Developing a Clinical Predictive Model for Nosocomial Bloodstream Infections from the Electronic Medical Record

Gonzalo Bearman MD,MPH

CORNELL UNIVERSITY

Joan and Sanford I. Weill Medical College
Outline

- Background
- Project goal and methods
  - Setting
  - BSI definitions and the EMR at NYPH-Cornell
  - Choosing an analytic method/statistical model
  - Variables of interest/methods
    - EMR screens and data repository
- Results
  - Frequencies and univariate analyses
  - Multivariate model
  - Predictive model and internal validation
- Conclusion and future steps
Background

Nosocomial bloodstream infections are associated with significant morbidity and mortality

- 250,000 cases of nosocomial BSI per annum
- Attributable mortality of 35%
- Associated with excessive LOS- 24 days
- Excess hospital cost of $40,000 per survivor

Nosocomial BSI Risk Factors

- Altered mental status
  - Head injury, coma, medications

- Malnutrition

- Diagnosis
  - DM, immunosuppressive conditions

- Medications
  - Antacids, H2 Blockers, Immunosuppressives
  - Antibiotics (protective)

- Mechanical ventilation

- Invasive catheters
  - Central venous, arterial, and urinary

- Cutaneous disorders - stasis ulcers, decubiti
Study Question: Is it possible to identify patients at high risk for nosocomial BSI using an electronic medical record?
Project Goal

- Three phase study:
  1. Development of a clinical predictive model (Cox Proportional Hazards) using *known risk factors* for nosocomial BSI *exclusively* from the electronic medical record (EMR).
    - **NOT A CONVENTIONAL CHART REVIEW**
  2. Validate the model prospectively
  3. Apply the model in real time via the EMR to identify patients prospectively for risk reduction measures
    - **High Risk** status displayed as Screen Alert with automated decision support

Today’s topic of interest is **PHASE ONE** of the study
Study Team

Cornell University Departments

- Gonzalo Bearman MD, MPH (ID, PH)
- Lewis Drusin MD, MPH (ID, PH)
- Nathaniel Hupert MD, MPH (PH)
- Michael Oppenheim MD (ID, Med Informatics)
- Eneida Mendonca MD (Med Informatics-Columbia University)
- Paul Christos MS (PH, Statistics)
- Maryam Behta PharmD (Pharmacy)
Setting

- New York Presbyterian Hospital of Weill Cornell Center
  - 850 Bed tertiary care medical center
  - Located in Manhattan, New York City
  - Multi-specialty services including medicine, surgery, cardiology, cardiothoracic surgery, obstetrics and gynecology, pediatrics, and bone and solid organ transplants.
  - Approximately 40,000-45,000 patient discharges per annum
Study Population

Population

- Hospitalized patients from June 1ˢᵗ 2001 to June 1ˢᵗ 2002
- Medical and surgical services, including MICU, CCU, SICU.
  - Exclusively areas of the hospital where the EMR is fully available/integrated (excludes Burn Unit, NICU, and OB/GYN)
Nosocomial BSI Definition

Case of nosocomial BSI defined by CDC criteria

- Newly positive blood culture 48 hours post admission
- Capture 1st episode of nosocomial BSI
- Exclude:
  - patients with positive blood cultures within the first 48 hours
  - Patients with same organism bacteremia pre and post 48 hours
  - Coagulase negative staphylococci (exclude 342 isolates)
Method

1. Randomly divide cohort in half
   - Events vs non-events
   - All variables from EMR

2. Develop a predictive model
   - Cox-Proportional Hazards Model
     - From $\frac{1}{2}$ of cohort (Test cohort)

3. Assign a ‘scoring’ system
   - From variables in Cox-Model
     - Score(s) will influence PPV+, Sensitivity and Specificity

4. Validation of model
   - From other $\frac{1}{2}$ of cohort (Validation Cohort)
   - Statistical comparison of PV+, Sensitivity, Specificity between the test and validation cohort
Choice of analytic technique

Why Choose a Cox Proportional Hazards Model?

- Person time data available from database
- Model appropriate for unequal observation periods
- Power and Efficiency
- Regression model of multivariate analysis
EMR: Medical Informatics

Sunrise Critical Care v1.3 (Eclipsys Corp.)
- Available on nearly all units
- 100% nursing, nutrition, respiratory therapy documentation
- 100% physician order entry
- Pharmacy records of all medications during LOS
- Laboratory, microbiology, pathology, and radiology records
- Data is entered in discrete form and free text
- Clinical data repository (real time)

Wealth of Information that can be queried!!!!
Variables of interest from EMR:

- **1. Demographics**
  - Age
  - Gender

- **2. Diagnoses**
  - Diabetes mellitus
    - Keyword search- MED
    - Or Serum glucose >120 x 3
    - Or Pt receiving diabetic medications
Variables of interest

2. Diagnoses

- Immunosuppressive conditions
  - HIV/AIDS, malignancies, lymphomas, rheumatologic conditions
    - No ICD-9 Codes are available in clinical database
    - Free text search and match from MED (Medical Entities Dictionary) keywords
  - Or Pt receiving Immunosuppressive medications and antiretrovirals
Demographics, Admitting Diagnoses, Hospital Location

Registration Data
- Medical Record #: [redacted]
- Patient Name: [redacted]
- Unit: [redacted]

Patient Information
- Birthdate: 08/04/1969
- Sex: M
- Marital Status: SINGLE
- Age: 32 yrs
- Race: [redacted]

Admission Data
- Admitting Diagnoses: LEUKEMIA
- Date of Admission: 08/01/82

Religious Affiliation: [redacted]
- Accompanied by: Friend
- Physician notified of pt arrival: Dr. Klinger
- Date/Time notified: 03/30/02 08:00

Demographics, Admitting Diagnoses, Hospital Location
3. Neurologic Status:

Nurses neurologic assessment

Capture: episode of altered mental status- lethargy or comatose
Variables of Interest

4. Mechanical Ventilation
   - Categorical
   - Continuous variable
   - 0-7 days
   - > 7 days
Variables of Interest

5. Medications

- Physician order entry system
  - Antibiotics
  - Sedatives
  - Immunosuppressives
  - H2 blockers, sucralfate or proton pump inhibitors
6. Central Venous Line: Type, type and site, # of days

7. Arterial Lines: Body site and # of days
## 8. Skin/Wound Status: from Nurse’s Assessment

**Search:** ulcers/pressure ulcers/diabetic ulcers

### Pressure Ulcer Type
1. Stage 1 (Reddened area not resolved within 30 mins of repositioning)
2. Stage 2 (Superficial skin break or blister)
3. Stage 3 (Skin break exposing subcutaneous tissue)
4. Stage 4 (Skin break exposing muscle and bone)
5. Eroder (Blackened tissue)

### Pressure Ulcer Stage (measured weekly in cm):
- Length, width, depth

### Pressure Ulcer Appearance
- Reddened
- Blue/purple/darkened skin
- Pink
- Beady red
- Clean
- Excoriated
- Blistered
- Open areas
- Scabbled
- Necrotic
- Slough tissue
- Spongy/swollen
- Wound margins separated from slough
- Granulation tissue
- Healing
- Undermining
- Tunneling
- Other appearance: free tent

### Pressure Ulcer Drainage
- No drainage
- Drainage present
- Amount
- Color

### Pressure Ulcer Odor
- No odor
- Mild odor
- Malodorous

### Pressure Ulcer Treatment (Stage 1-2)
- Turned and positioned
- Pillow/wedges used for positioning
- Heels elevated off bed surface
- Heel pads
- Elbow pads
- HOB at or below 30 degrees
- Skin kept dry
- Assisted with active ROM
- Assisted with passive ROM
- Incontinence kit
- Gel cushion used when in chair
- Gel dressing
- Medicated (medicated) applied
9. Nutritional Status as assessed by nutrition service.

Dichotomize: Malnourished or not malnourished
Results

- **Total Cohort**
  - 245 distinct NBSI cases
  - 29,954 “non cases”

- **Test Cohort**
  - 123 distinct NBSI cases
  - 14,977 “non cases”

- **Validation Cohort**
  - 122 distinct NBSI cases
  - 14,977 “non cases”
Age: (Test Cohort)
- Mean=55.5 years
- HR=1.003 P=0.262

Gender (Test Cohort):
- Male=.48 Female=.52
- HR (Males)=1.40 P=0.06

Age and Gender were not significant predictors on univariate analysis
Frequency/Univariate Analysis

- **Unconscious/lethargic (Test Cohort):**
  - N=1214 HR=2.25 P<0.001

- **Malnutrition (Test Cohort):**
  - N=264 HR=1.65 P=0.17

- **Diabetes (Test Cohort):**
  - N=2900 HR=3.09 P<0.001

Unconscious and DM have increased HR on Univariate analysis
Malnutrition is not a significant predictor
Immunosuppressive condition (Test Cohort):
  - N=1097 HR=1.67  P=0.05

Medications (Test Cohort):
  - PPI N=2893 HR=3.97 P<0.001
  - IMS med N=1540 HR=2.88 P<0.001
  - H2RA N=1420 HR=2.03 P<0.001
  - DM med N=1987 HR=3.68 P<0.001
  - Antibiotics N=4933 HR=8.18 P<0.001
  - Alt MS med N=7826 HR=6.78 P<0.001

Immunosuppressive condition was significant on univariate analysis.
All medications significant predictors on univariate analysis.
Frequency/Univariate Analysis

Mechanical Ventilation (Test Cohort)
- Number of patients N=877
- HR=2.04
- P<0.001

Mechanical ventilation has an elevated HR on Univariate Analysis
Frequency Analysis

**Invasive Venous Catheters (Test Cohort)**

- **Total Patients with catheters N=1887**
  - 0-6 days =1253
  - 7 or more days =634
- **Total Patients with Tunneled N =540**
  - 0-6 days =323
  - 7 or more days=217
- **Total Patients with Non-tunneled N =1193**
  - 0-6 days =916
  - 7 or more days=277
- **Total Patients with PICC N=401**
  - 0-6 days =219
  - 7 or more days=182

NT lines are the most common
The greatest number of lines are found in the 0-6 days category
Frequency Analysis

Central Venous Type/Site (Test Cohort)
(Total Patients with specific type/site)

- Tunneled CV mediport N=152
  - 0-6Days= 118 ; 7 or more days= 34
- Tunneled CV NS N=540
  - 0-6Days= 323; 7 or more days= 217
- Non-tunneled IJ N=857
  - 0-6Days=761 ; 7 or more days=96
- Non-tunneled SC N=393
  - 0-6Days=319 ; 7 or more days=74
- Non-tunneled Groin N=244
  - 0-6Days=191 ; 7 or more days=53

Catheter Type Other (Test Cohort) =1097
- 0-6Days= 1043; 7 or more days=54

NT IJ was the most common line
The majority of the lines fell into the 0-6 days subcategory
Univariate Analysis

- All Central Venous Catheters Grouped (Test Cohort):
  - Catheter 0-6 days:
    - HR=3.45
    - P<0.001
  - Catheter equal or >7 days:
    - HR=3.53
    - P<0.001

HR increases with time for all central lines grouped.
Univariate Analysis

**Tunneled (Test Cohort)**

- TCVP 0-6 days
  - HR=1.64 P<0.28

- TCVP equal or >7 days
  - HR=12.92 P<0.001

- Tunneled NS 0-6 days
  - HR=1.87 P<0.02

- Tunneled NS equal or >7 days
  - HR=2.62 P<0.001

TCVP has greatest associated HR
HR increases with time for both tunneled catheter categories
Univariate Analysis

- **Non-Tunneled (Test Cohort)**
  - NT 0-6 days: HR=3.04 P<0.001
  - NT equal or > 7 days: HR=2.82 P<0.001

- **PICC (Test Cohort)**
  - PICC 0-6 days: HR=1.54 P=0.16
  - PICC equal or > 7 days: HR=1.33 P=0.38

HR decreases with time for NT catheters
HR NOT significant for PICC
Univariate Analysis

Catheter Type Other (Test Cohort)

(distinct catheters)

- CTO 0-6 days:
  - HR=1.22
  - P=0.43

- CTO equal or > 7 days:
  - HR=1.46
  - P=0.34

CTO was Not a significant predictor of HR on Univariate Analysis
Univariate Analysis

Non-tunneled Catheter by TYPE and SITE and TIME (Test Cohort):

- IJ 0-6 days: HR=2.15 P<0.001
- IJ equal or > 7 days: HR-2.20 P<0.01
- SC 0-6 days: HR=2.88 P<0.001
- SC equal or > 7 days: HR-1.49 P=0.32 NS
- G/F 0-6 days: HR=1.94 P=.04
- G/F equal or > 7 days: HR-0.95 P=0.22 NS

HR Greatest for SC > G/F > IJ - HR increases with time for IJ
Frequency Analysis

Arterial Catheters (Test Cohort):

- Total Patients With Arterial Lines N=1671
  - Axillary N=72
    - 0-6Days= 41; 7 or more days=32
  - Radial N=593
    - 0-6Days= 463; 7 or more days=130
  - Groin N=188
    - 0-6Days= 150; 7 or more days=38

Radial arterial lines were the most common
The majority of arterial lines fell into the 0-6 days subcategory
Univariate Analysis

- **Total AL (Test Cohort)**
  - 0-6 days: HR=1.54 P<0.09
  - equal or > 7 days: HR=2.24 P<0.001

- **Axillary AL (Test Cohort)**
  - 0-6 days: HR=1.11 P=0.88
  - equal or > 7 days: HR=1.78 P=0.21

- **Radial AL (Test Cohort)**
  - 0-6 days: HR=1.62 P=0.08
  - equal or > 7 days: HR=1.90 P=0.04

- **Groin/Femoral AL (Test Cohort)**
  - 0-6 days: HR=2.38 P<0.02
  - equal or > 7 days: HR=1.00 P=0.89

G/F AL has the greatest associated HR
Total AL has increased HR in GE 7 days category
Frequency/Univariate Analysis

- Urinary Catheter (Test Cohort)
  - N (number of patients) = 4222
  - HR = 2.39
  - P < 0.001

Presence of Urinary Catheter increases HR of NBSI on Univariate Analysis
**Univariate Analysis**

**Cutaneous lesions (Test Cohort)**

- Skin ulceration (pressure ulcer/ulcer)
  - Total Patients with ulcer N=627
  - HR=1.97
  - P<0.001

Skin ulceration increases the HR of NBSI on Univariate Analysis
Summary

Univariate Analysis – increased HR:
- Unconsciousness, DM, immunosuppression
- All medications
- Mechanical ventilation
- Central Venous Catheters
  - Grouped: 0-6D and GE7D categories
  - Tunneled and non-tunneled
  - SC>G/F>IJ in 0-6 D category
- Arterial Catheters
  - All Art.Line GE7D grouped
  - G/F 0-6D category
  - Radial AL GE7D
- Urinary catheter and cutaneous ulceration
Multivariate Model Tree Test Cohort

<table>
<thead>
<tr>
<th>Variable</th>
<th>DF</th>
<th>Parameter Estimate</th>
<th>Standard Error</th>
<th>Chi-Square</th>
<th>Pr &gt; Chi Sq</th>
<th>Hazard Ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td>antimicrobial scat</td>
<td>1</td>
<td>1.52932</td>
<td>0.31897</td>
<td>22.9883</td>
<td>&lt; .0001</td>
<td>4.615</td>
</tr>
<tr>
<td>ctotal fg t0 lt7</td>
<td>1</td>
<td>0.57248</td>
<td>0.24149</td>
<td>5.6201</td>
<td>0.0178</td>
<td>1.773</td>
</tr>
<tr>
<td>ctotal fg e7</td>
<td>1</td>
<td>0.24388</td>
<td>0.25022</td>
<td>0.9500</td>
<td>0.3297</td>
<td>1.276</td>
</tr>
<tr>
<td>TCVP fg e7</td>
<td>1</td>
<td>2.23329</td>
<td>0.37438</td>
<td>35.5840</td>
<td>&lt; .0001</td>
<td>9.330</td>
</tr>
<tr>
<td>antibiotic db scat</td>
<td>1</td>
<td>0.75582</td>
<td>0.19968</td>
<td>14.3277</td>
<td>0.0002</td>
<td>2.129</td>
</tr>
</tbody>
</table>

Multivariate HR: TCVP>Antibiotics>DM> Catheter 0-6> catheter GE 7D

- Variables chosen by clinical relevance, significance on univariate analysis and by SAS step-wise selection
- Catheter GE 7D non-significance likely due to lack of power but retained due to pathophysiologic plausibility
Predictive Model

- Variables from multivariate analysis
- Predictive index
  - ‘Points’ derived from respective HR of Cox Proportional Hazards Model
    - HR is rounded to the nearest integer
  - Based on the cutoff points define:
    - Sensitivity and Specificity
    - Positive Predictive Value (PPV)
# Predictive Model

Predictive Model

Score = 5*antimicrobials_cat + 2*ctotal_fgt0lt7 + 1*ctotal_fge7 + 2*antidbs_cat + 9*TCVP_fge7

<table>
<thead>
<tr>
<th>Score</th>
<th>With Event</th>
<th>Without Event</th>
<th>Cutoff if &gt;</th>
<th>PV+</th>
<th>Sensitivity</th>
<th>Specificity</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>3</td>
<td>8637</td>
<td>0</td>
<td>0.019</td>
<td>0.976</td>
<td>0.577</td>
</tr>
<tr>
<td>1</td>
<td>4</td>
<td>146</td>
<td>1</td>
<td>0.019</td>
<td>0.944</td>
<td>0.587</td>
</tr>
<tr>
<td>2</td>
<td>4</td>
<td>1279</td>
<td>2</td>
<td>0.022</td>
<td>0.911</td>
<td>0.672</td>
</tr>
<tr>
<td>3</td>
<td>2</td>
<td>25</td>
<td>3</td>
<td>0.022</td>
<td>0.895</td>
<td>0.674</td>
</tr>
<tr>
<td>4</td>
<td>0</td>
<td>60</td>
<td>4</td>
<td>0.022</td>
<td>0.895</td>
<td>0.678</td>
</tr>
<tr>
<td>5</td>
<td>18</td>
<td>3103</td>
<td>5</td>
<td>0.051</td>
<td>0.750</td>
<td>0.885</td>
</tr>
<tr>
<td>6</td>
<td>8</td>
<td>197</td>
<td>6</td>
<td>0.053</td>
<td>0.685</td>
<td>0.898</td>
</tr>
<tr>
<td>7</td>
<td>31</td>
<td>1103</td>
<td>7</td>
<td>0.113</td>
<td>0.435</td>
<td>0.972</td>
</tr>
<tr>
<td>8</td>
<td>23</td>
<td>195</td>
<td>8</td>
<td>0.120</td>
<td>0.250</td>
<td>0.985</td>
</tr>
<tr>
<td>9</td>
<td>21</td>
<td>203</td>
<td>9</td>
<td>0.294</td>
<td>0.081</td>
<td>0.998</td>
</tr>
<tr>
<td>10</td>
<td>0</td>
<td>3</td>
<td>10</td>
<td>0.323</td>
<td>0.081</td>
<td>0.999</td>
</tr>
<tr>
<td>12</td>
<td>0</td>
<td>0</td>
<td>12</td>
<td>0.323</td>
<td>0.081</td>
<td>0.999</td>
</tr>
<tr>
<td>15</td>
<td>8</td>
<td>13</td>
<td>15</td>
<td>0.200</td>
<td>0.016</td>
<td>0.999</td>
</tr>
<tr>
<td>17</td>
<td>2</td>
<td>8</td>
<td>17</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**End Point= NBSI:** Score ranges from 0- 17. Positive Trend in PV + with increasing score. 5 Point cut-off to maximize sensitivity and specificity
**Validation of Model**

Score = 5*antimicrobialscat + 2*ctotalfgt0lt7 + 1*ctotalfge7 + 2*antidbscat + 9*TCVPfge7

<table>
<thead>
<tr>
<th>Score</th>
<th>With Event</th>
<th>Without Event</th>
<th>Cutoff if &gt;</th>
<th>PV+</th>
<th>Sensitivity</th>
<th>Specificity</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>6</td>
<td>8565</td>
<td>0</td>
<td>0.018</td>
<td>0.950</td>
<td>0.572</td>
</tr>
<tr>
<td>1</td>
<td>5</td>
<td>156</td>
<td>1</td>
<td>0.017</td>
<td>0.909</td>
<td>0.582</td>
</tr>
<tr>
<td>2</td>
<td>2</td>
<td>1297</td>
<td>2</td>
<td>0.021</td>
<td>0.893</td>
<td>0.669</td>
</tr>
<tr>
<td>3</td>
<td>3</td>
<td>23</td>
<td>3</td>
<td>0.021</td>
<td>0.868</td>
<td>0.670</td>
</tr>
<tr>
<td>4</td>
<td>0</td>
<td>74</td>
<td>4</td>
<td>0.021</td>
<td>0.868</td>
<td>0.675</td>
</tr>
<tr>
<td>5</td>
<td>21</td>
<td>3108</td>
<td>5</td>
<td>0.046</td>
<td>0.694</td>
<td>0.883</td>
</tr>
<tr>
<td>6</td>
<td>10</td>
<td>180</td>
<td>6</td>
<td>0.045</td>
<td>0.612</td>
<td>0.895</td>
</tr>
<tr>
<td>7</td>
<td>34</td>
<td>1153</td>
<td>7</td>
<td>0.086</td>
<td>0.331</td>
<td>0.972</td>
</tr>
<tr>
<td>8</td>
<td>16</td>
<td>178</td>
<td>8</td>
<td>0.088</td>
<td>0.198</td>
<td>0.983</td>
</tr>
<tr>
<td>9</td>
<td>18</td>
<td>207</td>
<td>9</td>
<td>0.125</td>
<td>0.050</td>
<td>0.997</td>
</tr>
<tr>
<td>10</td>
<td>0</td>
<td>9</td>
<td>10</td>
<td>0.154</td>
<td>0.050</td>
<td>0.998</td>
</tr>
<tr>
<td>12</td>
<td>0</td>
<td>4</td>
<td>12</td>
<td>0.171</td>
<td>0.050</td>
<td>0.998</td>
</tr>
<tr>
<td>15</td>
<td>4</td>
<td>19</td>
<td>15</td>
<td>0.167</td>
<td>0.017</td>
<td>0.999</td>
</tr>
<tr>
<td>17</td>
<td>2</td>
<td>10</td>
<td>17</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**End Point= NBSI:** Score ranges from 0-17. Positive Trend in PV + with increasing score. 5 Point cut-off to maximize sensitivity and specificity
Internal Validity: Chi-Square Test

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Test Statistic</th>
<th>P Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>PV +</td>
<td>0.6425</td>
<td>0.422</td>
</tr>
<tr>
<td>Sensitivity</td>
<td>0.9504</td>
<td>0.3296</td>
</tr>
<tr>
<td>Specificity</td>
<td>0.4382</td>
<td>0.5080</td>
</tr>
</tbody>
</table>

Statistical comparison between the predictive model and the validation cohort reveals NO statistically significant differences in PV+, Sensitivity, and Specificity.
Conclusion

- Study goal was to use the EMR to identify patients at high risk for nosocomial BSI using automated systems.
- Data was successfully extracted for 245 NBSI and 29,954 non-BSI patients.
- This cohort was randomly divided in half in order to develop a predictive model for NBSI.
- The predictive model was *internally* validated on the second half of the cohort.
- The model is *preliminary* and must be *prospectively* validated.
Based on Preliminary Model

Computerized screen alerts and automated decision support in real time

- Remind and guide physicians and staff:
  - Promptly remove any intravascular catheter that is not essential. (I A *)
  - For patients with need of IV access
    - Antimicrobial impregnated CVC (I B *)
    - Use a subclavian site for non-tunneled CVC (I A *)
    - Consider PICC or midline catheters (II *)
  - Medications:
    - Decrease excessive use of antibiotics.
    - Physician screen alerts indicating # of specific medication days

**Guidelines for the Prevention of Intravascular Catheter-Related Infections.**

**CID 2002:35, 1281-1307.**