A Hybrid System That Supports Public Reporting in Pennsylvania

RS Johannes, MD, MS
Cardinal Health
June 19, 2007
Serendipity & History

• In the mid 1980s, three people met
  – Dr. Allen Brewster, MediQual Founder
  – Dr. Donald Fetterolf, Highmark BC/BS
  – Ernie Sessa, NAHDO founder, 1st Exec Dir PHC4
    • They all shared a single belief, namely that the precision and face validity of risk adjustment depended upon use of clinical data beyond claims data

• Pennsylvania Health Care Cost Containment Council
  – Only state to perform uninterrupted annual public reporting of hospital performance data for 20 years
  – Pennsylvania publicly reports 50+ diseases
  – Cardinal Health has provided the data collection and risk adjustment methods over the entire period.
Four important themes

• Timing of clinical data is important to admission based severity stratification

• Laboratory data is both objective and powerful as a predictor

• Laboratory data is electronically available

• The face value of clinical data should not be underestimated
Clinical data and timing

• Provides for models that can better identify risk in the peri-admission period.
  – Frees the models from the criticism that late hospital stay events are used for adjustment
  – Better separation of comorbidities from complications
  – Position upheld by other published studies

• Is Admission (POA) coding the solution?
  – In order to add a POA flag, the code must first be present
  – How might this work in a situation such as hyponatremia?
Can ICD9 codes detect abnormal labs?

<table>
<thead>
<tr>
<th>Laboratory Abnormality</th>
<th>Sensitivity Admission Period</th>
<th>% Improvement from Full Hospital Stay</th>
</tr>
</thead>
<tbody>
<tr>
<td>Low serum sodium (&lt; 135 mEq/L)</td>
<td>11.8</td>
<td>2.1</td>
</tr>
<tr>
<td>High serum sodium (&gt; 145 mEq/L)</td>
<td>12.3</td>
<td>2.5</td>
</tr>
<tr>
<td>Low serum potassium (&lt; 3.5 mEq/L)</td>
<td>22.8</td>
<td>4.2</td>
</tr>
<tr>
<td>High serum potassium (&gt; 5.0 mEq/L)</td>
<td>18.9</td>
<td>3.9</td>
</tr>
<tr>
<td>Low serum hemoglobin (&lt; 12 g/dl (females), &lt; 14 g/dl (males))</td>
<td>9.6</td>
<td>0.6</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Laboratory Abnormality</th>
<th>Sensitivity of Claims Data (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Low serum sodium recorded &gt;10 times (≤ 135 meq/L)</td>
<td>30.0</td>
</tr>
<tr>
<td>High serum sodium recorded &gt;10 times (&gt;145 meq/L)</td>
<td>42.2</td>
</tr>
<tr>
<td>Low serum potassium recorded &gt;10 times (&lt; 3.5 meq/L)</td>
<td>19.5</td>
</tr>
<tr>
<td>High serum potassium recorded &gt;10 times (&gt; 5.0 meq/L)</td>
<td>20.4</td>
</tr>
<tr>
<td>Low serum hemoglobin Hgb recorded &gt;10 times (&lt;12 g/L for females or &lt;14 g/L for males)</td>
<td>9.5</td>
</tr>
</tbody>
</table>

Source: Cardinal Health Response to CMS 1488-P
© 2007 Cardinal Health, Inc.or one of its subsidiaries. All rights reserved.
# Lab data used for risk adjustment

<table>
<thead>
<tr>
<th>Hematology &amp; Coagulation</th>
<th>WBC</th>
<th>Hemoglobin</th>
</tr>
</thead>
<tbody>
<tr>
<td>% Bands</td>
<td>Platelets</td>
<td></td>
</tr>
<tr>
<td>PT &amp; INR</td>
<td>PTT</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Chemistry</th>
<th>Na</th>
<th>K</th>
</tr>
</thead>
<tbody>
<tr>
<td>BUN</td>
<td>Creatinine</td>
<td></td>
</tr>
<tr>
<td>Albumin</td>
<td>Ca</td>
<td></td>
</tr>
<tr>
<td>Total Bilirubin</td>
<td>AST</td>
<td></td>
</tr>
<tr>
<td>Alkaline Phosphatase</td>
<td>Glucose</td>
<td></td>
</tr>
<tr>
<td>Cardiac enzymes</td>
<td>BNP</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Blood Gases</th>
<th>pO₂</th>
<th>pCO₂</th>
</tr>
</thead>
<tbody>
<tr>
<td>pH</td>
<td>HCO₃</td>
<td></td>
</tr>
<tr>
<td>Base Excess</td>
<td>FiO₂</td>
<td></td>
</tr>
</tbody>
</table>
Lab values enable gradation of risk

Distribution of Albumin Levels and Mortality Rates

Source: CMS response 1488-P, June 2006
The Dimensions of Risk

Manual

Key Clinical Findings
Vital Signs
Laboratory

Automated

Claims

Additional Clinical Data
KCF: physical exam, imaging
Lab: BUN, Na, K, Albumin
Vital Signs: TPR &BP

Claims Data
ICD-9 diagnosis and procedure codes, patient demographics (age, gender, billing charges, disposition, etc.)
Recent studies: AHRQ

• AHRQ has a long standing interest in the risk adjustment process
  – When Dr. Mark McClellan was still at Stanford University, he led an AHRQ supported study to examine the merits of various methods exclusively relying on administrative data
  – This study quite unambiguously argued in favor of 3M APR-DRGs as the method of choice when using administrative data

• In 2003, PHC4 and MediQual began discussions with AHRQ
  – In 2005, a competitive contract was let to Abt Associates with Michael Pine as a collaborator
  – Results of this work were first presented at NAHDO in December 2006
  – First publications appeared in The American Surgeon and JAMA in December 2006 and January 2007
What did AHRQ / Abt Assoc do?

- They examined 3 years (2000-2003) of data provided to them by PHC4/MediQual and:
  1. Established an advisory panel with AHRQ
  2. Tested the question: *Once risk-adjustment models have utilized claims data maximally, is there any added benefit to clinical data?* Also addressed cost/benefit ratio.
  3. Studied 8 conditions (MI, CHF, Stroke, GI bleed, pneumonia, AAA repair, CABG, craniotomy)
  4. Developed a family of models (AgeOnly, Admin, Admin+POA, [Admin+POA]+Lab + [Admin+POA+Lab]+VS, [Admin+POA+Lab+VS]+KCFs)
Differences in IQI Hospital-Level Bias

Less bias with increasing amounts of clinical data

Upper Bound for Bias in Standard Deviations

- No Adjustment
- AgeOnly
- ADM
- ADM+POA
- + LAB
- + VS
- + KCFs

Percent Exceeding Upper Bound

> 0.5  > 1.0  > 1.5  > 2.0  > 2.5  > 3.0

Used with permission: Michael Pine & Assoc 2006
© 2007 Cardinal Health, Inc.or one of its subsidiaries. All rights reserved.
Recent studies: Cardinal Health

- Examined 3 years of data from 2000-2003 but divided the population into those hospitals that collected data electronically from those that did not.
- Tested the question: What are the relative merits of the types of data used for risk adjustment?
- Studied 6 conditions: (ischemic and hemorrhagic stroke, pneumonia, MI, CHF, and septicemia)
- This study has been accepted to Medical Care (Using Automated Clinical Data for Risk Adjustment: Development and Validation of Six Disease-Specific Mortality Predictive Models for Pay-For-Performance. Ying P Tabak, RS Johannes, Jeffrey H Silber)
# Results from the Cardinal Health Study

## Relative Contribution of Laboratory Variables in Relationship to Other Variables

<table>
<thead>
<tr>
<th>Disease Group</th>
<th>Lab / Age</th>
<th>Lab / ICD-9 Var.</th>
<th>Lab / VS</th>
<th>Lab / AMS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ischemic Stroke</td>
<td>1.61 (1.28-2.04)</td>
<td>7.02 (5.12-9.62)</td>
<td>3.18 (2.14-4.75)</td>
<td>0.33 (0.26-0.41)</td>
</tr>
<tr>
<td>Hem. Stroke</td>
<td>1.76 (1.45-2.14)</td>
<td>2.26 (1.70-2.99)</td>
<td>3.14 (2.29-4.29)</td>
<td>0.37 (0.29-0.48)</td>
</tr>
<tr>
<td>Pneumonia</td>
<td>1.00 (0.89-1.12)</td>
<td>3.59 (3.01-4.28)</td>
<td>2.52 (2.11-3.00)</td>
<td>4.60 (3.79-5.59)</td>
</tr>
<tr>
<td>AMI</td>
<td>0.96 (0.87-1.05)</td>
<td>67.4 (51.6-87.9)</td>
<td>2.65 (2.27-3.09)</td>
<td>3.38 (2.93-3.90)</td>
</tr>
<tr>
<td>CHF</td>
<td>2.35 (2.07-2.66)</td>
<td>14.1 (10.3-19.5)</td>
<td>2.93 (2.57-3.33)</td>
<td>4.01 (3.46-4.65)</td>
</tr>
<tr>
<td>Septicemia</td>
<td>2.75 (2.39-3.15)</td>
<td>8.03 (6.09-10.6)</td>
<td>2.53 (2.14-2.98)</td>
<td>7.66 (5.86-10.0)</td>
</tr>
</tbody>
</table>

Results are shown as Omega (ω) statistics with 95% confidence intervals.

All save Lab/Age for Pneumonia & AMI are p < 0.0001

Source: Medical Care (in press)
Evaluation of Critical Organ Systems

• Where are laboratory results robust?
  – Kidney: BUN, creatinine, Sodium, Potassium
  – Liver: alk phos, bilirubin, albumin, AST
  – Lung: Arterial blood gases, $O_2$ saturation
  – Heme: Hbg, WBC, bands, platelets, proTime, PTT
  – Endocrine: glucose, Sodium, Potassium

• Where might lab values use some of help?
  – Heart: BNP, Cardiac enzymes
  – Brain: WBC, CPK
### The Dimensions of Risk: Example

#### Congestive Heart Failure Model
- **2004-2005**
  - 120,745 cases
  - 4,377 deaths
- Mortality rate = 3.6%
- c-statistic = 0.802

#### Key Clinical Findings

#### Vital Signs

#### Laboratory

#### Claims

#### Available Electronically

#### Claims

#### Laboratory

<table>
<thead>
<tr>
<th>Variable</th>
<th>Odds Ratio</th>
<th>Lower 95CL</th>
<th>Upper 95CL</th>
</tr>
</thead>
<tbody>
<tr>
<td>Moderate or Severe Altered Mental Status</td>
<td>5.18</td>
<td>4.56</td>
<td>5.89</td>
</tr>
<tr>
<td>Mild Altered Mental Status</td>
<td>1.23</td>
<td>1.13</td>
<td>1.35</td>
</tr>
<tr>
<td>Systolic BP mm Hg ≤ 80</td>
<td>2.39</td>
<td>2.15</td>
<td>2.65</td>
</tr>
<tr>
<td>Systolic BP mm Hg (81-100)</td>
<td>1.73</td>
<td>1.59</td>
<td>1.88</td>
</tr>
<tr>
<td>Diastolic BP mm Hg ≤ 53</td>
<td>1.68</td>
<td>1.51</td>
<td>1.86</td>
</tr>
<tr>
<td>Diastolic BP mm Hg (54-62)</td>
<td>1.23</td>
<td>1.13</td>
<td>1.35</td>
</tr>
<tr>
<td>Oral Temp F ≤ 95 or Oral Temp F &gt; 100</td>
<td>1.32</td>
<td>1.19</td>
<td>1.46</td>
</tr>
<tr>
<td>Pulse (100-119)</td>
<td>1.46</td>
<td>1.34</td>
<td>1.58</td>
</tr>
<tr>
<td>Pulse ≥ 120</td>
<td>1.55</td>
<td>1.41</td>
<td>1.70</td>
</tr>
<tr>
<td>Albumin g/dL ≤ 2.4</td>
<td>2.12</td>
<td>1.85</td>
<td>2.44</td>
</tr>
<tr>
<td>Albumin g/dL (2.5 - 2.7)</td>
<td>1.64</td>
<td>1.43</td>
<td>1.89</td>
</tr>
<tr>
<td>&lt;CPK U/L ≤ 35 or CPK U/L&gt;500</td>
<td>1.22</td>
<td>1.13</td>
<td>1.32</td>
</tr>
<tr>
<td>Na meq/L ≤ 130</td>
<td>1.64</td>
<td>1.50</td>
<td>1.87</td>
</tr>
<tr>
<td>Na meq/L (131 - 135)</td>
<td>1.22</td>
<td>1.12</td>
<td>1.32</td>
</tr>
<tr>
<td>Na &gt; 145</td>
<td>1.46</td>
<td>1.26</td>
<td>1.69</td>
</tr>
<tr>
<td>BUN mg/dL (35-50)</td>
<td>1.66</td>
<td>1.53</td>
<td>1.81</td>
</tr>
<tr>
<td>BUN mg/dL (51 - 70)</td>
<td>2.22</td>
<td>2.02</td>
<td>2.45</td>
</tr>
<tr>
<td>BUN mg/dL &gt;70</td>
<td>3.53</td>
<td>3.19</td>
<td>3.91</td>
</tr>
<tr>
<td>pH ≤ 7.25</td>
<td>1.78</td>
<td>1.53</td>
<td>2.08</td>
</tr>
<tr>
<td>pH (7.26 - 7.33)</td>
<td>1.67</td>
<td>1.45</td>
<td>1.93</td>
</tr>
<tr>
<td>pCO2 Arterial ≤ 35 or pCO2 Arterial&gt;60</td>
<td>1.57</td>
<td>1.42</td>
<td>1.74</td>
</tr>
<tr>
<td>Composite Trop I &gt; 1 mg/ml or CPK-MB &gt; 9 IU</td>
<td>1.57</td>
<td>1.40</td>
<td>1.77</td>
</tr>
<tr>
<td>Anticoagulant PT/INR</td>
<td>1.38</td>
<td>1.28</td>
<td>1.48</td>
</tr>
<tr>
<td>Total bilirubin mg/dL &gt; 1.4</td>
<td>1.47</td>
<td>1.31</td>
<td>1.65</td>
</tr>
<tr>
<td>WBC &gt; 10.9</td>
<td>1.44</td>
<td>1.34</td>
<td>1.55</td>
</tr>
<tr>
<td>BNP ≤ 100 or proBNP ≤ 1000</td>
<td>0.39</td>
<td>0.23</td>
<td>0.66</td>
</tr>
<tr>
<td>BNP &gt; 2400 or proBNP &gt; 18000</td>
<td>1.37</td>
<td>1.21</td>
<td>1.55</td>
</tr>
<tr>
<td>Years older than 45</td>
<td>1.03</td>
<td>1.03</td>
<td>1.04</td>
</tr>
<tr>
<td>CHRONIC OBST AIRWAY DIS</td>
<td>1.17</td>
<td>1.09</td>
<td>1.25</td>
</tr>
<tr>
<td>AML SUBSEQUENT</td>
<td>2.33</td>
<td>1.53</td>
<td>3.54</td>
</tr>
<tr>
<td>CHRONIC PULMONARY HEART DIS</td>
<td>1.13</td>
<td>1.03</td>
<td>1.24</td>
</tr>
<tr>
<td>METASTATIC CANCER</td>
<td>2.08</td>
<td>1.81</td>
<td>2.40</td>
</tr>
</tbody>
</table>
MD acceptance: Enhanced face validity

Box 1: Common themes about what makes data feedback effective in the hospital setting

- Theme 1: Data must be perceived by physicians as valid to motivate change.
- Theme 2: It takes time to develop the credibility of data within a hospital.
- Theme 3: The value of feedback depends on data quality.
- Theme 4: Feedback must be aligned with local processes.

Data feedback efforts in quality improvement: lessons learned from US hospitals

Opinions on provider profiling: Telephone survey of stakeholders

Comparison of Administrative Data and Medical Records to Measure the Quality of Medical Care Provided to Vulnerable Older Patients

Table 3

1. Inherent problems with the use of billing and administrative databases for profiling.
2. Lack of time for physicians to review profiles and use the information they provide.
3. Missing data and inability to gather needed information (diagnoses, laboratory test values, physical findings, nonprescription drug use, herbal product use, etc.)
4. Performance scores based on administrative data may vary greatly over time.
5. Performance scores based on administrative data alone may not be as high as those obtained from medical records.
6. Higher performance scores based on administrative data alone may not reflect the true quality of care.
7. Performance scores based on administrative data alone may not reflect the true quality of care.
8. Performance scores based on administrative data alone may not reflect the true quality of care.
9. Performance scores based on administrative data alone may not reflect the true quality of care.
10. Performance scores based on administrative data alone may not reflect the true quality of care.
Advantages of clinical data

- Objective
- Precise
- Time-stamped
- Suffers from few missing data
- Not susceptible to being gamed
- Verifiable in medical literature
- Accepted by clinicians
- Opportunity for automated data collection
Thank You for Your Attention

Cardinal Health
Working together. For life.℠
SIRS Study Results

Cases in the ATLAS™ database with discharges during 2003 passing hospital-level and patient-level edits (n=1,009,753)

2003 cases with discharges after June 30, 2003 (to allow at least 6 months for the new SIRS ICD-9CM codes to be used). (n=490,163)

Cases that had a blood culture positive for bacteria during admission (DOS -1 → DOS 4) from hospitals that collected 5 days of KCF data and had at least one KCF collected (n=1,392)

Cases that met at least 2 SIRS criteria on the same day as the first positive blood culture (n=1,195)

Cases that met at least 2 SIRS criteria and had primary or secondary dx codes of 995.91 or 995.92 (n=43)

ICD-9 codes identify SIRS 39/915 = 4.3 %

Lab values miss ICD-9 code: 1/1,025 = 0.10 %

# of cases that had + blood culture and SIRS Dx, but no SIRS criteria based on lab values (n=1)

# of cases that had + blood culture and SIRS Dx, but no SIRS criteria based on lab values (n=3)

# of cases that met at least 2 SIRS criteria on the same day as the first positive blood culture (exclusions eliminated) (n=915)

# of cases that met at least 2 SIRS criteria on the same day as the first positive blood culture (n=110)

# of cases that did not meet at least 2 SIRS criteria on the same data as first positive blood culture (n=110)

# of cases that did not meet at least 2 SIRS criteria on the same data as first positive blood culture (n=197)

Evaluate Principal Dx codes for exclusions

Exclusions eliminated All Prin. DX codes included

All Principal dx codes included (n=1,392)

Lab values miss ICD-9 code: 3/1,392 = 0.21 %

ICD-9 codes identify SIRS 43/1,195 = 3.6 %

Lab values miss ICD-9 code: 1/1,025 = 0.10 %

# of cases that met at least 2 SIRS criteria on the same day as the first positive blood culture (exclusions eliminated) (n=915)

# of cases that met at least 2 SIRS criteria on the same day as the first positive blood culture (n=110)

# of cases that did not meet at least 2 SIRS criteria on the same data as first positive blood culture (n=110)

# of cases that did not meet at least 2 SIRS criteria on the same data as first positive blood culture (n=197)
Differences in IQI Hospital-Level Bias

Percent Exceeding Upper Bound

Upper Bound for Bias in Standard Deviations

- No Adjustment
- Level 0
- Level 1
- Level 1-POA
- Level 1-DRG
- Level 2
- Level 3
- Level 4

Used with permission: Michael Pine & Assoc 2006

© 2007 Cardinal Health, Inc. or one of its subsidiaries. All rights reserved.