College of Population Health Lectures, Presentations, Workshops

Jefferson College of Population Health

2-18-2011

Detecting Hospital Quality with PRIDIT.

Robert D. Lieberthal, PhD
School of Population Health, Thomas Jefferson University

Follow this and additional works at: https://jdc.jefferson.edu/hplectures

Part of the Health and Medical Administration Commons, and the Public Health Commons

Let us know how access to this document benefits you

Detecting hospital quality with *PRIDIT*

Robert D. Lieberthal

Jefferson School of Population Health

February 18, 2011

I could not have done this alone

- *PRIDIT* code from Richard Derrig and the Automobile Insurance Board of Massachusetts
- Funding from the Agency for Healthcare Research and Quality
- Lots of great input
    - Amol Navathe, Dan Polsky, Rachel Werner, and two anonymous referees at HSR
    - Attendees at the 2008 NRSA trainees conference, 2008 International Conference on Health Policy Statistics
    - Elaine Yuen and the students of PBH609

Outline

Outline

Hospital quality measures are hard to aggregate

- ▶ Hospital quality is assessed on process and outcome measures
  - ▶ Process measures include appropriate antibiotic use, frequent hand washing
  - ▶ Outcome measures include 30 day readmission rates, risk adjusted mortality
- ▶ It is hard to determine which observed measures of quality are good indicators of high quality hospitals
  - ▶ What is the relative importance of different measures?
  - ▶ How can we account for hospital characteristics like teaching status and ownership type?
  - ▶ Apparent high performance of hospitals could be a result of locating near a healthy population
- ▶ Quality should measure how much a hospital can improve a patient's health, not how healthy she was to begin with

Hospital Compare contains publicly reported hospital process measures

Process measure	Average		Jefferson	
	US	PA	Adherence	Patients (N)
Antibiotic timing	87%	88%	82%	303
Correct antibiotic	93%	93%	98%	302

Table: Hospital compare sample data, 7/1/2009-12/31/2009

Both measures contain some discretion

I used process measures and hospital characteristics

- 20 process measures of adherence to best practices from 0-100% at a single point in time
    - Heart attack (8 measures)
    - Heart failure (4 measures)
    - Pneumonia (6 measures)
    - Surgical infection prevention (2 measures)
- 5 other demographic variables from American Hospital Association data
    - Acute care or critical access hospital
    - Hospital ownership (govt, nfp, fp)
    - Emergency services
    - Accreditation
    - Teaching intensity (several levels)

Reporting data is optional for some hospitals, mandatory for others
(or Medicare would reduce their payments)
My sample included 4,217 hospitals that reported data

Heart attack measures contained lots of variation

Measure	Percent reporting	Average adherence
ACE inhibitor or ARB for LVSD	73%	80%
Aspirin at arrival	87%	92%
Aspirin at discharge	85%	89%
β-blocker at arrival	87%	85%
β-blocker at discharge	85%	87%
PCI $<$ 120 minutes post arrival	30%	64%
Smoking cessation	65%	79%
Thrombolytics $<$ 30 minutes post arrival	41%	30%

Heart failure quality measures were well reported

Measure	Average (%)	reporting (%)
ACE inhibitor or ARB for LVSD	80	89
Assessment of left ventricular function	80	93
Discharge instructions	52	83
Smoking cessation	74	81

Pneumonia measures were well reported

Measure	Percent reporting	Average adherence
Pneumococcal vaccination	94%	56%
Antibiotic(s) < 4 hours after arrival	93%	77%
Oxygenation assessment	94%	99%
Smoking cessation	83%	71%
Appropriate antibiotic(s)	84%	78%
Blood culture before antibiotic	84%	82%

Surgical infection measures were not well reported

Measure	Percent reporting	Average adherence
Antibiotic 1 hour before incision	35%	74%
Antibiotic stopped < 24 hours post surgery	35%	67%

Outline

Result is overall score

- ▶ Output on quality of hospitals and value of different variables
  - ▶ A relative ranking of all 4,217 hospitals in the dataset
  - ▶ A weighting system for the relative importance of quality indicators, demographic variables
- ▶ Example: Temple University Hospital scored 0.01419 (national average is 0)
- ▶ Example: Heart failure measure *patients given assessment of left ventricular function* was weighted 0.69731 (maximum score is 1)
- ▶ No negative weights
  - ▶ All measures were associated with positive quality–sometimes focus on one measure in isolation can hurt overall quality
  - ▶ If I had recoded the hospital characteristics, they would have been negative
- ▶ Small hospital bias caveats
  - ▶ Volume is not included
  - ▶ Hospitals did not report measures with N<25 observations
  - ▶ I imputed an average value for unreported variables

A few variables accounted for most of the variation in quality

- ▶ Patients given beta-blocker at arrival and at discharge
  - ▶ Well reported (∼85%)
  - ▶ Majority but not total adherence (∼85%)
- ▶ All 4 heart failure measures (esp. assessment of left ventricular function)
- ▶ Measures with total adherence not useful for measuring quality
  - ▶ Oxygen assessment for pneumonia–99% adherence!
- ▶ Surgical measures not well reported and so did not explain much variation
- ▶ All process measures positively associated with quality
- ▶ More teaching is better–no residency programs < some residency programs < full residency programs < residency and med school program

Heart attack measures were good indicators

Ranked bins from 1 (best) to 5 (worst)

	Full data set Importance	Clinical data only Importance
ACE inhibitor or ARB for LVSD	3	2
Aspirin at arrival	2	1
Aspirin at discharge	2	1
β-blocker at arrival	1	1
β-blocker at discharge	1	1
PCI $<$ 120 minutes post arrival	4	3
Smoking cessation	2	1
Thrombolytics $<$ 30 minutes post arrival	5	4

Heart failure measures were good indicators

	Full data set Importance	Clinical data only Importance
ACE inhibitor or ARB for LVSD	2	1
Assessment of left ventricular function	1	1
Discharge instructions	1	1
Smoking cessation	1	1

Pneumonia and surgery weren't as good

	Full data set Importance	Clinical data only Importance
Pneumonia		
Pneumococcal vaccination	2	1
Antibiotic(s) < 4 hours after arrival	5	3
Oxygenation assessment	4	3
Smoking cessation	2	1
Appropriate antibiotic(s)	3	2
Blood culture before antibiotic	4	3
Surgical infection prevention		
Antibiotic 1 hour before incision	3	3
Antibiotic stopped < 24 hours post surgery	5	4

Hospital demographics were good controls

|	All data Importance	Clinical data Importance
Acute care hospital	5	N/A
Government hospital	5	N/A
Private hospital	5	N/A
Accredited hospital	4	N/A
Emergency service available	5	N/A
Major teaching	4	N/A
Significant teaching	3	N/A
Any teaching	3	N/A

Hospital quality was evenly distributed
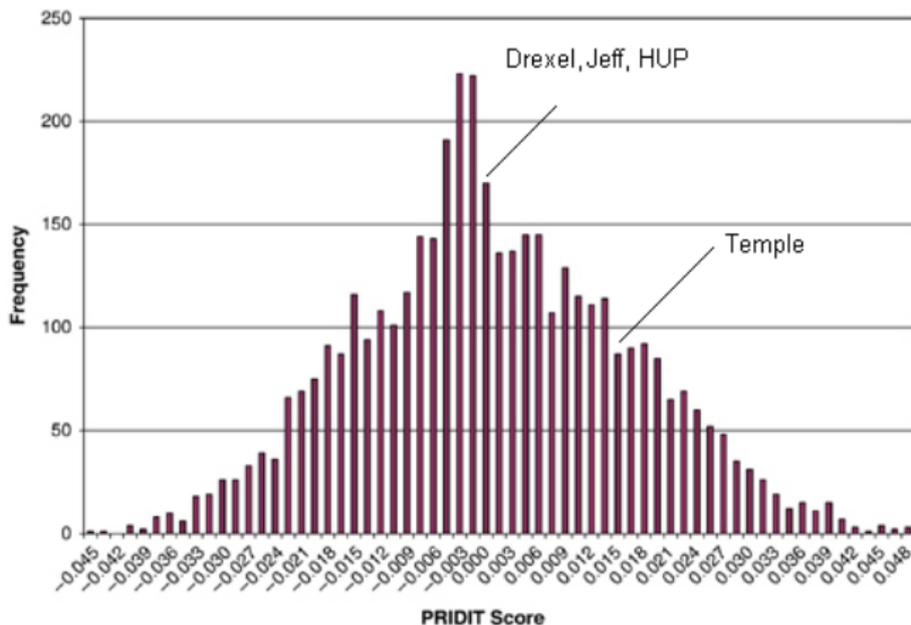
Lots of hospitals in the middle, a few "outliers" of high and low quality

How well does *PRIDIT* distinguish between hospitals?

- ▶ What is the difference between scoring 0.05 and 0.10?
- ▶ What is the value in raising the score from 0.05 to 0.10?
  - ▶ Pay for performance and provider networks–are they different enough to pay one more than the other?
- ▶ How does the standard error change across the range of *PRIDIT* scores?
- ▶ Importance differs by geography
  - ▶ Some areas only have one hospital–see if you need to make it better
 - ▶ Some areas have lots of hospitals–direct people to the best one

Sampling *PRIDIT* scores with replacement

- 1000 samples with replacement of the universe of scores
- Not bagging–I didn't recompute the scores in each sample
- U-shaped distribution of standard errors–much more confidence in median than high or low scores
- Quintile cutoffs: -0.012, -0.003, 0.0045, 0.0135
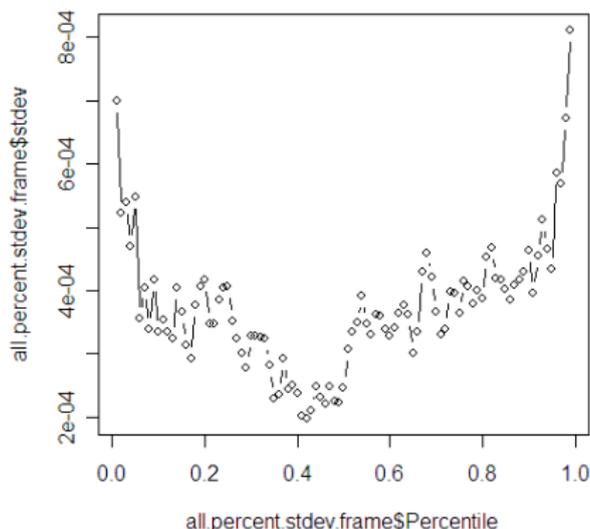
Standard errors are small relative to *PRIDIT* scores

- ► You can tell many Philadelphia hospitals apart within a 2 sd range (one sd in each direction)
- ► Hahneman, Presby and Jefferson are all close (Frankford and Nazareth are also only 4 miles apart)
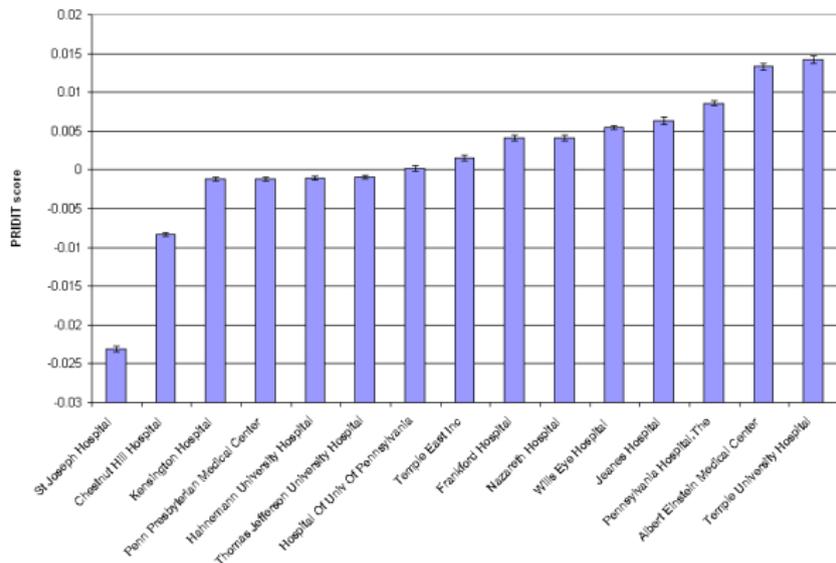
Validating *PRIDIT* with outcome measures in Hospital Compare

- Better scores should correlate with less readmission and mortality
- Readmission
  - Readmission is 30-day readmission rate
  - Theory: they should have fixed you the first time
  - Less than perfect indicators
    - Not a long look back period
    - If people die, they are never readmitted
- Mortality
  - Mortality is risk adjusted
  - Theory: they should have fixed you well enough so you don't die
  - Less than perfect indicators
    - Not a long look back period
    - It would be better to have raw scores and do the adjustments myself

Using multiple observations from Hospital Compare

- Multiple observations of the same hospital over time
  - Measure the stability of hospital rankings over time
  - Measure the relative importance of each measurement over time
  - More hospitals report more data over time
  - Hospital Compare has more measures now than when I first studied these hospitals
  - Use outcomes measures over time to rank quality and/or validate quality scores
- Measure the bootstrapped confidence intervals against multiple observations

Questions/comments/feedback

The extensions are works in progress. I have submitted the validation idea as a grant proposal to the Actuarial Foundation. Any feedback would be really useful to me. Thanks!