Data-Driven Medicine

Daphne Koller
Stanford University
Electronic Health Records: A New Opportunity

- Discrete Events: Laboratory Interventions: Medicines, Procedures
- Continuous physiologic measurements
- Progress notes
- Imaging Data
Outline

- Risk stratification for premature infants
  - Joint work with Suchi Saria, Anna Penn, Anand Rajani, and Jeff Gould

- Prognostic stratification for breast cancer
  - Joint work with Andy Beck, Matt van de Rijn, Ankur Sangoi, Samuel Leung, Robert Marinelli, Torsten Nielsen, Marc van de Vijver, Robert West
Continuous physiologic monitoring

What information is hiding here?

Recurring signatures discovered using unsupervised learning:

- Signatures 2 and 5 appear mostly in sick infants
- Signatures 3, 9 and 10 more common in healthy infants
- “Bad” signatures have lower entropy

Predicting Neonate High Morbidity

SNAP and SNAPPE-II – 12 hours
(Richardson et al, J. Pediatrics 2001)

CRIB - 12 hours
(International Neonatal Network, Lancet 1993)

Apgar - at birth

• Performs consistently better
• Dramatically outperforms Apgar, current standard of care:
  • **AUC 0.91 vs 0.69**
Invasive Laboratory Tests

- Tests added:
  - White Blood cell count
  - Band neutrophils
  - Hematocrit
  - Platelet Count
  - Blood gas measurements (\(\text{PaO}_2\) and \(\text{PaCO}_2\))
  - pH

- Invasive lab measurements do not add to Physiscore’s predictive value

## Comparison Summary

<table>
<thead>
<tr>
<th></th>
<th>Physiscore (Our tool)</th>
<th>APGAR (Standard of care)</th>
<th>CRIB</th>
<th>SNAP-II</th>
<th>SNAPPE-II</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Time from birth</strong></td>
<td>3 hours after birth</td>
<td>5 mins</td>
<td>12 hours</td>
<td>12 hours</td>
<td>12 hours</td>
</tr>
<tr>
<td><strong>Accuracy</strong></td>
<td>0.91</td>
<td>0.69</td>
<td>0.85</td>
<td>0.82</td>
<td>0.87</td>
</tr>
<tr>
<td><strong>Invasive testing</strong></td>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
<td>X</td>
</tr>
</tbody>
</table>

THE POSITION OF
HISTOLOGY IN THE PROGNOSIS OF
CARCINOMA OF THE BREAST.

By D. H. PATEY, M.S. Lond., F.R.C.S. Eng.,
SURGICAL REGISTRAR, MIDDLESEX HOSPITAL;
AND
R. W. SCARFF, M.R.C.S. Eng.,
ASSISTANT PATHOLOGIST, BLAND-SUTTON INSTITUTE OF PATHOLOGY,
MIDDLESEX HOSPITAL.

[April 21, 1928  801
THE LANCET,]
Cancer Grading System

Results of Investigation.

If the cases are subdivided into groups according to the histological index of malignancy determined as above, there are:

GROUP 1.—Slight histological malignancy, 16 cases.
- Alive and well: 11 cases
- Died of recurrence: 2 cases
- Died following operation: 1 case
- Died of other causes: 2 cases

GROUP 2.—Moderate histological malignancy, 12 cases.
- Alive and well: 5 cases
- Died of recurrence: 6 cases
- Died of other causes: 1 case

GROUP 3.—Marked histological malignancy, 22 cases.
- Alive and well: 5 cases
- Died of recurrence: 17 cases

Thus there does appear to be a progressive deterioration in the results as the histological index of malignancy increases. This supports the conclusions of Greenough, who found:

Group 1: 68 per cent. of cases were cured.
Group 2: 33 per cent. of cases were cured.
Group 3: 0 per cent. of cases were cured.

This same basic system is used today!
What information is hiding here?

Break Up Image into Superpixels

Classify superpixels

- Two main types of cells:
  - Epithelial cells – cancer
  - Stromal cells – surrounding tissue

- Train classifier to distinguish superpixels

(Classifier: epithelium vs stroma)
Define broad range of features

- Characterize epithelial and stromal cells
- Spatial relationships
- Contextual features

Clinical samples

- H&E stained tissue microarray images from breast cancer patients from 2 institutions

  - Netherlands Cancer Institute (NKI) – n=248
    - women less than 53 years with Stage I or Stage II breast cancer

  - Vancouver General Hospital (VGH) – n=328
    - a population-based cohort with a higher proportion of older women and women with more advanced disease.

Building a prognostic model

- Prognostic model trained to the binary target of 5 year survival.
- Algorithm: $L_1$-regularized logistic regression
- NKI dataset was used for building the model
  - NKI performance assessed by 8-fold cross-validation
- VGH dataset was never used for training, but only for model validation

Prognosis of 5-year survival


A. C-Path 5YS Model on the NKI Cohort (n=248)

B. C-Path 5YS Model on the VGH Cohort (n=328)

P<0.001
Significant stratification even within grade

Grade only weakly correlated with outcome in VGH dataset

11 most predictive features

Presence of stromal objects without nuclei

Average relative border of stromal spindled nuclear objects to stromal round nuclear objects

8 Epithelial Features

3 Stromal Features

Stromal features not used at all in current pathological analysis

Data Driven Medicine

- Huge amounts of medical data are now being collected every day
- Much of the previous work is hypothesis-based: testing specific, human-constructed hypotheses
- Unbiased, data-driven analysis can discover novel and important signatures
  - Short-term variability in neonate heart-rate
  - Importance of stromal tissue
- Can be used for prognosis, treatment guidance
- Can provide new insights on disease processes