Examining Quality: Are We Making Progress?

E.A. Kroch, PhD
Hospital Studies

- Hospital Quality Incentive Demo
- QUEST Quality Improvement Collaborative
- AHRQ R18 Evaluation of QUEST
- Sepsis Mortality in Relation to QUEST Participation
- Readmissions and NLP
- “Phenotyping” Value of Care
- Patient Perception (HCAHPS)
CMS/Premier HQID

- Launched in October 2003 – 260 Hospitals
  
  **Hypothesis:** Financial Incentives improve hospital quality performance

- Years 1 to 3: 34 (process) measures across 5 clinical conditions:
  - Heart Attack (AMI)
  - Heart Failure
  - Pneumonia
  - Heart Bypass Surgery (CABG)
  - Hip/Knee Replacement

- Years 4 to 6: 41 nationally recognized measures in multiple clinical conditions.
HQID Hospitals vs. Non-participants

HQID hospitals have higher quality ratings than non-participants

- HQID participants avg. 6.8% higher than Non-Participants

- Avg. improvement for HQID participants = 10.9%

- Avg. improvement for Non-participants = 9.3%

New England Journal of Medicine publication by Lindenauer et al. (February 2007) found that hospitals engaged in P4P achieved quality scores 2.6 to 4.1 percentage points above other hospitals due solely to the impact of P4P incentives.

A composite of 19 measures shared in common between HQID and Hospital Compare shows P4P hospitals performing above the nation as a whole
HQID Findings

- Post program surveys could not explain differences in success rates.

- Some weak evidence came from the HLQAT and AHRQ surveys of culture & leadership.

- Hospitals treating under-served populations were at a payment incentive disadvantage.

- By the 6th year of the demonstration, we found no statistically significant disparity for top performance awards.

- Findings suggested that implementation of payment incentive programs should be gradual to allow disadvantaged hospitals time to adjust.
Perform Improve Collab: QUEST

- Qua*li*ty
- Effic*ien*cy
- S*afe*t
- T*ran*sparency

- Drive rapid improvement in both cost and quality
- Shape policy and payment guidelines
- Develop and refine the data collection and analytical tools to support value-based purchasing, harm prevention and public reporting
**Quest sample of U.S. hospitals**

157 Quest charter members include urban/rural, large/small and teaching/non-teaching facilities across 31 states.

Bed size ranges:
- 22% - 150 beds or less
- 29% - 151-300 beds
- 25% - 301-450 beds
- 24% - 451 or more beds

- 70% Disproportionate share
- 38% teaching
- 14% rural

33% Safety Net
Hospital Attributes Evaluated

This study analyzed performance variation among QUEST charter members by several key attribute segments:

- poverty percent*
- median income*
- safety net status
- facility location (urban or rural)
- Geographic region (MW, NE, S, and W)
- teaching status (ACGME)
- staffed bed size
- nurse staffing levels
- level of cardiac resources
- Medicare percent

* From 2010 Census and based on county of hospital location
Summary

- Hospitals that serve lower SES populations have less success in achieving collaborative performance goals for mortality and evidence-based care.
  - Disparities in EBC dissipated over time.
  - Disparities in mortality did not dissipate over time.

- By contrast, hospitals that serve a higher percentage of under insured and uninsured patients had lower severity adjusted patient cost compared to hospitals that serve a more affluent patient population.

- Regional differences suggest that providers in some regions, especially the Midwest have greater success, even after controlling for SES of the populations they serve.
Implications

- Pay-for-performance and other payment reforms need to evaluate population bias due to differences in hospital characteristics and include these factors in the assessment of performance as appropriate.

- Regional variation suggest that national collaboratives where providers can learn from each other to implement best practices may help resolve some of these differences.
Identifying factors associated with successful quality improvement: AHRQ R18 Eval

Learning from the QUEST collaborative

- Why are some organizations more successful than others?
  - What factors and strategies?

- Can the strategies of high performing institutions be spread?
Among QUEST hospitals inpatient mortality rates fell
But the experience of individual hospitals has varied
Why are some organizations more successful at implementing new clinical practices than others?

Klein and Sorra 1996
Study methods

- **Brief survey of QUEST Transformation Leads**
  - Implementation strategies for specific clinical practices (e.g., RRT, Sepsis Bundle)
  - Level of implementation

- **Very brief survey of clinical and administrative leaders**
  - Perceptions of clinical practices
  - Level of implementation

- **We link survey findings to changes in risk adjusted hospital mortality rates**
Examples of questions concerning implementation strategies

- Domain: Setting goals
  
  “Implementation of Sepsis Bundles has been an explicit quality improvement goal for senior administrative leaders at our hospital”

- Domain: Engaging champions and sponsors
  
  “The Sepsis Bundle initiative had an engaged physician champion”

- Domain: Using data to guide improvement
  
  “Performance data on the use of Sepsis Bundles are reviewed with frontline clinical staff and clinical leaders”
The Participation Effect
Results for Sepsis Mortality:

- What association exists between participation in sepsis-targeted collaborative events and decline in risk-adjusted sepsis mortality

- Tracked QUEST member participation in several collaborative events over the course of the course of six years – 2008-2013.
### QUEST performance improvement methodology

<table>
<thead>
<tr>
<th><strong>Educational calls</strong></th>
<th>One hour calls that aim to provide education on various topics related to the QUEST domains</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Sprints</strong></td>
<td>A 90 day rapid cycle improvement webinar series to help drive improvement in specific indicators</td>
</tr>
<tr>
<td><strong>Collaboratives</strong></td>
<td>A six to nine month improvement initiative focused on a specific condition, disease state or process of care</td>
</tr>
<tr>
<td><strong>QUEST National Meeting</strong></td>
<td>Two face to face meeting per year</td>
</tr>
<tr>
<td><strong>Resources</strong></td>
<td>Clinical director assigned to each organization, QualityAdvisor™ &amp; SafetySurveillor®, Performance Improvement Portal™</td>
</tr>
</tbody>
</table>
### Participation Rates (variation)

<table>
<thead>
<tr>
<th>Participation Rate</th>
<th>N</th>
<th>Mean</th>
<th>Std Dev</th>
<th>5th Percentile</th>
<th>95th Percentile</th>
<th>Max</th>
</tr>
</thead>
<tbody>
<tr>
<td>Degree of Sepsis Participation</td>
<td>185</td>
<td>9%</td>
<td>17%</td>
<td>0%</td>
<td>36%</td>
<td>91%</td>
</tr>
<tr>
<td>Degree of QUEST Participation 2010</td>
<td>193</td>
<td>16%</td>
<td>15%</td>
<td>0%</td>
<td>48%</td>
<td>83%</td>
</tr>
<tr>
<td>Degree of QUEST Participation 2011</td>
<td>240</td>
<td>13%</td>
<td>13%</td>
<td>0%</td>
<td>41%</td>
<td>56%</td>
</tr>
<tr>
<td>Degree of QUEST Participation 2012</td>
<td>301</td>
<td>15%</td>
<td>15%</td>
<td>0%</td>
<td>46%</td>
<td>86%</td>
</tr>
<tr>
<td>Degree of QUEST Participation 2013</td>
<td>296</td>
<td>13%</td>
<td>14%</td>
<td>0%</td>
<td>40%</td>
<td>80%</td>
</tr>
</tbody>
</table>

- A member’s participation rate equals the number of QUEST educational events attended divided by the total number of events offered to QUEST hospitals in a given year.

- The sepsis participation rate was limited to educational events with a sepsis focus.
Participation Effects from the **Binary Models**

### Participation Effects by Severity and Alternative Interactions

<table>
<thead>
<tr>
<th></th>
<th>All Sepsis</th>
<th>Sepsis</th>
<th>Severe Sepsis</th>
<th>Septic Shock</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dep Mean</td>
<td>-0.685</td>
<td>-3.089</td>
<td>-1.146</td>
<td>4.657</td>
</tr>
<tr>
<td>Adj R-sq</td>
<td>0.318</td>
<td>0.171</td>
<td>0.392</td>
<td>0.468</td>
</tr>
<tr>
<td>Part-Mod1</td>
<td>-0.211</td>
<td>-0.117</td>
<td><strong>-0.782</strong></td>
<td><strong>-1.441</strong></td>
</tr>
<tr>
<td>Part-Mod2</td>
<td>-0.300</td>
<td>-0.133</td>
<td><strong>-0.765</strong></td>
<td><strong>-1.537</strong></td>
</tr>
<tr>
<td>Part-Mod3</td>
<td>-0.129</td>
<td>-0.097</td>
<td>-0.755</td>
<td><strong>-1.283</strong></td>
</tr>
<tr>
<td>Q-Flag</td>
<td>-0.509</td>
<td>-0.372</td>
<td>-0.310</td>
<td>-0.109</td>
</tr>
<tr>
<td>Q-Trend</td>
<td><strong>-0.192</strong></td>
<td>-0.060</td>
<td>0.002</td>
<td>-0.210</td>
</tr>
<tr>
<td>Nobs</td>
<td>4142</td>
<td>4081</td>
<td>3391</td>
<td>3555</td>
</tr>
</tbody>
</table>

**KEY:**
- p<0.10
- p<0.05
- p<0.01

- Model 1 is a linear trend effect
- Model 2 allows for unconstrained time effects
- Model 3 allows for cohort time effects
Discussion of Participation Effect

- Using participation rates did not alter results.
- Preliminary mortality results from an independent parallel study (funded by AHRQ R18 program):
  
  “It appears that participation in QUEST activities and use of data are the best predictors of mortality reduction…”

- Next steps
  - Introduce alternative participation rate metrics
  - Account for time lags
  - Segment observations by PDx (bacterial vs. other)
  - Other model specifications

- General discussion
Hospital Readmissions: The Challenge of Modeling Acute Inpatient Rehospitalizations

- Preventable (Avoidable) Hospitalizations
  - Since treatable chronic illness are responsible for many such hospitalizations, does their existence mean failure of the system?

- Ambulatory Care Sensitive Conditions
  - Inadequate access to primary care?

- Poverty and Social Determinants of Disease
  - Related to poverty and associated social determinants, and the strategies that must be deployed will have to respond to these needs.

- Natural Language Processing (NLP)
  - Enriching the information set with nontraditional clinical information in three domains: physical function, cognitive status, and psychosocial environment
NLP Domains and generated terms from focus groups.

<table>
<thead>
<tr>
<th>Group Term</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Cognitive Status</strong></td>
<td></td>
</tr>
<tr>
<td>Impaired Mental Status</td>
<td>Difficulty thinking, comprehending, or following directions</td>
</tr>
<tr>
<td>Limited Capacity/Competency</td>
<td>Impaired ability to make medical decisions</td>
</tr>
<tr>
<td>Mood Problems</td>
<td>Mood problems</td>
</tr>
<tr>
<td>Difficult Behavior</td>
<td>Patients with problematic interactions with care team staff</td>
</tr>
<tr>
<td>Negative Ideation</td>
<td>Suicidal or violent thoughts</td>
</tr>
<tr>
<td>Inadequate Health Literacy</td>
<td>Poor understanding of medical concepts/tasks</td>
</tr>
<tr>
<td>Positive Cognitive Issues</td>
<td>Statements consistent with adequate cognitive ability</td>
</tr>
<tr>
<td><strong>Physical Function</strong></td>
<td></td>
</tr>
<tr>
<td>Poor Physical Function</td>
<td>Significantly impaired physical function</td>
</tr>
<tr>
<td>Poor ADLs†</td>
<td>Difficulty completing Activities of Daily Living</td>
</tr>
<tr>
<td>Fall Risk</td>
<td>Gait or balance problems with increased risk for falling</td>
</tr>
<tr>
<td>Needs Assistance with Movement</td>
<td>Needs assistance with moving, transferring, and ambulation</td>
</tr>
<tr>
<td>Maximum Care</td>
<td>Needs maximum care, non-ambulatory, bed bound</td>
</tr>
<tr>
<td>Poor Condition</td>
<td>Frailty, malnutrition, or neglect</td>
</tr>
<tr>
<td>Pain</td>
<td>Chronic or uncontrolled pain</td>
</tr>
<tr>
<td>Decubitus Ulcers</td>
<td>Has bed sores</td>
</tr>
<tr>
<td>Positive Physical</td>
<td>Statements consistent with adequate physical function</td>
</tr>
<tr>
<td><strong>Psychosocial Support and Environment</strong></td>
<td></td>
</tr>
<tr>
<td>Care Management Program</td>
<td>A program in our hospital system for a small subset of high risk patients determined by CMS* who receive intensive outpatient case management.</td>
</tr>
<tr>
<td>Substance Abuse</td>
<td>Alcohol or other substance abuse</td>
</tr>
<tr>
<td>Domestic Abuse</td>
<td>Domestic abuse</td>
</tr>
<tr>
<td>At Risk</td>
<td>Patients previously stated to be high risk for adverse outcomes or offered outpatient programs for patients at risk for adverse outcomes</td>
</tr>
<tr>
<td>Feels Unsafe</td>
<td>Patient does not feel safe in their home environment</td>
</tr>
<tr>
<td>Unstable Housing</td>
<td>Inadequate, unstable, or absent permanent housing</td>
</tr>
<tr>
<td>Poor Social Support</td>
<td>Inadequate contact with friends, family or other positive social supports</td>
</tr>
<tr>
<td>Logistical Difficulties</td>
<td>Issues related to childcare, transportation to appointments, or scheduling follow-up</td>
</tr>
<tr>
<td>Jail</td>
<td>Current or prior incarceration</td>
</tr>
<tr>
<td>Financial Problems</td>
<td>Inadequate insurance, employment, or access to funds for medications/treatments</td>
</tr>
<tr>
<td>Positive Psychosocial</td>
<td>Statements consistent with adequate psychosocial support/supportive environment</td>
</tr>
</tbody>
</table>
Discriminatory Power is Moderate

C-Stat = 0.773

C-Stat = 0.682
Summary

- Patient-level modeling to control for readmission risk is moderate at the hospital level.
- Hospital discrimination rests mostly on the distribution of principal diagnoses.
- Patient-level modeling of most individual diagnoses (HF, AMI, PNE) has very little discriminatory power.
- Using our best models to control for patient readmission risk does not leave the remaining variation in readmission rates across hospitals attributable to hospital performance.
- One left out variable set is socio-economic demographics
“Phenotyping” Hospital Value of Care for Patients with Heart Failure

- Objective: To characterize hospitals based on patterns of their combined financial and clinical outcomes for heart failure hospitalizations longitudinally.

- Data: Detailed cost and administrative data on hospitalizations for heart failure from 424 hospitals in the 2005–2011 Premier database.

- Design: Using mixture modeling to identify groups of hospitals with distinct joint trajectories of RSC per hospitalization and RSMR, we assessed hospital characteristics associated with the distinct patterns using multinomial logistic regression.


- Conclusions: Hospitals vary substantially in the joint longitudinal patterns of cost and mortality, suggesting marked difference in value of care. Understanding determinants of the variation will inform strategies for improving the value of hospital care.
Value of HF Care (HSR paper)
Value of HF Care  
(HSR paper)

- Phenotype 1: Highest cost and low mortality (n = 14 hospitals, 3.2 percent of the hospitals in sample)
- Phenotype 2: High cost and low mortality (n = 86 hospitals, 20.4 percent of the hospitals in sample)
- Phenotype 3: Medium cost and low mortality (n = 147 hospitals, 34.6 percent of the hospitals in sample)
- Phenotype 4: Medium cost and high mortality (n = 26 hospitals, 6.2 percent of the hospitals in sample)
- Phenotype 5: Low cost and low mortality (n = 151 hospitals, 35.6 percent of the hospitals in sample)

Dotted lines reflect confidence intervals for each trajectory, based on the 95 percent confidence intervals of the point estimates in each year.
Value of AMI Care

Medical Care paper
Is Patient Experience More Than Just Perception?

HCAHPS Patient Satisfaction and Hospital-Acquired Infections

- Hospitals value patient satisfaction, but rarely do they consider patient perception an indicator of efficacious care.

- Hospital-acquired infection rates came from monitoring harm in the context of the QUEST improvement collaborative.
Hospital-Acquired Infections Evaluated

- Clostridium difficile (Cdif)
- Staphylococcus areus septicemia
- Central line associated blood stream infections
- Catheter associated urinary tract infections (CA-UTI)
- Ventilator associated pneumonia (VAP)
- Mediastinitis after coronary artery bypass procedures
- Infections following orthopedic procedures
- Infections following bariatric procedures
Composite Measures

- In addition to the individual hospital-acquired infection measures, 2 composite measures were evaluated:
  - Hospital-Acquired Infection Composite
    - Volume-based weighted average of 8 individual hospital-acquired infection measures
  - Harm Composite
    - Includes 25 measures given equal weight after adjusting for central tendency and dispersion, effectively turning each measure into a standardized score.
    - Measures include hospital-acquired infections, surgical or post-operative complications, and hospital wide harm.
Statistical Analysis

- Hospital-acquired infection rates and patient satisfaction scores were correlated at the hospital level.

- Infection rates were based on the populations at risk to make cross-hospital comparisons meaningful.

- Satisfaction scores were “top-box” response rates, e.g., percent of respondents who answered “always” to the question, “How often was your room … kept clean?”

- Pearson product-moment correlations, isolating partial effects from risk/acuity effects.
## Correlations – Individual Measures

<table>
<thead>
<tr>
<th></th>
<th>Cleanliness</th>
<th>Nurse Communication</th>
<th>Doctor Communication</th>
<th>Help Quickly</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cdiff</td>
<td>-0.189</td>
<td>-0.268</td>
<td>-0.357</td>
<td>-0.355</td>
</tr>
<tr>
<td>Septicemia</td>
<td>-0.208</td>
<td>-0.140</td>
<td>-0.232</td>
<td>-0.211</td>
</tr>
<tr>
<td>CLABSI</td>
<td>-0.222</td>
<td>-0.250</td>
<td>-0.259</td>
<td>-0.296</td>
</tr>
<tr>
<td>CA-UTI</td>
<td>0.064</td>
<td>0.019</td>
<td>0.010</td>
<td>0.056</td>
</tr>
<tr>
<td>VAP</td>
<td>-0.282</td>
<td>-0.156</td>
<td>-0.176</td>
<td>-0.184</td>
</tr>
<tr>
<td>Mediastinitis After CABG</td>
<td>-0.090</td>
<td>-0.014</td>
<td>-0.050</td>
<td>-0.099</td>
</tr>
<tr>
<td>Infection After Ortho Procs</td>
<td>-0.208</td>
<td>-0.201</td>
<td>-0.075</td>
<td>-0.146</td>
</tr>
<tr>
<td>Infection After Bariatric Procs</td>
<td>-0.045</td>
<td>-0.065</td>
<td>-0.043</td>
<td>-0.059</td>
</tr>
</tbody>
</table>

Correlations in bold red indicate \( p < .05 \).
Implications and Limitations

- Patient satisfaction scores can be useful indicators of the quality of hospital care such as efficacious infection control.

- Some anticipated relationships emerged – others did not.
  - Risk-adjusted LOS & Cost were negatively correlated with scores.
  - No significant effects for injury and pressure ulcers (QH 27 & 28)

- Study Limitations:
  - Administrative data
    - Currently compiling NHSN surveillance data for most of these conditions
  - Did not control for certain confounding factors, such as predisposition of patient scores as a function of type and severity of illness.
    - See study by Marc Harrison and James Merlino of the Cleveland Clinic (February, 2011).
    - Preliminary analyses showed a strong negative correlation between satisfaction scores and patient prognosis.
DISCUSSION?

- Hospital Quality Incentive Demo
- QUEST Quality Improvement Collaborative
- AHRQ R18 Evaluation of QUEST
- Sepsis Mortality in Relation to QUEST Participation
- Readmissions and NLP
- “Phenotyping” Value of Care
- Patient Perception (HCAHPS)