Electronic Medical Records Data for Individualized Health: Application to Low Risk Prostate Cancer

Yates Coley, PhD
Postdoctoral Fellow
Department of Biostatistics, JHSPH
rycoley@jhu.edu

August 13, 2015
Goal:
Build principled statistical models that integrate all the available data in order to inform clinical decision-making
Goal:
Build a statistical model that integrates all the available data to inform clinical decision-making

Statistical Challenges
1. Disease state not directly measured
2. Informative missingness
Active Surveillance of Prostate Cancer
Active Surveillance of Prostate Cancer

Key to Success: Correctly identify potentially lethal cancers
Active Surveillance of Prostate Cancer

Key to Success: Correctly identify potentially lethal cancers

Challenge: Disease state not directly measured
Active Surveillance of Prostate Cancer

**Key to Success:** Correctly identify potentially lethal cancers

**Challenge:** Disease state not directly measured
Active Surveillance of Prostate Cancer

**Key to Success:** Correctly identify potentially lethal cancers

**Challenge:** Disease state not directly measured
Age (years) | PSA (ng/mL) | Reclassification
---|---|---
64 | 1 | No
66 | 5 | No
68 | 10 | Yes
70 | | Biopsy Upgrading
72 | | Yes

©Hopkins inHealth 2015
True Prostate Cancer Status

Random Variability

True PSA

Measurement Error

Observed PSA

Biopsy Results

Measurement Error

©Hopkins inHealth 2015
Latent Class
- Indolent
- Lethal

Random Variability

True PSA

Measurement Error

Observed PSA

True Prostate Cancer Status

Measurement Error

Biopsy Results
Gold standard
Pathologic analysis observed in subset

True Prostate Cancer Status

Random Variability

Measurement Error

Observed PSA

True PSA

Biopsy Results

Measurement Error

©Hopkins inHealth 2015
Time-varying Biomarker

True Prostate Cancer Status

Random Variability

Measurement Error

Observed PSA

True PSA

Measurement Error

Biopsy Results

©Hopkins inHealth 2015
Discrete Time-to-Event

True Prostate Cancer Status

Random Variability

True PSA

Measurement Error

Observed PSA

Measurement Error

Biopsy Results

©Hopkins inHealth 2015
True Prostate Cancer Status

Observed PSA

Biopsy Results

Observed PSA

Biopsy Results

Time
True Prostate Cancer Status

Biopsy Results

Observed PSA

Individual-Level Random Effects

Biopsy Results
\[ L_i \propto P(\text{Cancer State}_i) \]

\[ \times f(\text{PSA}_i | X_i, Z_i, \text{Random Effects}_i) g(\text{Random Effects}_i | \text{Cancer State}_i) \]

\[ \times \prod_j P(\text{Biopsy Upgrade}_{ij} | W_{ij}, \text{Cancer State}_i) \]
We then proceed as above to get a re-weighted posterior for the latent variables of patient $k$.

\[
L_i \propto P(\text{Cancer State}_i) \\
\times f(\text{PSA}_i | X_i, Z_i, \text{Random Effects}_i) g(\text{Random Effects}_i | \text{Cancer State}_i) \\
\times \prod_j P(\text{Biopsy Upgrade}_{ij} | W_{ij}, \text{Cancer State}_i)
\]

**Pooled Logistic Regression**
We then proceed as above to get a re-weighted posterior for the latent variables of patient $k$. The likelihood of the model is given by:

$$L_i \propto P(\text{Cancer State}_i) \times f(\text{PSA}_i | X_i, Z_i, \text{Random Effects}_i) \times g(\text{Random Effects}_i | \text{Cancer State}_i) \times \prod_j P(\text{Biopsy Upgrade}_{ij} | W_{ij}, \text{Cancer State}_i)$$
We then proceed as above to get a re-weighted posterior for the latent variables of patient $k$.

\[
L_i \propto P(\text{Cancer State}_i) \times f(\text{PSA}_i | X_i, Z_i, \text{Random Effects}_i) \times g(\text{Random Effects}_i | \text{Cancer State}_i) \times \prod_j P(\text{Biopsy Upgrade}_{ij} | W_{ij}, \text{Cancer State}_i)
\]
Partially-latent class

\[ L_i \propto P(\text{Cancer State}_i) \]

\[ \times f(\text{PSA}_i | X_i, Z_i, \text{Random Effects}_i) g(\text{Random Effects}_i | \text{Cancer State}_i) \]

\[ \times \prod_j P(\text{Biopsy Upgrade}_{ij} | W_{ij}, \text{Cancer State}_i) \]

Bayesian Estimation:
Posterior Probability of Lethal Cancer
We then proceed as above to get a re-weighted posterior for the latent variables of patient $k$.

Partially-latent class

$$L_i \propto P(\text{Cancer State}_i)$$

$$\times f(\text{PSA}_i | X_i, Z_i, \text{Random Effects}_i) g(\text{Random Effects}_i | \text{Cancer State}_i)$$

$$\times \prod_j P(\text{Biopsy Upgrade}_{ij} | W_{ij}, \text{Cancer State}_i)$$

Challenge: Informative Missingness
True Prostate Cancer State (Latent)
True Prostate Cancer State (Latent)

Observed PSA

Surgical Removal (Observe True State)

Biopsy Results

Missing at Random

©Hopkins inHealth 2015
True Prostate Cancer State (Latent)

Observed PSA

Surgical Removal (Observe True State)

Biopsy Results

Missing NOT at Random
\[ L_i \propto P(\text{Cancer State}_i) \]

\[
\times f(\text{PSA}_i | X_i, Z_i, \text{Random Effects}_i) \ g(\text{Random Effects}_i | \text{Cancer State}_i) \\
\times \prod_j P(\text{Biopsy Upgrade}_{ij} | W_{ij}, \text{Cancer State}_i) \\
\times P(\text{Surgery}_{ij} | \text{History}_{ij}, W_{ij}, \text{Cancer State}_i) \\
\]
Prostate Cancer State (Latent)

Observed PSA

Receive Biopsy

Biopsy Results

Missing NOT at Random
$L_i \propto P(\text{Cancer State}_i) \times f(\text{PSA}_i | X_i, Z_i, \text{Random Effects}_i) \times g(\text{Random Effects}_i | \text{Cancer State}_i) \times \prod_j P(\text{Biopsy Here}_{ij} | \text{History}_{ij}, W_{ij}, \text{Cancer State}_i) \times P(\text{Biopsy Upgrade}_{ij} | \text{Biopsy Here}_{ij}, W_{ij}, \text{Cancer State}_i) \times P(\text{Surgery}_{ij} | \text{History}_{ij}, W_{ij}, \text{Cancer State}_i)$
**Diagnosis**

- P(Biopsy Upgrade): 40%
- P(Lethal PCa): 2%

**5 Years Follow-up**

- P(Biopsy Upgrade): 2%
- P(Lethal PCa): 2%

---

**Probability Reclassification**

- 0% 25% 50% 75% 100%

**PSA (ng/mL)**

- Age (years): 64 66 68 70 72
- 1 5 10 50

---

**Biopsy Performed**

- No Biopsy

---

©Hopkins inHealth 2015
Diagnosis

5 Years Follow-up

10 Years Follow-up

**P(Gleason 7+)**

*0% 100% 40% 55% 100%

Age (years)

PSA (ng/mL)

64 66 68 70 72

1 5 10 50

- Biopsy Performed
- No Biopsy

**P(Biopsy Upgrade)**

**P(Lethal PCa)**

©Hopkins inHealth 2015
Decision Support Tool

http://rycoley.shinyapps.io/prediction-app
Johns Hopkins BSPH Biostatistics

Aaron Fisher

JHMI Brady Institute of Urology

Bal Carter

Mufaddal Mamawala

Johns Hopkins SOM

Karthik Rao

Ken Pienta

Research sponsored by the Patrick C. Walsh Prostate Research Fund
AUC=0.75

AUC=0.72

Unadjusted
I.O.P.
Posterior $P(\text{Biopsy})$

Observed $P(\text{Biopsy})$