An Update on the Molecular Classification of Breast Cancer

Emad Rakha
MB BCh, MSc, MD, PhD, FRCPath
Professor and Consultant Pathologist
School of Medicine, University of Nottingham
Why do we need to classify breast?

* BC is a heterogeneous disease. Tumour with similar morphology show variable behaviour, outcome and response to therapy

* Stratification of BC into classes will help in
  
  Â Identifying patients whose prognosis is so good that adjuvant therapy would not be cost beneficial
  
  Â Identifying patients whose prognosis is so poor that a more aggressive adjuvant approach would be warranted
  
  Â Identifying patients likely to be responsive or resistant to particular forms of therapy
Presentation / Stage

* Metastatic BC (<10%):- systemic therapy based on predictive markers (ER and HER2)

* Locally advanced BC (<10%):- local control + systemic therapy based on predictive markers

* Early staged BC (>80%):- local control +/- systemic therapy based on prognostic and predictive markers

ER+ BC = Hormone therapy

HER2+ BC = Herceptin based on prognostic factors
Selecting Patients for Chemotherapy

- **Stage**
- **Biology**

Risk:
- Grade + receptors
- NPI / Adjuvant!online/Predict

Chemotherapy benefit estimate from proportional hazards eg; survival benefits after 10 years

* No single factor useful.
* Should be used in combination
* Complementing clinicopathological with molecular variables
Prognosis

- High risk: Chemotherapy
- Low risk: No chemotherapy

However, clinically indeterminate groups such as LN-/ER+/HER2- tumours: Additional prognostic tests are needed (Multigene Prognostic Assays)
Prognostic multigene signatures
Multigene signatures

Microarray and RT-PCR based assays

- 21 gene signature (Oncotype Dx)
- 70 gene signature (MammaPrint)
- 76 gene signature (Rotterdam)
- 50 genes: Risk of Recurrence (ROR) score (Prosigna)
- 12 genes (Endopredict) & Epclin
- 5 genes (Molecular grade index)
- 2 gene ratio (H/I™)
- 97 gene: Genomic grade index (MapQuant Dx)
- 14 genes (BreastOncPx)
- 14 gene signature (Celera Metastasis Score™)
Multigene signatures

IHC and and ISH based assays

- 4 gene signature (IHC4; ER, PR, HER2 and Ki67)
- 5 gene signature (Mammostrat)
- 9 gene signature (Mammostrat Plus; 5 + ER, PR, HER2 and Ki67)
- 5 gene signature (ProEx™ Br)
- 3 gene signature (eXagenBC™)

Signatures based on a biological process

- Wound-response signature (442 genes)
- Immune signatures (14 genes)
- Invasiveness Gene Signature (186 genes)
Recent ASCO guideline recommendation

In addition to ER, PR and HER2, there is sufficient evidence of clinical utility for the biomarker assays [Oncotype DX, EndoPredict, PAM50, Breast Cancer Index, and urokinase plasminogen activator and plasminogen activator inhibitor type 1 in HR+/HER2- LN- groups and can be used.

These assays should not be used to guide treatment decision in LN+, HER2+ or triple negative cancer.

No other molecular test (including ki67) should be used to direct treatment decision.
**Onco-type DX™ 21-Gene Recurrence Score (RS) Assay**

16 Cancer and 5 Reference Genes (RT-PCR) From a pool of 250 genes

**PROLIFERATION**
- Ki-67
- STK15
- Survivin
- Cyclin B1
- MYBL2

**ESTROGEN**
- ER
- PR
- Bcl2
- SCUBE2

**INVASION**
- Stromelysin 3
- Cathepsin L2

**HER2**
- GRB7
- HER2

**CD68**

**GSTM1**

**BAG1**

**REFERENCE**
- Beta-actin
- GAPDH
- RPLPO
- GUS
- TFRC

**RS =**

+ 1.04 x Proliferation Group Score
+ 0.47 x HER2 Group Score
- 0.34 x ER Group Score
+ 0.10 x Invasion Group Score
+ 0.05 x CD68
- 0.08 x GSTM1
- 0.07 x BAG1

**Category** | **RS (0-100)**
--- | ---
Low risk | RS <18 (50%; 7%)
Int risk | RS ≥18 - <31 (14%)
High risk | RS ≥31 (27%, 30%)

**Onco type DX™ Clinical Validation: RS as Continuous Predictor**

My RS is 35. What is the chance of recurrence within 10 years?

- **Low-Risk Group**: No chemotherapy
- **Intermediate-Risk Group**: My RS is 35. What is the chance of recurrence within 10 years? 25%
- **High-Risk Group**: Chemotherapy

95% CI
Main point: The Oncotype DX® Recurrence Score® is correlated with distant recurrence rate at 10 years, hormone therapy benefit, and chemotherapy benefit. There is a continuous biology of breast cancer that is revealed by the Recurrence Score. There are clear underlying phenotypes that correspond to Low Recurrence Score Disease and High Recurrence Score Disease.

Distant recurrence rate at 10 years with 5 years of tamoxifen treatment: the higher the score, the higher the risk of distant recurrence.
Hormone therapy benefit: the lower the score, the greater the impact of tamoxifen given for 5 years on proportion of patients recurrence-free at 10 years.
Chemotherapy benefit: the higher the score, the greater the proportion of patients distant recurrence-free at 10 years.
The precision of the Recurrence Score with a tight confidence interval in the lower Recurrence Score ranges, allows oncologists to confidently identify low Recurrence Score disease. Since a low Recurrence Score value is associated with low risk of distant recurrence and low benefit from adding chemotherapy to tamoxifen, many oncologists will choose to forego chemotherapy. Conversely, a high Recurrence Score value is associated with a high risk of distant recurrence and significant benefit from adding chemotherapy to tamoxifen, prompting many oncologists to add chemotherapy. Approximately 25% of patients will fall into the intermediate Recurrence Score range. Although the continuous nature of the assay results allows oncologists to ascertain distant recurrence risk for these patients, there are multiple ways to apply the information. Patients with a “high” intermediate Recurrence Score value may be considered differently from those with a “low” intermediate Recurrence Score value. Nonetheless, the information provided by the Recurrence Score result can be very helpful in making treatment decisions when viewed in the context of other patient-specific factors, including patient concern, age, tumor size and grade, and conditions that may increase the risk of chemotherapy-associated toxicity.


Rakha Emad, 03/01/2016
EndoPredict

It uses FFPE tissue and RT-PCR

It measures expression of 8 genes plus 4 control:
- 3 proliferation-associated genes (BIRC5, UBE2C, DHCR7),
- 5 HR-associated genes (RBBP8, IL6ST, AZGP1, MGP, STC2), plus
- 3 normalisation genes (CALM2, OAZ1 and RPL37A) and 1 DNA reference gene (HBB)

The EP score is calculated to give a scale from 0 to 15. EP scores of <5 are designated as 'low risk' and EP scores of 5 or more as 'high risk'
EndoPredict

Å The EPclin score is calculated by adding clinical data about tumour size and nodal status to the EP score

Å EPclin score predicts the likelihood of 10-year distant metastasis in women with ER+ HER2- BC

Å Is used to identify tumour types that will not benefit from chemotherapy

Å One EndoPredict test costs £1000 if performed in a local laboratory or £1500 if it is sent to the distributor's laboratory compared to >$3,000 for Oncotype DX
PAM50

- Formalin-fixed, paraffin-embedded tissue and RT-PCR or nanotechnology-based nCounter digital gene expression platform

- mRNA expression of 50 genes + 5 housekeeping control genes

- Categorization of tumors into the four intrinsic subtypes with prediction of outcome (lum A, lum B, HER2-enriched, and basal-like) with high classification agreement with microarray “intrinsic” subtyping

- ROR (Prosigna) for risk stratification
IHC4 score

• Semiquantitative expression of ER, PR, HER2 and Ki67

• Retrospective analysis of TransATAC Trial revealed that the 4 markers when combined can produce score with high prognostic relevance (IHC4)

• IHC4+C= 4 markers+ grade, size, LN, age and endocrine therapy

IHC4 Comparison with Oncotype DX

Cuzick et al JCO 2011(32):4273-8
Several comparison studies and prospective clinical trials have been carried out and some are ongoing to determine their clinical utility.

**Oncotype DX:** TAILORx (Trial Assigning Individualized Options for Treatment (Rx), and RxPONDER (Rx for Positive Node (N1), Endocrine Responsive Breast Cancer)

**MammaPrint:** MINDACT (Microarray In Node-negative and 1-3 Node-Positive Disease May Avoid ChemoTherapy) that accrued 6,600 patients between 2006 and 2011.

**In the UK**

**OPTIMA:** Optimal Personalised Treatment of early breast cancer using Multi-parameter Analysis (A trial of a diagnostic test in ER+ HER2- LN+ BC)
These are the 3 on-going randomised trials of multi-parameter assays in ER+ve breast cancer:

**TAILORx**

- Randomised women with node--ve disease with a RS of 11-25 to chemotherapy or not in addition to endocrine therapy and is expected to report in 2015. It should provide prospective validation for the use of Oncotype as a predictive tool for selecting chemotherapy and additional information on the test threshold.

**MINDACT**

- Must surely be one of the most complex studies ever conducted in breast cancer. It used the MammaPrint assay, performed on fresh tissue but now also available for use on formalin-fixed paraffin-embedded tumour blocks. Its primary objective is to demonstrate that MammaPrint is superior to conventional risk assessment with "Adjuvant! Online" in identifying women who are likely to benefit from chemotherapy. MINDACT includes women with 1-3 involved nodes and should report at about the same time as TAILORx.

**RxPONDER**

- Is similar in design to TAILORx but is for women with 1-3 involved nodes and allows randomisation of all patients with RS ≤25. It is currently recruiting.
## Multiparameter assays evaluated in OPTIMA prelim

<table>
<thead>
<tr>
<th>Test</th>
<th>Prosigna (ROR_PT)</th>
<th>MammaPrint</th>
<th>IHC4</th>
<th>IHC4 - AQUA</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Low risk</td>
<td>High risk</td>
<td>Low risk</td>
<td>High risk</td>
</tr>
<tr>
<td>Oncotype DX</td>
<td>N=298</td>
<td>N=297</td>
<td>N=257</td>
<td>N=270</td>
</tr>
<tr>
<td>Low risk</td>
<td>63</td>
<td>19</td>
<td>60</td>
<td>23</td>
</tr>
<tr>
<td>High risk</td>
<td>3</td>
<td>15</td>
<td>2</td>
<td>15</td>
</tr>
<tr>
<td>Prosigna ROR_PT</td>
<td>N=295</td>
<td>N=254</td>
<td>N=268</td>
<td></td>
</tr>
<tr>
<td>Low risk</td>
<td>53</td>
<td>9</td>
<td>54</td>
<td>9</td>
</tr>
<tr>
<td>High risk</td>
<td>13</td>
<td>25</td>
<td>18</td>
<td>19</td>
</tr>
<tr>
<td>MammaPrint</td>
<td></td>
<td>N=254</td>
<td>N=269</td>
<td></td>
</tr>
<tr>
<td>Low risk</td>
<td></td>
<td>51</td>
<td>9</td>
<td></td>
</tr>
<tr>
<td>High risk</td>
<td></td>
<td>22</td>
<td>18</td>
<td></td>
</tr>
<tr>
<td>IHC4</td>
<td></td>
<td></td>
<td>N=244</td>
<td></td>
</tr>
<tr>
<td>Low risk</td>
<td></td>
<td>57</td>
<td>14</td>
<td></td>
</tr>
<tr>
<td>High risk</td>
<td></td>
<td>4</td>
<td>25</td>
<td></td>
</tr>
</tbody>
</table>
OPTIMA prelim results

• Disagreement between multi-parameter tests at an individual patient level.

• OPTIMA prelim unable to identify a “best” test.

• Expected cost-effectiveness highly dependent on model assumptions
Molecular taxonomy (intrinsic subtypes)

Gene expression

Expanded IHC

RT-PCR

*karyotype, LOH, aCGH……

*Next Generation sequencing
Microarray-based gene expression analysis

Perou et al In 2000
>1700 genes

Each row is a gene
Each column is a sample
Green: <median
Black: =median
Red: >median

Rt panel: cell lines
Left panel: tissue

Dendrogram: similarities in the expression patterns

Molecular portraits of human breast tumours

Charles M. Perou*, Therese Sorlie†, Michael B. Eisen*,
Molecular/intrinsic subtypes

• Subsequent studies: several molecular subclasses of clinical relevance

• Luminal class: classified into at least two (Lum A and Lum B) based on proliferation and HER2 status

• Basal class was classified into several molecular subgroups including
  - Claudin-low,
  - Molecular apocrine subtype,
  - Interferon-enriched,
  - Basal-like I and basal-like II,
  - Mesenchymal stem-like and
  - Immunomodulatory subtypes
Molecular Subtypes and Prognosis

"Intrinsic" gene set on 78 single tumor samples

Basal-like = 4  ERBB2+ = 5  Normal Breast  Luminal Subtype B = 2  Luminal Subtype A = 1

SURVIVAL

DISEASE-FREE SURVIVAL

sensitive analysis was done on the set of 51 doxorubicin treated patients only

Sorlie T et al, PNAS 2001
Clinicopathologic surrogate definition

Luminal A-like
- ER+, HER2-, Ki67 low, PgR high
- Low-risk molecular signature (if available)

Luminal B-like
- HER2-negative:
  - ER+, HER2- and either Ki67 high or PgR low
  - High-risk molecular signature (if available)
- HER2-positive:
  - ER-positive, HER2-positive, any Ki67, any PgR

HER2-positive (non-luminal):
- HER2+, ER and PgR absent

Basal-like/Triple-negative
- ER and PgR absent, HER2-negative

Annals of Oncology 26: v8–v30, 2015
The integrative transcriptomic and genomic subgroups have distinct clinical outcomes

**METABRIC: 2012**

**2000 BC**

Genomics driven classification of BC based on an integrative analysis of gene expression (>34k) and genome-wide copy number alterations

10 molecular classes / Integrative clusters

The genomic and transcriptomic architecture of 2,000 breast tumours reveals novel subgroups

Christina Curti1,2,3, Schrøer P, Shah1,2,4, Sun, Fei, Ding Chiu1,2,4, Gulisa Turashvili1,2,4, Oscar M, Rueda1,2,3, Mark J, Dunning2.
Highly significant association with patient outcome

Log Rank \( p < 0.0001 \)

Survival in months

Log Rank \( p < 0.0001 \)
NPI+

Underpinning Hypothesis

The performance of prognostic factors may differ in the different biological classes of breast cancer

The NPI formula may therefore need to be redeveloped / refined for each biological class
Concept of NPI +

A two tier approach:

• Initial identification of the biological class using a minimized panel of 10 biomarkers assessed IHC

• Application of a predefined bespoke NPI-like formula relevant to specific class of the patients tumour
Biomarkers used for classification

- Oestrogen Receptor (ER)
- Progesterone Receptor (PR)
- Cytokeratin 5/6
- Cytokeratin 7/8
- EGFR (HER1)
- c-erbB2 (HER2)
- c-erbB3 (HER3)
- c-erbB4 (HER4)
- P53
- Mucin 1
Seven Breast Cancer Phenotypic Classes

Breast Cancer

ER+
- Luminal CKs+
  - Luminal
    - PgR+
      - HER3+ HER4+
        - Luminal A Class 1 370 (34.4%)
      - HER3- HER4-
        - Luminal N Class 2 146 (13.6%)
    - PgR-
      - HER3+ HER4+
        - Luminal B Class 3 123 (11.5%)
  - HER2+
    - HER2
      - ER+
        - HER2/ER+ Class 4A 60 (5.6%)
      - ER-
        - HER2/ER- Class 4B 85 (7.9%)
  - ER-
    - p53+
      - Basal – p53 altered Class 5 126 (11.7%)
    - p53-
      - Basal – p53 normal Class 6 87 (8.1%)

Not classified = 76 cases (7.2%)
NPI+

After molecular classification we applied conventional prognostic variables to each class:

» Tumour Size (in cm.) \((Sz)\)
» Tumour Stage \((St)\)
» Number of positive nodes \((N)\)
» Mitotic Index \((M)\)
» Lymphovascular invasion (coded) \((LVI)\)
» Progesterone Receptor status code \((PgR)\)
» Ki67
Formulæ for each Class

1. \((0.41 \times \text{Positivenodes}) + (1.03 \times \text{Mitotic Index})\)
2. \((1.51 \times \text{Stage}) + (0.66 \times \text{Mitotic Index})\)
3. \((0.23 \times \text{Positive nodes}) + (0.85 \times \text{Mitotic Index})\)
4. \((0.14 \times \text{Positive nodes}) + (0.33 \times \text{Size})\)
5. \((0.15 \times \text{Nodes positive}) + (1.23 \times \text{VI})\)
6. \((0.80 \times \text{Stage}) + (0.32 \times \text{Mitotic Index})\)
7. \((1.28 \times \text{Stage}) + (0.87 \times \text{VI})\)
NPI+
Genome Sequencing

- Human Genome Project (Lander, 2001; completed 2003)
  → 3 GB, 13 years, $3 billion USD

- Celera Genome Sequencing Project (Venter, 2001)
  → 3 GB (5.11-fold coverage), <2 years, $300 million USD

- NGS, 454 Roche (Wheeler, April 2008)
  → 3 GB (7.4-fold coverage), 2 months, $60,000 USD

- NGS, 2012
  → entire genome, 1 day, < $1000 USD (twice cost of HER2 FISH)
# Applications of NGS

- **Genetic analysis**
  - Whole genome sequencing
    - Mutations
    - Copy number
    - Somatic rearrangements
  - Targeted genome sequencing

- **Transcriptomic analysis**
  - Digital gene expression
    - Expression profiling
  - RNA sequencing
    - Expression profiling
    - Identification of novel splice variants
    - Mutation analysis in coding genes
    - Fusion genes and 'readthroughs'
  - Small RNA profiling
    - Unbiased miRNA profiling
    - Identification of novel species

- **ChIP sequencing**
- **Epigenomics**
  - Methylation profiling
- **Functional studies**
  - shRNA screen deconvolution

## Advantages

- More efficient interrogation of structural variations and study of clonal evolution
- Deciphering intratumoural heterogeneity
- Potential for identification of previously unrecognized cancer related genes and aberrations
What have we learnt from NGS?

Mutational Landscape of BC at base-pair level

>30k somatic mutations are identified. Overall, BC shows a mean of 57 (range 5–374) somatic mutations per cancer

Mutation frequencies are similar to ovarian or renal clear cell carcinomas, but lower than bladder urothelial or lung squamous cell carcinoma

*Only* P53, PIK3CA and GATA3 *are* mutated in >10% of BC

Differentiate between Driver and Passenger mutations & Founder mutations and Progressor mutations
What have we learnt from NGS?

Identified a very large genetic diversity among different breast tumours

Improved our knowledge of intertumour heterogeneity and deciphered intratumoural heterogeneity

Mutated cancer genes can be grouped into certain pathways; p53, HER2, PI3K,...) so some of these genetically different tumours could be phenotypically similar due to mutations in the same pathway

Revealed the complex landscape of genetic rearrangement of different molecular classes
Intra-tumour genetic heterogeneity

- Tumours are composed of tumours cells
  - Diverse phenotype
  - Distinct genetic aberrations
- Although subclonal mutations present in all tumours, a dominant clone 50-95% of tumour always present: clonal and subclonal mutations

Tackling intratumor genetic heterogeneity

A
- Tumour cell with mutation 1
- Tumour cell clone with mutations 1+2
- Tumour cell clone with mutations 1+3

B
- Not actionable?
Improved knowledge of cancer biology

Improved prognostication  Prediction of response

Refined classification
Personalised medicine