Quality of Care Measurement: Risk-Adjustment Methods

Randall D. Cebul, MD
Professor of Medicine, Epidemiology & Biostatistics
Case Western Reserve University at
MetroHealth Medical Center
rdc@po.cwru.edu

Objectives

- Context(s)
- Purposes of Risk Adjustment
- Settings that require risk adjustment
- Typical components of risk adjustment
- Performance measures for risk-adjustment
- Limitations/pitfalls

Contexts in Which Risk-Adjustment is Important

- Observational Studies
  - Measuring the effect of an intervention (or, more generally, an "exposure") from non-experimental data
- Quality of Care
  - How you "perform" depends on the hand you're dealt.
- Report Cards
  - How does your [practice, hospital, MCO] measure up?
  - Cleveland Health Quality Choice (CHQC) Program

Quality of Care: Who's Being Measured?

- Individual Providers
- Provider Organizations
  - Hospitals
  - Managed Care Organizations/HMO's
  - Practices

Today's *COMPETING HOSPITAL NEWS*:

"Local Teaching Hospital losing colon cancer treatment race to Community General"....

Data from the hospital association, 2000

- 63% of Community General's patients achieved remissions after colon cancer treatment.
- 44% remission rate at Local Teaching Hospital for treatments of colon cancer

The reporter from *COMPETING HOSPITAL NEWS* calls you (the well known health outcomes / statistics expert) for your opinion about why Local Teaching Hospital may be doing so poorly.

She says she has more data from the hospital association.

What do you want to know?

Table 1

Proportion of Cases by Stage of Cancer
(Strong *Selection Bias*)

Cancer Stage	Community General	Local Teaching Hospital	Total
I	98 (49%)	49 (10%)	147 (21%)
II	48 (24%)	151 (30%)	199 (28%)
III	52 (26%)	302 (60%)	354 (51%)
Total	198 (100%)	502 (100%)	700 (100%)

Table 2

Remission Rates by Cancer Stage
(Risk-stratified Outcomes)

Cancer Stage	Community General		Local Teaching Hospital		Total	
I	80/98	(82%)	40/49	(82%)	120/147	(82%)
II	28/48	(58%)	88/151	(58%)	116/199	(58%)
III	16/52	(31%)	94/302	(31%)	110/354	(31%)
Total	124/198	(63%)	222/502	(44%)	346/700	(49%)

Headlines in Tomorrow's COMPETING HOSPITAL NEWS

Commentary* in the *Medical Times Gazette.
February 13, 1864, page 187

The following comment was a reply to comparisons of death rates across English hospitals by London's office of the Registrar General:

"Any comparison which ignores the difference between the Apple-cheeked farm-laborers at Stoke Pogis (probably for Rheumatism and sore legs), and the wizzened, red-herring-like Mechanics of Soho or Southwark, who come from a London Hospital, is fallacious."

From CHQC Reports (1991-98) - Introductory Comments:

"Severity-Adjustment for Fairness

One important aspect of this report on hospital quality is that the data are severity-adjusted. That means that the statistical model takes into account patient-specific variables that are beyond the control of the hospitals, such as the age of the patient, the severity of the illness, chronic health status, etc. …

This … feature enables fair comparisons of the same diagnosis at all of the hospitals because it takes into consideration pre-existing factors that may alter the treatment and outcome of care."

What are the 3 General Types of Measures Used to Assess Quality (Donabedian)?

1.

2.

3.

What are the 3 General Types of Measures Used to Assess Quality?

1. Structure (applies more to organizations)
   - Facilities
   - Credentials
   - Etc
2. Process (of Care)
   - Appropriate test(s) ordered
   - Appropriate treatment(s) administered
   - Adherence to accepted guidelines for care
3. **Outcomes (of Care) – principal focus of session**
   - **Quantity of Life (Mortality)**
   - **Quality of Life**
   - **Satisfaction**

What is the ideal study design for making inferences about quality of care from outcomes of care?

Ideal Study Design for Making Inferences about Quality from Outcomes

- Random assignment of "*at-risk*" patients to providers or provider organizations
- Large (enough) sample size
- Appropriate time frame to capture outcome of interest (e.g., mortality, morbidity, functional status, satisfaction)
- Appropriate and reliably measured outcome
  - These design standards minimize bias due to selection [random assignment] & false inferences due to inadequate power [sample size, duration], noise [sample size, reliable measurement], or inappropriate outcome

"At-risk" patients: who/what are the subjects of quality measurement?

- Usually –
  - particular kinds of care (e.g., flu shots, mammograms)
  - particular conditions (e.g., diabetes, congestive heart failure, stroke, pregnancy)
  - particular procedures (e.g. surgery, angioplasty)
- Less commonly –
  - All enrollees of a health plan, or patients receiving care at a hospital

Most Outcomes-focused Reports of Quality Use Observational Data, Not RCTs

- Patient registries (SEER)
- Administrative databases (MEDPAR)
- Quality performance databases (HEDIS)
- Organizational databases (GM, Ford, MCOs)
- Institutional clinical databases
- Regional hospital report card databases (Dayton, Cleveland, Cincinnati)

Risk Adjustment

- Purpose: to level the playing field in making comparisons of outcomes, either against a benchmark, or across individual providers, hospitals, or provider groups.
- "The rationale for risk adjustment is to remove one source of this variation [in outcomes across providers], leaving residual differences to reflect quality"
- "The underlying assumption is that outcomes reflect a mix of factors: **patient outcomes =**
  effectiveness of treatments +
  patient risk factors that effect response to treatment + quality of care +
  random chance."

  Iezzoni LI. *JAMA*. 1997; 278: 1600-1607.

Consequences of NOT leveling the playing field

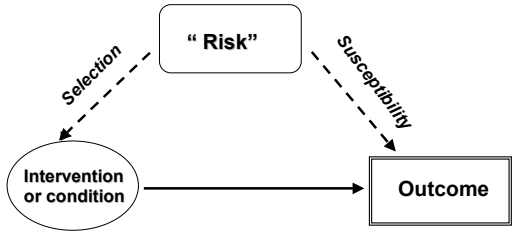
- Inappropriate inference of high quality:
  - E.g.; patients cared for by a given provider are "less sick" than those cared for by others – outcomes are better Not because of higher quality, but because patients were more likely *a priori* to have good outcomes
- Inappropriate inference of low quality:
  - E.g.; patients cared for by a given provider are "sicker" than those cared for by others – outcomes are worse Not because of lower quality, but because patients were more likely *a priori* to have bad outcomes

Risk-Adjustment is necessary whenever comparisons are being made from observational data:

- Across providers (hospitals, MCOs)
- Between a provider and a benchmark result*
- Over time, if selection changes (relevant to CQI)

* unless the benchmark is an absolute

Risk adjustment is necessary when susceptibility to outcomes varies across providers' patients, especially if there is differential selection by risk-related factors



Components of Risk to Consider in Adjusting

- Central Questions:
  - Is the factor intrinsic to the patient before the encounter (pre-existing - "provider not responsible for it")?
  - Is the factor associated with the outcome of interest?
  - Might the factor be associated with selection for the intervention/encounter?
- "The Big 3"
  - Sociodemographic factors
  - Comorbidity
  - Severity of illness

The Big 3 Factors in Risk Adjustment
Ignore These at Your Own Peril…

- Sociodemographic factors:
  - Age, sex, race/ethnicity*, SES, (genetic characteristics)
- Comorbidity
  - The importance, number, (and severity) of co-existing conditions
- Severity of Illness
  - Acute clinical stability; importance, extent, and severity of principal diagnosis

* See discussion for considerations re: including race/ethnicity

Other Factors to Consider in Risk Adjustment

- Importance depends on what outcomes are being measured & the empirical relationship of factor to outcomes
- Pre-care:
  - Function (physical, cognitive, psych, psychosocial)
  - Health Status
  - Quality of life
  - Attitudes and preferences

Illustrative Methods to Adjust for Comorbidity

Measure	Outcome of Interest
• Comorbidity Index – 3 subscales for comorb. – Severity, complications, & functional status – Score: 0-4	Hospital Mortality
• Charlson Comorbidity Index – Weights (1-6) for 19 Comorbid illnesses	Hospital Mortality

Conditions and Weights in Charlson Index

Condition	Weight
Myocardial Infarction	1
CHF	
Peripheral Vasc Disease	
Cerebrovascular Disease	
Dementia	
COPD	
Coll Vascular Disease	
Ulcer Disease	
Mild Liver Disease	
Diabetes	
Hemiplegia	2
Moderate-severe Renal Disease	
Diabetes with end organ disease	
Any Tumor	
Leukemia	
Lymphoma	
Moderate-severe liver disease	3
Metastatic solid tumor	6
AIDS	

Illustrative Severity of Illness Adjustment Methods

Measure	Outcome, Data Source
• APACHE – Score: 0-71	Hospital mortality, ICU Pts – Medical record review
• Dis. Staging (DS-Clin) – Score: 1-3, 420 Dxs	Complications, Dx-specific – Medical record review
• DS-Scale – Score: 0-100	Hospital Mortality, All Dx – D/c abstract
• Medisgroups – Score: 0-4	Hospital Mortality, Generic – Medical record review
• HCFA Model	Hospital Mortality, Generic – D/c abstract

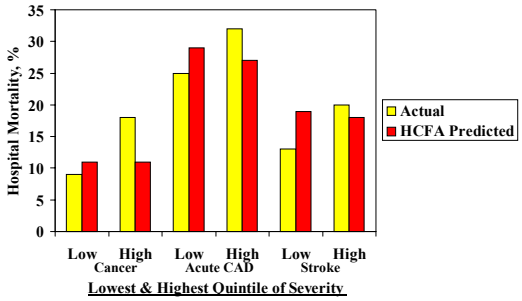
The HCFA Mortality Adjustment Method

- Purpose: to risk-adjust hospital death rates for public reporting of hospital performance
- Application: Observed/Expected deaths (where "expected" is based on patient characteristics in the risk-adjustment model)
- Model Equation:

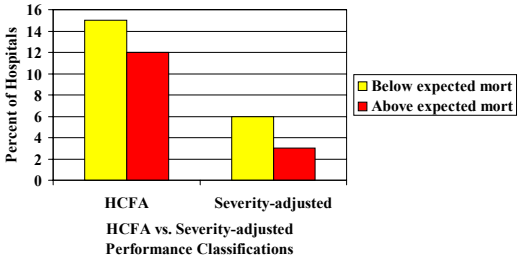
$$Ln\ P/(1-P) = B_o + B_1\,(\text{N of prior adm}) + B_2\,(\text{sex}) + B_3\,(\text{transfer status}) + B_i\,(\text{age group}) + B_i\,(\text{comorbidity category})$$

- What's missing?

HCFA's model was insensitive to differences in severity of illness: generally, overestimating mortality in low severity & overestimating in high severity patients



Severity-adjustment reduced the frequency of "extreme" classifications (9%) as compared with the HCFA model (27%)



HCFA Mortality Model: In Memorium 1986-1991

Components of Risk-Adjustment: Summary

- Risk adjustment of outcomes may require detailed data pertaining to patient features not routinely acquired & reported – the specific data may depend on the outcome being evaluated
- In evaluating mortality, The Big 3 should always be included – other issues (e.g., transfer status, special services attracting extreme cases) also should be considered
- Incomplete adjustment may result in inappropriate classification of performance

How do We Measure the Performance of Risk-adjustment Methods?

- Statistical- distribution of results: Does the method yield the expected number of high and low performance outliers?
- Convergent validity: Are there other data to support the inferences? How does the measure compare with others?
- Statistical methods for predicting binary outcomes: ROC area, c-statistic (0.5-1.0)

Measures to Predict In-hospital Stroke Mortality:
C-Statistic (ROC Curve Area)

R-A Measure	c-statistic
MedisGroups (MG)	0.87
CHQC	0.86
Physiol Score (PS)	0.84
Disease Staging (DS)	0.74
PMC Severity Score (PMC)	0.73
APR-DRGs	0.77

Measures to Predict In-hospital Stroke Mortality Result in DifferentRisk Estimates:
Odds of Death Predicted by Measures A & B*

Measure A	Measure B	A>B	A<B
MG	PS	8.1%	9.7%
MG	DS	16.2%	41.6%
MG	APR-DRGs	13.7%	34.7%
PS	DS	15.7%	36.3%
DS	APR-DRGs	22.6%	18.7%

•A>B – odds of death >2.0 higher by A than B
•A<B - odds of death 0.5 or lower by A than B

Consequences of inadequate or inaccurate adjustment

- Inappropriate inference of high quality
- Inappropriate inference of low quality

- Re: stroke – will your hospital fail to get appropriate credit for the sickness of its patients because of inadequate or inaccurate adjustment?

Other Limitations of Risk-Adjustment Methods

- Race/ethnicity/insurance
- Cost of Data
- In-hospital mortality as a measure of quality
  - vs. 30-day
  - vs. Other endpoints
  - accounting for DNR status

Other Limitations of Risk-Adjustment Methods

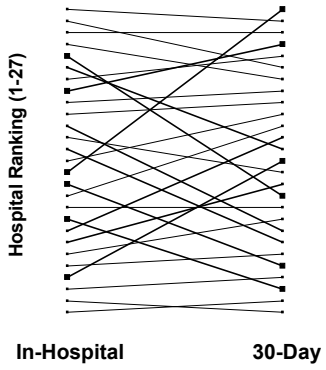
- The Race/ethnicity/insurance Conundrum
  - Minorities and uninsured patients may receive lower-quality hospital care than others
  - Adjusting providers' care for these patient characteristics may mask these important differences in quality; but
  - These characteristics also may be associated with poorer outcomes independent of quality
  - What are the options?

Other Limitations of Risk-Adjustment Methods

- Cost of Data
  - Administrative data alone: lower cost-lower quality
  - Linking lab data to administrative data: better quality-higher cost
  - Linking lab + pharmacy data: better quality & higher cost still
  - Medical records abstractions: CHQC model, higher cost still- $17-20/record; $70-134K/hospital

Other Limitations of Risk-Adjustment Methods

- In-hospital mortality as a measure of quality
  - vs. 30-day



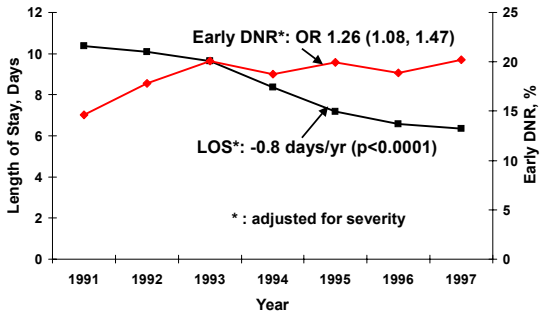
Other Limitations of Risk-Adjustment Methods

- In-hospital mortality as a measure of quality
  - vs. 30-day
  - vs. Other endpoints
    - Functional status (eg, stroke)
    - Readmission rates (eg, CHF)

Other Limitations of Risk-Adjustment Methods

- In-hospital mortality as a measure of quality
  - vs. 30-day
  - vs. Other endpoints
  - accounting for DNR status

TRENDS IN LOS AND RATES OF EARLY DNR



Summary, cont'd - Increasing Mortality for Hospitalized Stroke Patients in Cleveland

3. **Early DNR decisions dramatically influenced mortality rates:**
   - **1 in 5 patients documented as Early DNR**
   - **26% increase in the rate of Early DNR decisions**
   - **60% of all deaths by 30 days**
4. **Increasing stroke mortality rates in Cleveland between 1991-1997 are explained by changes in practice patterns - decisions re: Length of Stay and Early DNR - the appropriateness of which are not certain.**

Conclusions Re: Endpoints & Risk-adjustment for Stroke in Cleveland

1. **If mortality is used to reflect Quality of Care for hospitalized stroke patients:**
   - **30-day, not in-hospital**
   - **must account for Early DNR decisions**
2. **More research is necessary to determine whether the observed increase in severity-adjusted mortality represents poorer care; more efficient care; or more compassionate care.**

Cleveland Health Quality Choice:
In Memorium
1991-1998

From Lisa Iezzoni[1]:

"Designing a clinically reasonable but logistically feasible risk adjustment method is challenging and demands trade-offs. The accumulated evidence suggests that strong inferences about quality should not be made on risk-adjusted mortality rates alone. However, the risk of not risk adjusting is that information - albeit imperfect - will be summarily dismissed. Opportunities will be lost for stimulating the introspection required to improve quality of care."

"The Risks of Risk Adjustment". *JAMA.* 1997; 278:1600-1607.