Databases in Clinical Research
Overview

• Background: History and utility of clinical data repositories

• Strategies: Integrating the outcomes tracking database into clinical workflow

• Brigham and Women’s Catheterization Laboratory Database: Architecture, Advances, Limitations

• Examples of Data Exploration: Risk models, “drilling down”, and device safety monitoring
Need for Clinical Data Repositories

• Randomized clinical trials are gold standard for testing a hypothesis, but there are significant limitations:
  
  • generalizability
  
  • timeliness
  
  • cost $$$
Cost of Randomized Clinical Trials

• Estimated cost of RCT:

  Drug Trial: $15,000/patient

• 1000 patient trial: $15MM

• Simply too expensive to answer every relevant clinical question with prospective blinded RCT.
Clinical Registries

• While RCT’s test hypotheses, the real world of clinical practice is a registry.
  
  All patients (generalizability)
  Dynamic (timeliness)

• Significant *Potential* cost savings when automated clinical registry (database system) bundled with other functional requirements

  clinical reporting, billing, inventory control
History of Successful Clinical Registries

- Duke Database
- Washington Heart Center
- Beth Israel Hospital, Boston
- Cleveland Clinic
- Mayo Clinic
- Massachusetts General Hospital
Why Clinical Cardiology?

• High volume clinical sites

• High event rates – death, MI, revascularization, rehospitalization, etc.

• High profile

• High cost to study
Applications of Clinical Databases:

- **Clinical Research:**
  - Retrospective “Hypothesis Generator”
  - Data mining
  - Prospective automated CRF
  - Risk prediction modeling

- **Quality Assurance:**
  - Interprovider variability
  - Benchmark review – ACC NCDR

- **Business and Operations Review** – Turnover times, referral patterns

- **Regulatory Requirements** – State DPH
Overview

• Background: History and utility of clinical data repositories

• Strategies: Integrating the outcomes tracking database into clinical workflow

• Brigham and Women’s Catheterization Laboratory Database: Architecture, Advances, Issues

• Examples of Data Exploration: Risk models, “drilling down”, and device safety monitoring
Strategies for Maintaining Clinical db

Three Strategies:

- Prospective/retrospective off-line chart review
- Data extraction w/ supplemental chart review
- Complete integration into electronic record system
Clinical Database Strategies:

- Parallel Chart Review
  - independent of clinical process
  - focus on data quality
  - maintain current workflow
  - requires team of coders
  - COST $$$

- Hybrid Strategy
- Fully Integrated
  - purely prospective
  - integral part of routine workflow
  - lowest cost (??)
  - data quality issues
  - data management
Integration Dimensions:

- Workflow
- Systems

- Stand Alone Retrospective
- Hybrid Cath Lab Only
- Fully Integrated EMR

Required to achieve "synergy savings"
Multi-Use Function of Clinical Cath Lab Databases:

- Clinical Outcomes Tracking Database:
  - Retrospective Clinical Research
  - Quality Assurance
  - Administrative reporting

- Clinical report generation (structured reporting; transcription templates)

- Technical and Professional Billing

- Inventory Management

- Increased complexity of database with each additional functional layer.
Information Flow Integrated Into Care Process

Care Stage

- Procedure Request
- Pre-Cath Evaluation
- Procedure
- Post-Cath Evaluation
- In-Hospital Follow-up
- Post-DC Follow-up

Output

- Images
- Report

Function

- Order Placement
- Scheduling
- Inventory Control
- Image Distribution
- Report Generation
- Integrated Billing
- Clinical Event Monitoring
Evolutionary Growth in DB Design: BWH CCL DB

- Version 1
- Version 2
- Version 3

# Fields

- Clinical
- Admin
- DB Utility
Functions Supported in Cath Lab:

- Clinical Documentation
- Clinical Outcomes Database (Research)
- Technical Billing
- Professional Billing
- Inventory Management
- Clinical/Quality Improvement Database
- Administrative Database Functionality
- State Reporting (DPH)
DB: Core to Supporting Multiple Functions

- Clinical Documentation
- Technical Billing
- Professional Billing
- Inventory Management
- Clinical/Quality Improvement Database
- Procedure Scheduling
- Administrative Database Functionality
- Image archiving
Tension within Medical Informatics

Database Requirements
-- structure data entry
-- limited vocabulary
-- fixed meaning
-- no free text entry
-- focus on consistency

Clinical Communication Requirements
-- unstructured
-- unlimited vocabulary
-- variable meaning
-- frequent revision
-- focus on interpretation

Clinical Documentation
Overview

• Background: History and utility of clinical data repositories

• Strategies: Integrating the outcomes tracking database into clinical workflow

• Brigham and Women’s Catheterization Laboratory Database: Architecture, Advances, Issues

• Examples of Data Exploration: Risk models, “drilling down”, and device safety monitoring
System Architecture: Phase I

BICS

ICIS

WITT

SQL 7 Transactional DB

Demographics

Report

Report

Demographics
System Architecture: Phase II

- BICS
  - ICIS
    - WITT
      - SQL 7 Transactional DB
        - SQL7 DB Analytical
          - Administrative Users
        - Demographics
          - Report
    - Demographics
      - Report
System Architecture: Phase III

BICS

ICIS

WITT

SQL7 DB

SQL 7

Transactional

DB

Tracking Query
-- vital status
-- d/c date, labs, meds
-- surgery, visits

Administrative

Users

Authorized

Users

Web Server

Report

Demographics

Report

Demographics
System Architecture: Phase IIIb

- **BICS**
  - Demographics
  - Report
  - Tracking Query
    - vital status
    - d/c date, labs, meds
    - surgery, visits

- **ICIS**
  - Demo
  - Report

- **WITT**
  - SQL 7 Transactional DB

- **SQL7 DB Analytical**
  - Administrative Users
  - Authorized Users

- **ACC-NCDR Export Application**
  - Application
  - DPH ACC Submissions
Relational DB Schema: Overview

Key:
- ○ ○: one to one
- ○ -: one to many
- - -: many to many
- □: dictionary table
Relational DB Schema: Lesion Treated
Overview

• Background: History and utility of clinical data repositories

• Strategies: Integrating the outcomes tracking database into clinical workflow

• Brigham and Women’s Catheterization Laboratory Database: Architecture, Advances, Issues

• Examples of Data Exploration: Risk models, “drilling down”, and device safety monitoring
Applications of BWH CCL Database

Risk Prediction Model Development

Drilling Down – Novel Risk Factors

Retrospective Device Safety Analysis

Monthly QA – Cath Lab M&M

Operations Management
Risk Models: Background

• Unadjusted Mortality Rates – Published 1999-2000
  - NY State PTCA Registry Model: 0.9%
  - NNE Cooperative Model: 1.1% 0.6%
  - Holmes et al (Mayo Clinic): 1.6%
  - Moscucci et al (Univ. Michigan): 3.3% 3.4%

• Risk prediction models help adjust for severity of illness
  _ providers assess quality of care – improve process
  _ State / public compare institutions and providers
  _ researchers assess effect of changes in care

See Hannan JAMA 277(11); Holmes Circ 102:517;
Moscucci JACC 34(3); O’Conner JACC 34(3)
Logistic and Score Models for Death

Logistic Regression Model

<table>
<thead>
<tr>
<th></th>
<th>Odds Ratio</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age &gt; 74yrs</td>
<td>2.53</td>
<td>0.01</td>
</tr>
<tr>
<td>B2/C Lesion</td>
<td>1.93</td>
<td>0.08</td>
</tr>
<tr>
<td>Acute MI</td>
<td>1.83</td>
<td>0.20</td>
</tr>
<tr>
<td>Class 3/4 CHF</td>
<td>8.14</td>
<td>0.00</td>
</tr>
<tr>
<td>Left main PCI</td>
<td>6.59</td>
<td>0.02</td>
</tr>
<tr>
<td>Stent Use</td>
<td>0.50</td>
<td>0.08</td>
</tr>
<tr>
<td>Cardiogenic Shock</td>
<td>8.33</td>
<td>0.00</td>
</tr>
<tr>
<td>Unstable Angina</td>
<td>1.69</td>
<td>0.17</td>
</tr>
<tr>
<td>Tachycardic</td>
<td>2.77</td>
<td>0.04</td>
</tr>
<tr>
<td>Chronic Renal Insuf.</td>
<td>2.71</td>
<td>0.05</td>
</tr>
</tbody>
</table>
# Logistic and Score Models for Death

## Logistic Regression Model

<table>
<thead>
<tr>
<th>Condition</th>
<th>Odds Ratio</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age &gt; 74yrs</td>
<td>2.53</td>
<td>0.01</td>
</tr>
<tr>
<td>B2/C Lesion</td>
<td>1.93</td>
<td>0.08</td>
</tr>
<tr>
<td>Acute MI</td>
<td>1.83</td>
<td>0.20</td>
</tr>
<tr>
<td>Class 3/4 CHF</td>
<td>8.14</td>
<td>0.00</td>
</tr>
<tr>
<td>Left main PCI</td>
<td>6.59</td>
<td>0.02</td>
</tr>
<tr>
<td>Stent Use</td>
<td>0.50</td>
<td>0.08</td>
</tr>
<tr>
<td>Cardiogenic Shock</td>
<td>8.33</td>
<td>0.00</td>
</tr>
<tr>
<td>Unstable Angina</td>
<td>1.69</td>
<td>0.17</td>
</tr>
<tr>
<td>Tachycardic</td>
<td>2.77</td>
<td>0.04</td>
</tr>
<tr>
<td>Chronic Renal Insuf.</td>
<td>2.71</td>
<td>0.05</td>
</tr>
</tbody>
</table>
## Logistic and Score Models for Death

**Logistic Regression Model**

<table>
<thead>
<tr>
<th>Factor</th>
<th>Odds Ratio</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age &gt; 74yrs</td>
<td>2.53</td>
<td>0.01</td>
</tr>
<tr>
<td>B2/C Lesion</td>
<td>1.93</td>
<td>0.08</td>
</tr>
<tr>
<td>Acute MI</td>
<td>1.83</td>
<td>0.20</td>
</tr>
<tr>
<td>Class 3/4 CHF</td>
<td>8.14</td>
<td>0.00</td>
</tr>
<tr>
<td>Left main PCI</td>
<td>6.59</td>
<td>0.02</td>
</tr>
<tr>
<td><strong>Stent Use</strong></td>
<td>0.50</td>
<td>0.08</td>
</tr>
<tr>
<td>Cardiogenic Shock</td>
<td>8.33</td>
<td>0.00</td>
</tr>
<tr>
<td>Unstable Angina</td>
<td>1.69</td>
<td>0.17</td>
</tr>
<tr>
<td>Tachycardic</td>
<td>2.77</td>
<td>0.04</td>
</tr>
<tr>
<td>Chronic Renal Insuf.</td>
<td>2.71</td>
<td>0.05</td>
</tr>
</tbody>
</table>

---

**Notes:**
- Logistic regression model for predicting death risk.
- Odds ratios indicate the increased risk of death associated with each factor.
- p-values indicate statistical significance of each factor.
Logistic and Score Models for Death

### Logistic Regression Model

<table>
<thead>
<tr>
<th>Condition</th>
<th>Odds Ratio</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age &gt; 74yrs</td>
<td>2.53</td>
<td>0.01</td>
</tr>
<tr>
<td>B2/C Lesion</td>
<td>1.93</td>
<td>0.08</td>
</tr>
<tr>
<td>Acute MI</td>
<td>1.83</td>
<td>0.20</td>
</tr>
<tr>
<td>Class 3/4 CHF</td>
<td>8.14</td>
<td>0.00</td>
</tr>
<tr>
<td>Left main PCI</td>
<td>6.59</td>
<td>0.02</td>
</tr>
<tr>
<td><strong>Stent Use</strong></td>
<td><strong>0.50</strong></td>
<td><strong>0.08</strong></td>
</tr>
<tr>
<td>Cardiogenic Shock</td>
<td>8.33</td>
<td>0.00</td>
</tr>
<tr>
<td>Unstable Angina</td>
<td>1.69</td>
<td>0.17</td>
</tr>
<tr>
<td>Tachycardic</td>
<td>2.77</td>
<td>0.04</td>
</tr>
<tr>
<td>Chronic Renal Insuf.</td>
<td>2.71</td>
<td>0.05</td>
</tr>
</tbody>
</table>

### Risk Score Model

<table>
<thead>
<tr>
<th>Beta coeff</th>
<th>Risk value</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.927</td>
<td>2</td>
</tr>
<tr>
<td>0.659</td>
<td>1</td>
</tr>
<tr>
<td>0.601</td>
<td>1</td>
</tr>
<tr>
<td>2.097</td>
<td>4</td>
</tr>
<tr>
<td>1.886</td>
<td>3</td>
</tr>
<tr>
<td>-0.683</td>
<td>-1</td>
</tr>
<tr>
<td>2.120</td>
<td>4</td>
</tr>
<tr>
<td>0.522</td>
<td>1</td>
</tr>
<tr>
<td>1.020</td>
<td>2</td>
</tr>
<tr>
<td>0.996</td>
<td>2</td>
</tr>
</tbody>
</table>
ROC Curves: Death Models
Validation Set: 1460 Cases

Artificial Neural Networks are non-linear mathematical models which incorporate a layer of hidden “nodes” connected to the input layer (covariates) and the output.
Artificial Neural Networks are non-linear mathematical models which incorporate a layer of hidden “nodes” connected to the input layer (covariates) and the output.
Artificial Neural Networks are non-linear mathematical models which incorporate a layer of hidden “nodes” connected to the input layer (covariates) and the output.
ROC Curves: Death Models
Validation Set: 1460 Cases

Risk Score of Death: BWH Experience
Unadjusted Overall Mortality Rate = 2.1%

Applications of BWH CCL Database

Risk Prediction Model Development

Drilling Down – Novel Risk Factors

Retrospective Device Safety Analysis

Monthly QA – Cath Lab M&M

Operations Management
# MACE Models: Impact of No-Reflow

<table>
<thead>
<tr>
<th>Condition</th>
<th>Odds Ratio</th>
<th>p-value</th>
<th>Beta Coefficient</th>
<th>Risk Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age &gt; 74yrs</td>
<td>1.40</td>
<td>0.16</td>
<td>0.337</td>
<td>0</td>
</tr>
<tr>
<td>B2/C Lesion</td>
<td>2.56</td>
<td>0.00</td>
<td>0.939</td>
<td>2</td>
</tr>
<tr>
<td>Acute MI</td>
<td>2.99</td>
<td>0.00</td>
<td>1.096</td>
<td>2</td>
</tr>
<tr>
<td>Class 3/4 CHF</td>
<td>3.61</td>
<td>0.00</td>
<td>1.283</td>
<td>3</td>
</tr>
<tr>
<td>Left main PCI</td>
<td>2.30</td>
<td>0.28</td>
<td>0.831</td>
<td>2</td>
</tr>
<tr>
<td>Stent Use</td>
<td>0.58</td>
<td>0.03</td>
<td>-0.539</td>
<td>-1</td>
</tr>
<tr>
<td>Cardiogenic Shock</td>
<td>3.33</td>
<td>0.01</td>
<td>1.202</td>
<td>3</td>
</tr>
<tr>
<td>USA</td>
<td>2.69</td>
<td>0.00</td>
<td>0.989</td>
<td>2</td>
</tr>
<tr>
<td>Tachycardic</td>
<td>1.36</td>
<td>0.44</td>
<td>0.311</td>
<td>0</td>
</tr>
<tr>
<td>No Reflow</td>
<td>2.90</td>
<td>0.01</td>
<td>1.044</td>
<td>2</td>
</tr>
<tr>
<td>Unscheduled</td>
<td>1.49</td>
<td>0.08</td>
<td>0.396</td>
<td>0</td>
</tr>
<tr>
<td>Chronic Renal Insuff.</td>
<td>1.58</td>
<td>0.23</td>
<td>0.457</td>
<td>1</td>
</tr>
</tbody>
</table>
63yo male 4yrs s/p 4v CABG.

Presents with NQWMI W/ lateral ST depress
Posis Angiojet: Rheolytic Thrombectomy
Direct Stenting of Culprit Lesion
Risk of In-Hospital Complication

TIMI Flow Rates Improved Significantly

Lack of Effective Treatment: BWH Experience

Risk of Death or Myocardial Infarction

Applications of BWH CCL Database

- Risk Prediction Model Development
- Drilling Down – Novel Risk Factors
- Retrospective Device Safety Analysis
- Monthly QA – Cath Lab M&M
- Operations Management
Patients receiving a closure device experienced a 44% reduction in vascular complications.

This effect was preserved in those patients receiving gp 2b3a inhibitors.

Applications of BWH CCL Database

Risk Prediction Model Development

Drilling Down – Novel Risk Factors

Retrospective Device Safety Analysis

Monthly QA – Cath Lab M&M

Operations Management
Coronary Procedures by Month

<table>
<thead>
<tr>
<th>Month</th>
<th>Diagnostic</th>
<th>Intervention</th>
<th>Total Cases</th>
</tr>
</thead>
<tbody>
<tr>
<td>Aug</td>
<td>375</td>
<td>189</td>
<td>384</td>
</tr>
<tr>
<td>Sept</td>
<td>330</td>
<td>165</td>
<td>342</td>
</tr>
<tr>
<td>Oct</td>
<td>364</td>
<td>159</td>
<td>371</td>
</tr>
<tr>
<td>Nov</td>
<td>346</td>
<td>159</td>
<td>356</td>
</tr>
<tr>
<td>Dec</td>
<td>294</td>
<td>159</td>
<td>296</td>
</tr>
<tr>
<td>Jan</td>
<td>452</td>
<td>198</td>
<td>458</td>
</tr>
<tr>
<td>Feb</td>
<td>333</td>
<td>148</td>
<td>337</td>
</tr>
<tr>
<td>Mar</td>
<td>383</td>
<td>167</td>
<td>388</td>
</tr>
<tr>
<td>Apr</td>
<td>433</td>
<td>185</td>
<td>437</td>
</tr>
<tr>
<td>May</td>
<td>428</td>
<td>185</td>
<td>436</td>
</tr>
<tr>
<td>June</td>
<td>452</td>
<td>180</td>
<td>458</td>
</tr>
<tr>
<td>July</td>
<td>447</td>
<td>202</td>
<td>456</td>
</tr>
</tbody>
</table>
## Planned vs. Ad Hoc PCI

<table>
<thead>
<tr>
<th></th>
<th>Jan</th>
<th>Feb</th>
<th>Mar</th>
<th>April</th>
<th>May</th>
<th>June</th>
<th>July</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ad Hoc PCI</td>
<td>125</td>
<td>103</td>
<td>118</td>
<td>115</td>
<td>106</td>
<td>121</td>
<td>132</td>
</tr>
<tr>
<td>Planned PCI</td>
<td>70</td>
<td>78</td>
<td>52</td>
<td>67</td>
<td>78</td>
<td>50</td>
<td>66</td>
</tr>
</tbody>
</table>

**Ad Hoc PCI rate**
- January: 33%
- February: 40%
- March: 36%
- April: 31%
- May: 30%
- June: 30%
- July: 35%
Internal vs. External MD Volume

<table>
<thead>
<tr>
<th>Month</th>
<th>Internal</th>
<th>External</th>
</tr>
</thead>
<tbody>
<tr>
<td>Aug</td>
<td>77%</td>
<td>23%</td>
</tr>
<tr>
<td>Sep</td>
<td>76%</td>
<td>24%</td>
</tr>
<tr>
<td>Oct</td>
<td>75%</td>
<td>25%</td>
</tr>
<tr>
<td>Nov</td>
<td>73%</td>
<td>27%</td>
</tr>
<tr>
<td>Dec</td>
<td>69%</td>
<td>31%</td>
</tr>
<tr>
<td>Jan</td>
<td>66%</td>
<td>34%</td>
</tr>
<tr>
<td>Feb</td>
<td>73%</td>
<td>27%</td>
</tr>
<tr>
<td>Mar</td>
<td>77%</td>
<td>29%</td>
</tr>
<tr>
<td>Apr</td>
<td>71%</td>
<td>29%</td>
</tr>
<tr>
<td>May</td>
<td>74%</td>
<td>26%</td>
</tr>
<tr>
<td>June</td>
<td>72%</td>
<td>28%</td>
</tr>
<tr>
<td>July</td>
<td>72%</td>
<td>28%</td>
</tr>
</tbody>
</table>
Post-Procedural Events for July, 2002

• Significant events *reported* during July, 2002:

<table>
<thead>
<tr>
<th>Category</th>
<th>Count</th>
</tr>
</thead>
<tbody>
<tr>
<td>Death</td>
<td>3</td>
</tr>
<tr>
<td>Stroke</td>
<td>1</td>
</tr>
<tr>
<td>CABG</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>(perforation of LCx)</td>
</tr>
<tr>
<td>MI*</td>
<td>6</td>
</tr>
<tr>
<td></td>
<td>(1 SAT)</td>
</tr>
<tr>
<td>TVR</td>
<td>2</td>
</tr>
<tr>
<td>Vascular</td>
<td>7</td>
</tr>
<tr>
<td></td>
<td>(1 transfusion reported, 1 PSA req. surg)</td>
</tr>
<tr>
<td>Renal</td>
<td>3</td>
</tr>
<tr>
<td>CHF</td>
<td>1</td>
</tr>
</tbody>
</table>

* MI defined as total CK-MB>3x ULN in patient w/o index AMI.
Applications of BWH CCL Database

Risk Prediction Model Development

Drilling Down – Novel Risk Factors

Retrospective Device Safety Analysis

Monthly QA – Cath Lab M&M

Operations Management
One-third of total case volume is noted as inpatient source in WITT.
The case volume is distributed according the usual 80/20 rule. Nearly 80% of cases is referred from 20% of the MDs.
Rules for Designing an Outcomes Database

• Understand workflow in detail. Identify immutable points (most of these depend on perspective).

• Incremental design – identify successful milestones

• Open architecture – use ODBC compliant relational databases as backbone

  Systems integration is most complex challenge

• Goal of distributed information availability.

• Identify implementation team. Responsibilities, project plan, regular operational meeting.
Acknowledgements

Cardiovascular Division

Jeff Popma, MD
Andrew Selwyn, MD
Campbell Rogers, MD
Charles Lin, MBA
Benjamin Paul

Decision Systems Group

Lucila Ohno-Machado, MD PhD
Robert Greenes, MD PhD
Aziz Boxwala, MD PhD

Partners Information Systems

Mark Nightingale
Thank You!