Biomarker Development: Case Study of Ovarian Cancer Recurrence

Christina Kendziorski, Kevin Eng
Janet Rader, and William Bradley

Department of Biostatistics and Medical Informatics
University of Wisconsin-Madison
Biomarker development

- Decide on a clear end-point.
  - What will the impact be?
  - What is the best in the area?
  - What is under development?

- Extra care with normalization and qc: batch effects, study designs

- Moving to clinic? Keep it simple

- Validate in independent populations
  - Cross validation has its advantages, but…
  - Independent populations are better
The Cancer Genome Atlas Project

...started in 2005 with ovary, lung, and brain

...can we use this data to help guide ovarian cancer treatment at recurrence?
Ovarian Cancer Overview

- 5th leading cause of cancer death among American women
- ~22,000 new cases in 2011 with ~14,000 deaths
- 5 year survival < 50% (all types)
Overall goals: prognosis and prediction of treatment response

75% of patients present in Stage III, IV

Surgery and platinum based chemotherapy

75% recur

25% cured

Platinum
Bevacizumab
Docetaxel
Doxorubicin
Gemcitabine
Navelbine
Tamoxifen
Topotecan
Expression based predictors of recurrence

- Oncotype DX for breast cancer recurrence (2005)
  - 21 gene signature measured via RT-PCR
  - 70 gene signature measured via microarrays
- Oncotype DX for colon cancer recurrence (2010)
- Coloprint for colon cancer recurrence (2012)
- Oncotype DX in trials for renal and prostate cancers
Best predictors of OC overall survival and recurrence

- Cytoreduction with age, stage, grade, CA-125
Biomarker development

- Decide on a clear end-point.
  - What will the impact be?
  - What is the best in the area?
  - What is under development?

- Extra care with normalization and qc: batch effects, study designs

- Validate in independent populations
  - Cross validation has its advantages, but…
  - Independent populations are better

- Moving to clinic? Keep it simple
Prognosis of Breast Cancer: MammaPrint

- 70 gene microarray that predicts breast cancer recurrence
- Approved by the FDA, February, 2007
- van’t Veer et al., *Nature*, 2002

- Time’s Best Invention of the Year, 2007
Dendogram - Mammaprint Data

dd

hclust ("p", "complete")
Dendogram - Mammaprint Data (POS10 Removed)
Dendogram - Mammaprint Data (POS10 and NEG37 Removed)
Mammaprint Diagnostics: Observational Model

Median
Mammaprint Diagnostics: Observational Model
CV Diagnostic - Mammaprint Data
CV Diagnostic - Mammaprint Data

![Graph showing the relationship between variance and transcript values.](image-url)
Best predictors of OC overall survival and recurrence

- Cytoreduction with age, stage, grade, CA-125
Comparable predictors of OC overall survival

**Cytoreduction**
- Low: 38%
- High: 62%

**TCGA193**
- Low: 27%
- High: 74%

**CLOVAR100**
- Low: 25%
- High: 76%

Signature x Cytoreduction
That’s progress…but they don’t tell you about underlying biology and/or how to treat
Core Signaling Pathways in Human Pancreatic Cancers Revealed by Global Genomic Analyses


There are currently few therapeutic options for patients with pancreatic cancer, and new insights into the pathogenesis of this lethal disease are urgently needed. Toward this end, we performed a comprehensive genetic analysis of 24 pancreatic cancers. We first determined the sequences of 23,219 transcripts, representing 20,661 protein-coding genes, in these samples. Then, we searched for homozygous deletions and amplifications in the tumor DNA by using microarrays containing probes for $\sim 10^6$ single-nucleotide polymorphisms. We found that pancreatic cancers contain an average of 63 genetic alterations, the majority of which are point mutations. These
- 24 human pancreatic tumor samples
  - SNP/CNV chips assessed genomic amplifications and deletions
  - Sequence obtained for approximately 25% of the genome
  - Gene expression quantified using next-generation sequencing

- Integrative Analysis indicates
  - Tumors contain an average of 63 genetic alterations
  - Specific genes altered in each tumor are largely different
  - Specific pathways altered are similar, with 12 pathways altered in 67%-100% of tumors.

  - Jones et al., Science, September 2008
“The pathway perspective helps bring rudimentary understanding to a very complex disease…the best hope for therapeutic development may lie in the discovery of agents that target the physiologic effects of altered pathways and processes rather than their individual gene components.”

- Jones et al., Science, September 2008
Goal is to develop a predictor of survival and recurrence that accommodates pathway structure, allows for personalized assessment of risk, and tells us something about biology...how to treat
Pathway Index Model and Data

- KEGG annotations map genes to pathways
- For each pathway,
  - Fit a lasso penalized Cox proportional hazards model to expression
  - Signs of coefficients reflect “susceptibility” (S) or “resistance” (R) genes
  - Construct a risk index:
    
    Average expr. of S genes - Average expr. of R genes
  - Combine pathway indices (comment on this later)
- For each of ~ 500 TCGA patients,
  - mRNA via Affy’s HT- HGU133A (~12,000 genes)
  - clinical information (age, stage, grade, surgery, treatment…)
- ~ 500 patients for validation from Tothill et al. (GEO: GSE9891) and Yoshihara et al., (GEO: GSE17260)
Pathway Index Model

Define a pathway
Pathway Index Model

Select genes
Pathway Index Model

Susceptibility genes

Diagram showing the relationship between various genes involved in mismatch repair, such as POLD1, POLD2, POLD3, RFC2, RFC3, RFC4, RFC5, RPA3, SSBP1, MSH2, MSH6, and MLH3.
Pathway Index Model

Susceptibility genes

Resistance genes

Diagram showing the relationships between various genes involved in mismatch repair.
Pathway Index Model

\[ Z_{ik} = \overline{X}_{ik}^S - \overline{X}_{ik}^R \]
### Patient-specific risk profile

<table>
<thead>
<tr>
<th>Patients</th>
<th>Apoptosis</th>
<th>Hedgehog</th>
<th>Vegf</th>
<th>Jak-STAT</th>
<th>...</th>
<th>MAPK</th>
</tr>
</thead>
<tbody>
<tr>
<td>TCGA-13-0758</td>
<td>0.449</td>
<td>-0.267</td>
<td>-0.803</td>
<td>-0.260</td>
<td>...</td>
<td>0.538</td>
</tr>
<tr>
<td>TCGA-09-0364</td>
<td>0.382</td>
<td>0.575</td>
<td>0.064</td>
<td>-0.650</td>
<td>...</td>
<td>-0.144</td>
</tr>
<tr>
<td>TCGA-13-0800</td>
<td>1.315</td>
<td>0.926</td>
<td>-0.565</td>
<td>0.036</td>
<td></td>
<td>-0.910</td>
</tr>
</tbody>
</table>

### PSRP: patient-specific risk profile

<table>
<thead>
<tr>
<th>Patients</th>
<th>Apoptosis</th>
<th>Hedgehog</th>
<th>Vegf</th>
<th>Jak-STAT</th>
<th>...</th>
<th>MAPK</th>
</tr>
</thead>
<tbody>
<tr>
<td>TCGA-13-0758</td>
<td>Risk+</td>
<td>Risk-</td>
<td>Risk-</td>
<td>Risk-</td>
<td>...</td>
<td>Risk+</td>
</tr>
<tr>
<td>TCGA-09-0364</td>
<td>Risk+</td>
<td>Risk+</td>
<td>Risk+</td>
<td>Risk-</td>
<td>...</td>
<td>Risk-</td>
</tr>
<tr>
<td>TCGA-13-0800</td>
<td>Risk+</td>
<td>Risk+</td>
<td>Risk-</td>
<td>Risk+</td>
<td>Risk+</td>
<td>Risk-</td>
</tr>
</tbody>
</table>

Eng et al., Statistics in Medicine, 2012
Patient specific indexes

Few patients share same set of abnormal pathways

…..which makes it difficult to infer common structure
...but the number of S+ pathways may be informative
Tothill et al., CCR 2008
Bild et al., Nature, 2006
Group patients by pathway and treatment

Log-rank test $p = 0.0093$

$n=32$ Drug- Path R-

$n=44$ Drug- Path R+

$n=12$ Drug+ Path R-

$n=21$ Drug+ Path R+
NT6: A six-gene signature derived from the Pathway-Index Model
NT6 is derived from the neurotrophin growth factor pathway
NT6 prognosis is comparable with leading methods

CLOVAR100: 100 gene signature from Verhaak et al., Journal of Clinical Investigation, 2013
NT6 predicts progression-free survival

Cox Model $p=0.048$
$n=498$

Spline Model $p=0.018$
$n=501$
NT6 is associated with the development of platinum resistance

A2008: Platinum Sensitive

OVCAR5: Platinum Resistant

IGROV1: Platinum Sensitive

IGROV8: Platinum Sensitive
NT6 knock-down slows cell growth
NT6 stratifies patients

Patient NT6 Indexes
NT6 stratifies patients with differential treatment response

<table>
<thead>
<tr>
<th>Treatment</th>
<th>All</th>
<th>Gp1</th>
<th>Gp2</th>
<th>Gp3</th>
<th>Gp4</th>
<th>Gp5</th>
<th>Gp6</th>
<th>p-val</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pt./Taxol</td>
<td>80 n=47</td>
<td>392 n=4</td>
<td>52 n=16</td>
<td>576 n=5</td>
<td>30 n=9</td>
<td>97 n=6</td>
<td>413 n=7</td>
<td>0.006</td>
</tr>
<tr>
<td>Pt./Gemzar</td>
<td>88 n=18</td>
<td>NA</td>
<td>34 n=4</td>
<td>165 n=3</td>
<td>124 n=5</td>
<td>62 n=3</td>
<td>60 n=3</td>
<td>0.210</td>
</tr>
<tr>
<td>Doxil</td>
<td>36 n=31</td>
<td>17 n=3</td>
<td>292 n=4</td>
<td>NA</td>
<td>36 n=7</td>
<td>164 n=4</td>
<td>35 n=8</td>
<td>0.022</td>
</tr>
<tr>
<td>Topotecan</td>
<td>31 n=26</td>
<td>275 n=2</td>
<td>30 n=10</td>
<td>63 n=1</td>
<td>29 n=9</td>
<td>1482 n=1</td>
<td>52 n=3</td>
<td>0.13</td>
</tr>
</tbody>
</table>
NT6+ measured via RT-PCR from paraffin embedded tissues
Are changes in NT6 meaningful?

Pre-Treatment NT6

NT6 High
NT6 Low

Days to first recurrence

Percent change in NT6

-10%  -5%   0%   +5%   +10%
Ovarian Cancer Overview

- 5th leading cause of cancer death among American women
- ~22,000 new cases in 2011 with ~14,000 deaths
- 5 year survival < 50% (all types)
Very recent results

Non-augmented: 52% increased risk of recurrence with increasing NT6
Augmented: 6% decreased risk
Biomarker development - recap

- Decide on a clear end-point (treatment response for OC recurrence)
  - What will the impact be? (potentially big as most OC patients recur)
  - What is the best in the area? (cyto predicts recurrence, but not tr)
  - What is under development? (nothing that we know of…)

- Extra care with normalization and qc: batch effects, study designs
  - TCGA has data from multiple hospitals collected over decades
  - Sicker patients get different treatments

- Moving to clinic? Keep it simple
  - Cox PH Lasso regression gives 6 genes; RT-PCR assay to measure

- Validate in independent populations
  - We considered two independent populations during development
  - We are collecting patients to continue to validate