Measuring the quality of intensive care

Gordon D. Rubenfeld, MD MSc
Professor of Medicine, University of Toronto
Chief, Program in Trauma, Emergency, and Critical Care
Sunnybrook Health Sciences Centre
Outline

• What are the challenges of measuring the quality of intensive care?
  – Are quality measures valid?
  – Are quality measures and reporting effective?
  – Are quality measures and reporting safe?
• Should we be reporting more publically?
Identifying the best doctors and hospitals looks easy.
Risk Adjusted Outcome

The Standardized Mortality Ratio

it’s just long division

< 1 = good, > 1 = bad

$$SMR = \frac{\text{Observed deaths}}{\text{Predicted deaths}} = \frac{117}{133} = 0.88 (0.68 - 1.3)$$
Risk Adjusted Outcome
The Standardized Mortality Ratio

it’s just long division
< 1 = good, > 1 = bad

\[
SMR = \frac{\text{Observed deaths}}{\text{Predicted deaths}} = \frac{117}{133} = 0.88 (0.68 - 1.3)
\]

\[
\sum \hat{a}_0 + \hat{a}_1 \text{APS}_{\text{day}1} + \hat{a}_3 \text{CHRONICHEALTH}
\]
Risk Adjusted Outcome

The Standardized Mortality Ratio

it’s just long division

< 1 = good, > 1 = bad

\[
\text{SMR} = \frac{\text{Observed deaths}}{\text{Predicted deaths}} = \frac{117}{133} = 0.88 \ (0.68 - 1.3)
\]

95 – 162 0.88 (0.5 – 1.8)

\[
\sum \hat{a}_0 + \hat{a}_1 \text{APS}_{\text{day1}} + \hat{a}_3 \text{CHRONICHEALTH}
\]
Every ICU cohort looks like this.
Problems with risk adjusted mortality

If you can’t reduce your numerator, increase your denominator

• Upcoding for quality
  – In response to outcome assessment of CABG, coding of comorbidities jumped abruptly at some hospitals, including a 1 year jump in “COPD” from 1.8% to 52.9%
  – “Good” hospitals more likely to overcode severity of illness.

Problems with risk adjusted mortality

*If you don't like your SMR, use a different risk model*

- Comparing hospital rankings for 5 severity models with similar statistical performance in 3 diseases in 108 hospitals
  - Good agreement ($\kappa > 0.7$) in only 11 pairs of assessments
- 40 Dutch ICUs compared using 3 risk adjustment models (APACHE 2, SAPS 2, MPM 2)
  - 21 ICUs were identified as performance outliers by one of the 3 models
  - Only 1 ICU was identified by all 3 models as an outlier

Iezzoni LI. *JAMA*. 1997;278:1600-1607.
Bakshi-Raiez, et al, *CCM* 2007; 35:
Problems with risk adjusted mortality

*If you don’t like your SMR, transfer out not in*

- **Transfer in**
  - Patients transferred to academic medical centers’ ICUs have double the mortality predicted by APACHE score (even though this model contains a variable for transfer)
  - *Importing* patients increases SMR

- **Transfer out**
  - Death is attributed to the receiving hospital
  - *Exporting* patients reduces SMR

Chest. 2007 Jan;131(1):68-75
Problems with risk adjusted mortality

*If you don’t like your SMR, spin the wheel again*

- Several studies use simulated populations of hospitals with “perfect” risk adjustment
  - 56-82% of the excess death rates were explained by random variation
  - 65% of bad outlier hospitals in any given year were due to chance
- Data on > 1000 cases improve accuracy
- Outliers in > 1 year improve accuracy

*JAMA. 1990;264:484-90.*
*Hosp Health Serv Adm. 1997;42:3-15.*
*Med Care. 1996;34:737-53.*
Case volumes for most procedures are too low to use mortality as outcome ... the same is probably true for many ICUs.

Does SMR correlate with other measures of quality?
Even US News and World Reports ranking does not correlate with process of care.

Compared 41 US News “Best Cardiac Hospitals” to 733 other hospitals by process of care.

Williams, et al, Circulation. 2006 Aug 8;114(6):558-64
Level 1 Trauma Centers
Variation in outcome for severe injury

In most cases, discrepancies between observed and expected outcomes should serve to trigger a review of the process of care to determine if quality problems in fact exist.
[...] when used as a measure of quality for individual hospitals, risk-adjusted mortality rates are seriously inaccurate. Publication of hospital mortality rates misinforms the public about hospital quality.
Interim conclusions

- Risk adjusted outcome (SMR) is unreliable, gameable, and does not correlate with other measures of quality.
- Chance plays a large part in assigning rank particularly in smaller centers.
- Risk adjusted outcome (SMR) does not always correlate with other measures of quality.
Does measuring and reporting quality have an effect on patients or physicians?
Effect of publication of quality measures

- **On patients**
  - CABG patients unaware of outcome data
  - More likely to base decisions on anecdotal reports than data

- **On physicians**
  - Minimal effect on referral
  - Perception that it affected surgeon willingness to do high-risk cases

- **On hospitals**
  - Depends on results – no one complains (or does anything about) low SMR
  - High SMR – complaints about data quality

Quasi-randomized trial of performance rating data in Wisconsin

24 hospitals Alliance Group

Public report

91 hospitals in Wisconsin

Confidential report

No report

Caveats apply: public report hospitals were not a randomized group, administrative data reports, could not distinguish activities in response to reports from pre-existing activities

Health Affairs. 2003 Mar-Apr;22(2):84-94
Validity of performance ratings are in the eye of the beholder

- Perception of validity of performance rating and hospital performance correlated ($r=0.13$) but only in public reporting hospitals

In general we don’t believe performance ratings, but we really don’t believe them if we look bad and other people will see the results

Health Affairs. 2003 Mar-Apr;22(2):84-94
Public report hospitals more likely to have quality improvement activities

“Low” performing public report hospitals more likely to engage in QI

<table>
<thead>
<tr>
<th>Number of activities</th>
<th>Public-report hospitals</th>
<th>Private-report hospitals</th>
<th>No-report hospitals</th>
</tr>
</thead>
<tbody>
<tr>
<td>5</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Obstetric care: Public-report hospitals with 5 activities exceed the other two categories.

Cardiac care: Public-report hospitals with 3 activities lead the other two categories.
Effectiveness of Public Report Cards for Improving the Quality of Cardiac Care
The EFFECT Study: A Randomized Trial

- Cluster RCT in 86 hospitals in Canada
- Early or delayed public release of cardiac report cards
- **No effect** on 12 AMI or 6 CHF processes of care
- **But**
  - *Increased* quality improvement activities at early intervention
  - *Improved* (small absolute) mortality in AMI and CHF with poor LV function

JAMA. 2009;302(21):2330-2337
If public reporting of quality measures is questionnably effective – is it at least safe?

- Improvements in CABG outcomes with public reporting associated with decreased access
  - By high risk patients
  - By racial minorities perceived to be at higher risk

- Focus on process measures without selective review may lead to indiscriminate overuse
  - 90% of patients who did not receive colorectal screening were excluded appropriately
  - Time to antibiotics in ED – PIP/TAZO for CHF?

- Perception that using complications as quality measure will lead to decreased screening for complication

JAMA. 2005 Mar 9;293(10):1239-44.
What about critical care?

- Evidence based process measures are increasingly problematic
- Mortality is distal and much of what we do is palliative
- VAP and CR-BSI are ... VAP and CR-BSI
Some issues may be hospital specific

Sunnybrook
TRAUMA, EMERGENCY & CRITICAL CARE

+ Night time discharges
A roadmap for the future?

- Measure Quality
- Report Privately
- Ignore Results

- Utter Nihilism
- Exit Only

- Measure Quality
- Report Publicly
- Debate Endlessly
  - TO EAST
  - 401
Conclusions

- Domains of quality (access, structure, process, outcome, patient experience) do not converge and may be orthogonal.
- The SMR (or ranking based on it) is misleading and is neither a reliable nor valid measure of quality.
- Public reporting of outcome data can have harmful effects that must be considered but may be necessary to be effective.
- What you measure is not nearly as important as what you do with what you measure.
Conclusions

• Only intensivists may care about the “quality” of intensive care – patients and payers probably care more about the quality of care overall

• The best (and safest) way to use quality measures is not known
  – Public or private reporting
  – Administrative data or locally collected
  – Explicit tools, coaching, data feedback in what form?
  – Pay for performance?
Recommendations for measuring and improving ICU performance

- Quality should be assessed along multiple domains.
- Benchmarking absolute performance between ICUs should be discouraged in favor of comparing improvement in performance.
- If benchmarking, assess as outlier only after considering multiple risk adjustment tools and serial status.
- Given relative weakness of evidence in critical care, ICUs should be allowed to select their own performance measures from a menu.
We cannot ignore quality data even if we disagree on its use and validity.
Put the data you have uncovered to beneficial use.

Slides and Conflict Vitae available on request
gordon.rubenfeld@sunnybrook.ca