: 99% (±SE 1%) vs. G2: 80% (±SE 5%); $p=0.036$; univariate HR: 19.1 (95% CI 2.5 to 148; $p=0.005$)
At 10 years: G1: 90% (±SE 4%) vs. G2: 69% (±SE 6%); $p=$NR; univariate HR: NR
Over entire follow-up period: univariate HR: 2.0 (95% CI 1.0 to 4.0; $p=0.04$)
Prediction of early BCSD
BCSS
At 5 years: adjusted HR: 14.4 (95% CI 1.7 to 122.2; $p=0.01$)
At 10 years: adjusted HR: 4.4 (95% CI 1.4 to 13.6; $p=0.01$)
Subgroup analyses: BCSS in hormonal therapy-naive patients (untreated)
At 5 years: adjusted HR: 10.8 (95% CI 1.2 to 94.7; $p=0.03$)
GEP compared with current risk classifications
<table>
<tr><th></th><th colspan="2">MammaPrint test ($n=148$) (n)</th><th></th></tr>
<tr><th>Adjuvant! Online[a]</th><th>Low risk ($n=91$)</th><th>High risk ($n=57$)</th><th>Discordant finding</th></tr>
<tr><td>Low risk ($n=74$)</td><td>62[b]</td><td>12</td><td>41 (28%)</td></tr>
<tr><td>High risk ($n=74$)</td><td>29</td><td>45[b]</td><td></td></tr>
</table>
a Patients were classified as having low clinical risk when the 10-year OS rate as predicted by Adjuvant! Online was >88% for ER+ tumours and >92% for ER– tumours.
b These values were summed to obtain concordant findings.</td><td>The MammaPrint test can accurately select postmenopausal patients at low risk of breast cancer-related death within 5 years of diagnosis and can be of clinical use in selecting postmenopausal women for adjuvant chemotherapy</td><td>Data on distant metastasis as first event</td></tr>
</table>

Study	Outcomes/end points	Results	Authors' conclusions	Comments
Na *et al.* (2011)[100]	Comparison of risk prediction by the MammaPrint test with that by the St Gallen criteria, NIH guidelines[101] and Adjuvant! Online	**Gene expression profiling compared with current risk classifications** The MammaPrint test identified five patients with a low-risk prognosis signature and 31 patients with a high-risk prognosis signature. Clinical risk was concordant with the prognosis signature for 29 patients according to the St Gallen guidelines; 30 patients according to the NIH guidelines[101] and 23 patients according to Adjuvant! Online	The results of the MammaPrint test for Korean patients with breast cancer were somewhat different from those identified in Europe. This difference should be studied to determine whether or not there is a gene disparity between Asians and Europeans. Further large-scale studies with a follow-up evaluation are required to assess whether or not the use of the MammaPrint test can predict the prognosis of Korean patients with breast cancer	

	MammaPrint test ($n=36$) (n)		
	Low risk ($n=5$)	High risk ($n=31$)	Discordant finding
St Gallen criteria[a]			
Low risk ($n=6$)	2[b]	4	7 (19%)
High risk ($n=30$)	3	27[b]	
[c]NIH guidelines[101]			
Low risk ($n=5$)	2[b]	3	6 (17%)
High risk ($n=31$)	3	28[b]	
Adjuvant! Online[d]			
Low risk ($n=14$)	3[b]	11	13 (36%)
High risk ($n=22$)	2	20[b]	

a St Gallen risk classification guideline according to Goldhirsch *et al.*:[95] a low clinical risk was defined as possessing all of the following criteria: ER+ and/or PR+ status, tumour size ≤2 cm, histological grade I and age ≥35.
b These values were summed to obtain concordant findings.
c Low risk for the LN– group was defined as a tumour size <1 cm and a favourable histological subtype such as tubular and mucinous cancer.
d Patients were classified as having low clinical risk when the 10-year OS rate as predicted by Adjuvant! Online was >88% for ER+ tumours and >92% for ER– tumours.

Comparison of the clinicopathological features with those of previous validation studies
Data reported but not extracted

NR, not reported.