Data Quality in Healthcare
Comparative Databases

Steve Meurer PhD, MBA/MHS
Vice President, Clinical Data & Informatics

University HealthSystem Consortium

- A member owned and governed consortium of academic medical centers
  - This relationship is what makes us unique
  - Approximately 90% of all major not for profit academic medical centers are UHC members
  - Affiliate hospitals are welcome and increasing in numbers (we currently have over 150 associate member hospitals)
  - Nearly 140 members and affiliates subscribe to the CDB
- UHC began in 1984, and has had only 2 CEOs
- UHC provides comparative databases, associated services, a Group Purchasing Organization, and networking opportunities
“Healthcare’s single most important issue is its inability to improve”

Don Berwick

Reasons for this are many, but a major hurdle is that very little quality data is perfect

HOWEVER, Imperfect data can be very useful in providing direction for improvement efforts ... only if you understand the imperfections
R2 x I3 = Change

Using data to tell a story / motivate improvement

- Relationships
- Resources
- Information
- Incentives
- Innovation

1. Is the data accurate?
2. Do you have appropriate comparisons / targets?
3. Is the data adjusted properly?
4. Do you have the necessary data?
5. Is the data analyzed correctly?
6. Is the data presented correctly (both in print and word)?

Source/Scope of CDB Data

Scope
- Inpatient Discharges
- Outpatient (Currently in R&D) will include ED, observation, chemo/rad therapies, and selected ambulatory procedures
- Three years of rolling data available online

Source
- CPDF – data feed for both CDB and CRM (line item detail)
- Monthly submission
Data Quality

1. Does the data smell or look fishy?
   1. UHC has developed an automated process that examines member data and spits out data quality reports
      1. These reports will look at all variables and ask whether they are within a target range
      2. If a variable is not within the target and does not effect overall statistics, the data still passes
      3. If a variable is not within the target and effects overall statistics, the data is returned to the member to be fixed

2. Is the data an accurate reflection of clinical practice?

Data Feeds

- Adm Reg Data
- Billing Data
- Medical Staff Data
- Coding
- Documentation / Medical Record
- "Clinical Data"
- Abstracting
- Core Measures
- CPDF
- CDB
- CRM
- IA
- Reports

Data Quality Check & Reports
Is the data an accurate reflection of clinical practice?

Administrative vs. Clinical Data

- Debate on the usefulness of administrative data
- Clinical data requires analysis of the chart and can be very expensive
- Administrative data also comes from analysis of the chart
- The chart is a result of the clinician’s (mainly physicians) documentation
### Similarities & Differences

<table>
<thead>
<tr>
<th>Administrative Data</th>
<th>Clinical Data</th>
</tr>
</thead>
<tbody>
<tr>
<td>• From medical record of discharged patient</td>
<td>• From a medical record &amp; other IT systems</td>
</tr>
<tr>
<td>• Began as a financial process</td>
<td>• Individualized by the nature of the project</td>
</tr>
<tr>
<td>• Completed by educated coders</td>
<td>• Usually completed by clinicians</td>
</tr>
<tr>
<td>• Uses a standardized methodology</td>
<td>• Individualized by the nature of the project</td>
</tr>
<tr>
<td>• Does not include values or test results</td>
<td>• Could include values or test results</td>
</tr>
</tbody>
</table>

*The medical record is the place where clinicians take the results of tests and document the patient’s condition*

---

### Literature Review

- 'Administrative data outperformed single-day chart review for comorbidity measure'.

- 'Enhancement of claims data to improve risk adjustment of hospital mortality’

- 'Developing data production maps: meeting patient discharge data submission requirements’

- 'Comparison of administrative data and medical records to measure the quality of medical care provided to vulnerable older patients’
What Variables Can be Investigated

- Risk Adjusted Outcomes – Observed and Expected LOS, Mortality and Cost
- Other variables include: Complications, Readmissions, AHRQ PSIs, Charge, CMI

Performance based on:
- Hospitals
- Product Lines
- DRGs & MS-DRGs
- Diagnoses / Procedures
- Physicians
- Discharge Date/Month/Year
- Patient Demographics

Resource Utilization*:
- Blood Products
- Drugs
- Imaging Tests
- ICU
- Med/Surg Supplies
- Pharmacy
- * CRM

Items that may be different between administrative and clinical data

Uses of CDB / CRM Data

1. Ongoing consistent reports for meetings
   - Scorecards
   - Examining a DRG per meeting
   - Standard agenda items on Medical Staff Meetings, Leadership Meetings, Board Meetings

2. Improvement Initiatives
   - Drill down from scorecards
   - Answering a question
   - Improvement Priorities

3. Research

4. Improve accuracy of documentation & coding
2008 Data Quality Related Projects

- MS DRGs (complete)
  - Developed for resource use and are derived from a grouper
- Present on Admission
  - Must be consistently documented
- Bringing in ‘clinical data’ (e.g. lab results)
  - Infection Control Tool
- Shortening time frame for submission & return of data
- Download re-architecture
- Adding nursing units and physician names
- Post hospital mortality
  - Currently use phone follow up &/or master death file

3 Forms of Expression

- Management Reports
- Quality & Accountability Study
- CDB Online Data Tools
Quality & Accountability Study

- Three years

- Beginning to get traction as the most statistically based ranking on quality

- Measures include: mortality (aggregate and by product line), core measure (did each patient receive all measures), AHRQ patient safety indicators with the highest signal ratios, & equity (core measures by race, gender & SES)

Excellent improvement seen from 2006 to 2007

3's or below in no domains! 8 on 2

Kid/pan tx and plastic surg

Improving and Good Core Measure Scores

<table>
<thead>
<tr>
<th>Year</th>
<th>Score (Plants)</th>
<th>Overall</th>
<th>Mortality</th>
<th>Effectiveness</th>
<th>Safety</th>
<th>Equity</th>
<th>Efficiency</th>
<th>PI Center</th>
</tr>
</thead>
<tbody>
<tr>
<td>2006</td>
<td>64.6 (46)</td>
<td>52.6% (21)</td>
<td>48.9% (16)</td>
<td>58.5% (85)</td>
<td>92.4% (53)</td>
<td>48.9% (23)</td>
<td>42.5% (1)</td>
<td></td>
</tr>
<tr>
<td>2007</td>
<td>64.2 (36)</td>
<td>58.2% (42)</td>
<td>65.0% (22)</td>
<td>59.4% (55)</td>
<td>109.0% (13)</td>
<td>43.8% (15)</td>
<td>No data</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Domain</th>
<th>2006</th>
<th>2007</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mortality: O/E Ratio</td>
<td>3.00 (0)</td>
<td>2.77 (0)</td>
</tr>
<tr>
<td>Individual Product Line</td>
<td>10</td>
<td>10</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Domain</th>
<th>2006</th>
<th>2007</th>
</tr>
</thead>
<tbody>
<tr>
<td>Effectiveness</td>
<td>48.2% (3)</td>
<td>39.4% (1)</td>
</tr>
<tr>
<td>Individual Product Line</td>
<td>10</td>
<td>10</td>
</tr>
</tbody>
</table>

MIT Information Quality Industry Symposium, July 16-17, 2008
### Management Reports

#### Key Indicator Report (KIR)
- Clinical Indicators Report (COR)
- Hospital Quality Measures Report (HQMR)
- Quality & Safety Management Report (QSMR)
- Efficiency Management Report (EMR)
- Supply Chain Report (SCR)

- Semi-static reports you receive quarterly
- KIR can be thought of as a balanced scorecard
- Widely dispersed among the membership
- The more databases you are in, the more data you will receive

---

### Expressions – Q&A Study

#### 3’s or below in the following PSIs: Death in low mortality DRGs

<table>
<thead>
<tr>
<th>Year</th>
<th>Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>2006</td>
<td>1.97 (2)</td>
</tr>
<tr>
<td>2007</td>
<td>1.43 (2)</td>
</tr>
</tbody>
</table>

#### Efficiency: Cost or O/E Ratio (Score)

<table>
<thead>
<tr>
<th>Year</th>
<th>Labor Cost</th>
<th>Supply Cost</th>
</tr>
</thead>
<tbody>
<tr>
<td>2006</td>
<td>0.18 (2)</td>
<td>1.04 (2)</td>
</tr>
<tr>
<td>2007</td>
<td>0.16 (2)</td>
<td>1.08 (2)</td>
</tr>
</tbody>
</table>

#### Patient Centeredness: Domain Weight: 0%

### 2006: NICAPs 0 (0 or 10)
- NICAPs 0 (0 or 10)
- No NICAPs

### Safety PSIs

- Death in low mortality DRGs
- Stroke prevention
- Preventing falls
- Preventing pressure sores
- Preventing venous thromboembolism
- Preventing adverse drug events
- Preventing infections
- Preventing hospital-acquired pneumonia
- Preventing ventilator-associated pneumonia
- Preventing catheter-related bloodstream infections
- Preventing central line infections
- Preventing iatrogenic perforation
- Preventing postoperative hemorrhage
- Preventing nosocomial sepsis
Clinical Outcomes Report Face Page – Qtr 4

Green/Red dots based on rank

- Half red dots in Qtr and Year in Med Onc & CT Surg
- Green/Red dots based on rank

Clinical Outcomes Report Drill Down on Med Onc

2 of the last 4 above expected
This represents 31 deaths in the 4th quarter 2007
### CDB Interface Default Report

**Volume, LOS, ICU, Complications, Mortality**

#### LOS Summary By HCO

<table>
<thead>
<tr>
<th>HCO</th>
<th>Cases</th>
<th>Risk Pool Cases</th>
<th>% With Complication(s)</th>
<th>% Deaths (Obs)</th>
<th>% Deaths (Exp)</th>
<th>Mortality Index</th>
<th>% Early Deaths</th>
</tr>
</thead>
<tbody>
<tr>
<td>22231</td>
<td>8008</td>
<td>18.32</td>
<td>2.07 **</td>
<td>2.07</td>
<td>0.77</td>
<td>0.67</td>
<td></td>
</tr>
<tr>
<td>23203</td>
<td>13,116</td>
<td>19.27</td>
<td>2.74 **</td>
<td>3.11</td>
<td>0.88</td>
<td>0.69</td>
<td></td>
</tr>
<tr>
<td>30,074</td>
<td>266</td>
<td>14.31</td>
<td>2.10</td>
<td>2.10</td>
<td>0.84</td>
<td>0.57</td>
<td></td>
</tr>
<tr>
<td>32,618</td>
<td>13,715</td>
<td>14.33</td>
<td>1.65 **</td>
<td>1.87</td>
<td>0.88</td>
<td>0.52</td>
<td></td>
</tr>
<tr>
<td>55,817</td>
<td>18,063</td>
<td>13.24</td>
<td>1.34 **</td>
<td>2.17</td>
<td>0.62</td>
<td>0.38</td>
<td></td>
</tr>
<tr>
<td>30,673</td>
<td>13,146</td>
<td>19.41</td>
<td>2.63 **</td>
<td>3.45</td>
<td>0.76</td>
<td>1.09</td>
<td></td>
</tr>
<tr>
<td>21,774</td>
<td>9,421</td>
<td>16.98</td>
<td>1.97 **</td>
<td>2.74</td>
<td>0.72</td>
<td>0.54</td>
<td></td>
</tr>
<tr>
<td>52,817</td>
<td>24,244</td>
<td>15.53</td>
<td>1.82 **</td>
<td>1.99</td>
<td>0.92</td>
<td>0.45</td>
<td></td>
</tr>
<tr>
<td>44,094</td>
<td>15,757</td>
<td>14.53</td>
<td>2.25</td>
<td>2.15</td>
<td>1.05</td>
<td>0.75</td>
<td></td>
</tr>
<tr>
<td>44,586</td>
<td>18,289</td>
<td>19.14</td>
<td>1.73</td>
<td>1.83</td>
<td>0.95</td>
<td>0.40</td>
<td></td>
</tr>
<tr>
<td>59,435</td>
<td>32,159</td>
<td>20.06</td>
<td>1.62 **</td>
<td>1.93</td>
<td>0.84</td>
<td>0.49</td>
<td></td>
</tr>
<tr>
<td>36,363</td>
<td>13,039</td>
<td>17.97</td>
<td>1.38 **</td>
<td>1.59</td>
<td>0.87</td>
<td>0.29</td>
<td></td>
</tr>
<tr>
<td>23,589</td>
<td>12,586</td>
<td>17.4</td>
<td>1.94 **</td>
<td>2.58</td>
<td>0.75</td>
<td>0.75</td>
<td></td>
</tr>
</tbody>
</table>

---

**Time Frame is CY 2007**
MSDRG 163 – Chest Px w/ MCC exp of 36%
DRG exp of 24%
20 day LOS exp of 20 days
SOI and ROM of extreme

High c-value of .924, close to 20,000 cases in the model

Although administrative data has no results, it will include all conditions that are diagnosed from notes and results
Data Quality Study

- Goal is to evaluate whether the data in the CDB is an accurate reflection of clinical practice
- Used the 5 Chicago area academic medical centers
- Studied the data quality reports as well as global reports from the CDB
- 5 variables for each organization were chosen and contact with the member determined if the variance was real, an artifact of coding or documentation or something other
Study Summary

- UHC found the data discrepancies were mostly an effect of documentation and coding practices. In particular, they resulted from:
  - institutional emphasis on particular product lines,
  - documentation/coding of secondary diagnoses based on impact on reimbursement,
  - patient population, and
  - institutional patient safety/quality programs.

<table>
<thead>
<tr>
<th>4. MS-DRG</th>
<th>a</th>
<th>b</th>
<th>c</th>
<th>d</th>
<th>e</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>781 Other antepartum diagnoses w medical complications</td>
<td>90.0</td>
<td>85.4</td>
<td>92.0</td>
<td>90.52</td>
<td>79.3</td>
<td>86.4</td>
</tr>
<tr>
<td></td>
<td>4%</td>
<td>0%</td>
<td>4%</td>
<td>%</td>
<td>%</td>
<td>8%</td>
</tr>
<tr>
<td>782 Other antepartum diagnoses w/o medical complications</td>
<td>9.96</td>
<td>14.6</td>
<td>7.96</td>
<td>9.48</td>
<td>20.6</td>
<td>13.5</td>
</tr>
<tr>
<td></td>
<td>%</td>
<td>0%</td>
<td>%</td>
<td>%</td>
<td>%</td>
<td>2%</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>2. Fluid and Electro Disorders</th>
<th>a</th>
<th>b</th>
<th>c</th>
<th>d</th>
<th>e</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Comorbidity</td>
<td>N = 22,374</td>
<td>N = 26,969</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fluid disorders (n)</td>
<td>4,070</td>
<td>4,718</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Percent of All Cases</td>
<td>18.2%</td>
<td>17.5%</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

©2007 University HealthSystem Consortium
2. Characteristics of Tobacco Use

<table>
<thead>
<tr>
<th>ICD-9 Code</th>
<th>a</th>
<th>b</th>
<th>c</th>
<th>d</th>
<th>e</th>
</tr>
</thead>
<tbody>
<tr>
<td>All Cases</td>
<td>N = 22,374</td>
<td>N = 26,969</td>
<td>N = 15,008</td>
<td>N = 22,056</td>
<td>N = 32,380</td>
</tr>
<tr>
<td>v1582 - hx tobacco use (n)</td>
<td>3611</td>
<td>2612</td>
<td>152</td>
<td>4998</td>
<td>82</td>
</tr>
<tr>
<td>v1582 - hx tobacco use (%)</td>
<td>16.14%</td>
<td>18.53%</td>
<td>0.55%</td>
<td>%</td>
<td>0.47%</td>
</tr>
<tr>
<td>3051 - tobacco use disorder (n)</td>
<td>2478</td>
<td>2367</td>
<td>1037</td>
<td>3405</td>
<td>230</td>
</tr>
<tr>
<td>3051 - tobacco use disorder (%)</td>
<td>11.08%</td>
<td>12.63%</td>
<td>1.53%</td>
<td>%</td>
<td>3.20%</td>
</tr>
</tbody>
</table>

Clinical Data would not pick this up as it is an effect of documentation

The average number of diagnoses coded per case

<table>
<thead>
<tr>
<th>Diagnoses Profile</th>
<th>HCO</th>
<th>Cases</th>
<th>Mean #</th>
<th>Max #</th>
</tr>
</thead>
<tbody>
<tr>
<td>140088</td>
<td>22,374</td>
<td>10.072</td>
<td>59</td>
<td></td>
</tr>
<tr>
<td>140119</td>
<td>26,969</td>
<td>9.253</td>
<td>55</td>
<td></td>
</tr>
<tr>
<td><strong>140150</strong></td>
<td><strong>15,008</strong></td>
<td><strong>6.380</strong></td>
<td><strong>25</strong></td>
<td></td>
</tr>
<tr>
<td>140276</td>
<td>22,056</td>
<td>9.929</td>
<td>77</td>
<td></td>
</tr>
<tr>
<td>140281</td>
<td>32,380</td>
<td>8.230</td>
<td>30</td>
<td></td>
</tr>
</tbody>
</table>

This hospital does not seem to be giving itself 'credit' for the severity of their patients

This will also negatively effect reimbursement
Summary

• For use in performance improvement, administrative data (if proper checks are in place) can be an effective portrayal of clinical practice

• In addition, the CDB can assist a hospital in improving the accuracy of administrative data quality and accuracy