Assessment of Linkage of SEER Breast Cancer Cases to Oncotype Dx Tests

NAACCR Annual Conference
St Louis, MO

Valentina Petkov, MD, MPH

June 14th, 2016
Objectives

- Background
- Linkage of SEER Breast Cancer (BC) cases to Oncotype DX results: Methods and linkage evaluation
- Comparison of registry collected and linked Oncotype DX
- Examples of recent research with the linked data
- Future directions
Background: SEER and Multigene Signature Tests (MGT) for BC

- Collected since 2010 under SSF22&23
  - SSF22: test name
  - SSF23: test result

- 2010-2012 SEER data
  - 10-13% of all BC cases had MGT
  - 24-30% if restricted to cases meeting the guidelines
  - 94% of all MGT were Oncotype DX
Background: Oncotype DX

- 21 gene assays (16 prognostic + 5 control genes)
- Developed and validated first for LN-, HR+, HER2- BC
  - In clinical use since 2004
- Validated for LN+, HR+, HER2- BC in 2008
- Oncotype DX DCIS in 2012 (16 genes:11+5)
- Genomic Health, Inc. (GHI) – the only lab performing the test
Background: Clinical Utility of Oncotype DX

- Prognostic – risk of distant recurrence
- Predictive – benefit of chemotherapy

- NCCN and ASCO recommended the test in 2008 for LN-
- NCCN recommended the test for LN+ in 2015
- ASCO did not recommend for LN+ (2016)
- Recurrence Score (0-100)
- Risk categories
  - Current: Low (<18); Interm. (18-30); High (>30)
  - TAILORx: Low (<12) Interm (12-25); Hight (>25)
Linkage SEER - GHI Data

- Linked
  - 411,585 GHI test orders processed 2004-2013
  - 649,311 SEER BC cases dx 2004-2012
- Each SEER registry was linked with all GHI cases
- Linkage methods
  - LinkPlus initially (cut-off of 7)
  - SAS algorithm to refine the match
  - Registry adjudication of uncertain matches
Linkage Evaluation

- 2000 randomly selected cases from:
  - Non-matches with LinkPluse score of 5&6
  - Matches per SAS algorithm
  - Non-matches per SAS algorithm
- Proportionally distributed based on registry case #
- Review by registries did not identify any errors
Linkage Evaluation (cont.)

- Compare SSF22=10 in 4 registries to having a match in GHI data
  - SAS algorithm classified 103 cases with Oncotype DX in SEER as non-matches in GHI data
  - Manual review rejected 680 cases with Oncotype DX in SEER as non-matches
  - In total 2112 (8.3%) cases with Oncotype DX in SEER were not matched
Linkage Evaluation (cont.)

- Reasons:
  - DCIS issue
  - Matching variables quality
  - Data errors in registry data
  - Different adjudication practices
    - Significant variability in proportion of cases w/ Oncotype DX in SEER rejected as non-matches
      - Range 1% to 19%
What did we gain?

67,842 test results

7,804 orders/tests flagged

Recurrence Score Distribution

<table>
<thead>
<tr>
<th>Percent</th>
<th>0</th>
<th>10</th>
<th>20</th>
<th>30</th>
<th>40</th>
<th>50</th>
<th>60</th>
</tr>
</thead>
<tbody>
<tr>
<td>low risk</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>55.2</td>
<td></td>
</tr>
<tr>
<td>intermed risk</td>
<td>36.7</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>high risk</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>8.1</td>
</tr>
</tbody>
</table>

Distribution of flagged orders or tests

- Non-Unique - Patient had multiple tumors: 20.1%
- Outside Analysis Criteria: 36.6%
- Failed Test - Inadequate tissue sample: 21.7%
- Cancelled Test - Cancelled by physician or patients: 21.5%
Quality Assessment SSF22&23

- Inclusion criteria
  - Cases dx 2010-2012
  - In situ excluded

- Questions
  - Completeness
  - Accuracy
  - Reporting bias
# Completeness or Added Value

<table>
<thead>
<tr>
<th>Oncotype Dx tests manually collected and provided through linkage</th>
<th>Test ordered</th>
<th>Test result available</th>
<th>Percent w test results</th>
</tr>
</thead>
<tbody>
<tr>
<td>2010-2012 cases with Oncotype Dx</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Manually collected by SEER (N cases)</td>
<td>25,427</td>
<td>23,992</td>
<td>94.4</td>
</tr>
<tr>
<td>Provided through linkage (N cases)</td>
<td>39,983</td>
<td>36,059</td>
<td>90.2</td>
</tr>
<tr>
<td>Only manually collected (N cases)</td>
<td>2,112</td>
<td>1,769</td>
<td>83.8</td>
</tr>
<tr>
<td>Only provided through linkage (N cases)</td>
<td>16668</td>
<td>15005</td>
<td>90.0</td>
</tr>
<tr>
<td>Proportion of tests provided by linkage but not captured in SEER</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Added value (% increase from baseline)</td>
<td>57.2</td>
<td>50.3</td>
<td>n/a</td>
</tr>
</tbody>
</table>
## Discrepancies between registry collected and provided through linkage Oncotype DX

<table>
<thead>
<tr>
<th>Results reported by registry (based on SSF23)</th>
<th>RS risk categories per GHI</th>
<th>Flagged Tests per GHI</th>
<th>Not linked</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Low</td>
<td>Intermed</td>
<td>High</td>
</tr>
<tr>
<td>Low risk</td>
<td>11746</td>
<td>414</td>
<td>13</td>
</tr>
<tr>
<td>Intermediate risk</td>
<td>71</td>
<td>6974</td>
<td>68</td>
</tr>
<tr>
<td>High risk</td>
<td>45</td>
<td>44</td>
<td>1516</td>
</tr>
<tr>
<td>Test done; result not in chart</td>
<td>384</td>
<td>240</td>
<td>37</td>
</tr>
<tr>
<td>Result unknown</td>
<td>123</td>
<td>82</td>
<td>21</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>12369</strong></td>
<td><strong>7754</strong></td>
<td><strong>1655</strong></td>
</tr>
</tbody>
</table>
Correlation between registry collected and linkage Oncotype DX scores

kappa = 94.6
Summary of discrepancies

- 8.3% w/ Oncotype DX in SEER did not find match in linkage data
- 4.9% RS did not match
- 2.7% RS discrepancy placed the case in different risk group
- 3.5% of cases w/ Oncotype Dx test in but no result in SEER were supplemented w RS
- 0.1% had results reported in SEER but no result is provided through the linkage (tests were flagged as cancelled or failed)
Reporting Bias

- Minimal differences between manually collected and linkage only Oncotype DX in age groups, race, SES index, marital status, insurance, SEER SS, grade, tumor size and hormonal status

- Cases with linkage only provided Oncotype Dx tend to have more missing information compared to cases with registry collected Oncotype DX particularly in the treatment related variables (chemo and radiation)
Examples of Recent Research w/ linked data

Five-year estimates of breast cancer-specific mortality by Recurrence Score group in node negative, HR+, HER2- BC

Petkov VI et al; npj Breast Cancer, June 2016
Examples of Recent Research w/ linked data

5-year BCSM by age group and RS group in Tested and Untested Patients

Petkov VI, ASCO 2016
Examples of Recent Research w/ linked data

Breast cancer-specific mortality at 5 years, by Recurrence Score group and number of positive nodes.

Roberts M, ASCO 2016
Examples of Recent Research w/ linked data

Oncotype DX dissemination by SEER registry in LN0 or LN1mic, HR+ female BC

Cronin K, ASCO 2016
Future Directions

➢ 2016 linkage is underway
   ▪ Linking 2013 BC cases to GHI 21 gene assay tests from 2012-2014
   ▪ Linking 2011-2013 BC cases to GHI 16 gene assay tests form 2011-2014

➢ Linkage improvements
   ▪ Improved SAS algorithms decreased registry review by 2/3
   ▪ Registry review of uncertain matches to include additional variables: date of surgery, MRN, facility #, registry collected Oncotype DX
   ▪ Planned survey with linkage adjudicating staff to better understand differences in accepting/rejecting a match
Future Directions

- Plan to add Oncotype DX to SEER-Medicare linkage
- Oncotype DX incorporated in staging in AJCC8
- ASCO guidelines recommended in 2016
  - PAM50
  - Breast Cancer Index
  - EndoPredict
  - Urokinase plasminogen activator (uPA) and plasminogen activator inhibitor type 1 (PAI-1)
## MGT in Early Stage BC

<table>
<thead>
<tr>
<th>Multigene Signature tests</th>
<th>Methodology</th>
<th>Targeted patients</th>
<th>Test results</th>
<th>Test available</th>
<th>Guidelines</th>
<th>Notes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Oncotype DX®, Genomic Health, Redwood City, CA, USA</td>
<td>RT-PCR based 21 gene assay (16+5 control)</td>
<td>HR+ w/o mets</td>
<td>score 0-100; low (&lt;18), intermed (18-30) and high (&gt;30) risk distant recurrence</td>
<td>2004-LN-; 2008-LN+; 2012-DCIS</td>
<td>NCCN-2008, 2015-LN+ ASCO-2007, St Gallen</td>
<td>TAILORx prospective trial re benefit chemo in RS 11-25.</td>
</tr>
<tr>
<td>PAM-50 ROR®, Prosigna; Nanostring Technologies, Seattle, WA, USA</td>
<td>Nanostring technology; quantify mRNA expression of 50 genes used in the PAM50 molecular classification algorithm and a series of housekeeping genes</td>
<td>stage I/II (including one to three positive nodes), ER-positive breast cancer in postmenopausal women treated with adjuvant endocrine therapy</td>
<td>recurrence score which reflects but does not explicitly report the intrinsic breast cancer subtype (luminal A, luminal B, etc.)</td>
<td>9/2013</td>
<td>ASCO, St Gallen (ESMO)</td>
<td>FDA approved, Sep-2013; EU clearance; can be performed by local pathology laboratories although it requires an expensive piece of equipment, the Nanostring nCounter Dx Analysis System (Nanostring Technologies)</td>
</tr>
<tr>
<td>Breast Cancer Index®, BioTheranosics, San Diego, CA, USA</td>
<td>two independent biomarkers, the HOXB13:IL17BR ratio and a five-gene molecular grade index that primarily consists of proliferation-related genes</td>
<td>ER+, LN-</td>
<td>score; also likelihood of benefit from extended endocrine therapy</td>
<td></td>
<td>ASCO, St Gallen (ESMO)</td>
<td>Better than ICH4 and Oncotype Dx in predicting late recurrence</td>
</tr>
<tr>
<td>EndoPredict®, Sividen Diagnostics GmbH, Koln, Germany</td>
<td>RT-PCR based assay (8 cancer genes + 3 control genes)</td>
<td>ER positive, HER2 negative</td>
<td>Low and high risk of recurrence, including late recurrence</td>
<td></td>
<td>ASCO, St Gallen (ESMO)</td>
<td>Marketed in Europe as diagnostic kit; Epclin risk score combines test and LN status and tumor size</td>
</tr>
<tr>
<td>Urokinase Plasminogen Activator (uPA) and Plasminogen Activator Inhibitor type 1 (PAI-1)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>ASCO</td>
<td>fresh frozen tissue</td>
</tr>
<tr>
<td>MammaPrint®, Agilent, Amsterdam, the Netherlands</td>
<td>Microarray based; measures the mRNA expression of 70 genes</td>
<td>&lt;61 years of age with stage I/II, LN- or LN+ (1-3); HR+/--; HER2+</td>
<td>Low and high risk of recurrence</td>
<td>St Gallen (ESMO)</td>
<td>Approved by FDA and EU as prognostic for 1-5 y recurrence in target patients; not good for long term recurrence; Results from MINDACT pending</td>
<td></td>
</tr>
<tr>
<td>Genomic Grade Index MapQuant Dx, Ipsogen, France</td>
<td>microarray-based assay that measures the expression of 97 genes to assign a molecular grade.</td>
<td>ER positive, intermediate grade</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ICH4 index</td>
<td>multivariate model that uses semiquantitative information from immunohistochemical assessment of ER, PR, HER2 and Ki67</td>
<td></td>
<td>Risk score calculated by a formula</td>
<td></td>
<td>in the absence of standardized quantification of each of the four variables applying the formula to local pathology results could be highly misleading.</td>
<td></td>
</tr>
<tr>
<td>Mammaostrat , Clarient, a GE Healthcare company, Aliso Viejo, CA</td>
<td></td>
<td>HR+</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Acknowledgment

- SEER Registries
- Genomic Health Clinical Lab
- IMS
- Nicki Schussler
- Will Howe
Thank You

Questions?

Contact Valentina Petkov at petkovvi@mail.nih.gov