Public Reporting of Surgical Outcomes

Benjamin D. Kozower, MD, MPH

Professor of Surgery

Co-Director, Surgical Outcomes Research Center

Washington University

St. Louis, MO

Disclosures

Chair, STS General Thoracic Surgery Database
Member, STS Quality Measurement Taskforce

Disclaimers

"All models are wrong, some models are useful"

George Box

History of Outcomes Research & Public Reporting

"The trustees of hospitals should see to it that an effort is made to follow up each patient they treat, long enough to determine whether the treatment given has permanently relieved the condition or symptoms complained of."

Dr. Ernest Amory Codman, 1918

Problems with Current Methods for Ranking Surgeons and Hospitals …..

" I have just heard that 50% of United Kingdom cardiac surgeons are below average, ... this has got to stop!"

The Rt Hon Frank Dobson
Former Labor Secretary for Health

To Your Health

Need a new knee? Heart valve? Back surgery? This Web site could help you find the top surgeons near you.

By Lena H. Sun July 14 Follow @bylenasun

SURGEONRATINGS.ORG by CONSUMERS' CHECKBOOK

Surgeons with Good Results for

Gallbladder Removal Surgery

Don't see your doctor? Here's why

Sort by Distance

5 surgeons found within 50 miles of zip code 22901 - Search Again

Click here to see how big the differences are

Advice & Explanations

Select to Compare	Name (Click doctor name for more information)	Better Outcomes: Fewer Deaths	Better Outcomes: Fewer Deaths or Other Bad Outcomes	Recommended Most By Surveyed Doctors	Board Certifications	Surgeon's Hospitals (✓ = Hospital high-rated for outcomes for this surgery or for this surgeon)
☐	Adams, Reid B. Lee St Charlottesville, VA 22908 434-924-0000 (5.6 miles)	★★★	---	---	• Surgery	University of Virginia Medical Center
☐	Faulkenberry, William L. 70 Medical Center Cir #213 Fishersville, VA 22939 540-332-5999 (24.0 miles)	---	★★★	---	• Surgery	Augusta Health
☐	Branum, Gene D. 2006 Health Campus Dr Harrisonburg, VA 22801 540-689-5800 (30.0 miles)	---	★★★	---	• Surgery	Sentara RMH Medical Center
☐	Garwood, Robert Allen 2006 Health Campus Dr Harrisonburg, VA 22801 540-689-5800 (30.0 miles)	---	★★★	---	• Surgery	Sentara RMH Medical Center
☐	Oates, Thomas M. 2006 Health Campus Dr Harrisonburg, VA 22801 540-689-5800 (30.0 miles)	---	★★★	---	• Pediatrics • Surgery	Sentara RMH Medical Center Danville Regional Medical Center

How confidently we say the surgeon's results are better-than-average: Very confidently = ★★★★★ Confidently = ★★★★ Probably = ★★★

Surgeon Scorecard

by Sisi Wei, Olga Pierce and Marshall Allen, ProPublica, Updated July 15, 2015

Guided by experts, ProPublica calculated death and complication rates for surgeons performing one of eight elective procedures in Medicare, carefully adjusting for differences in patient health, age and hospital quality. Use this database to know more about a surgeon before your operation.

READ OUR STORY
Making the Cut: Why Choosing the Right Surgeon Matters Even More Than You Know

METHODOLOGY
Read how we calculated complications and the key questions we considered.

EDITOR'S NOTE
Why ProPublica is naming surgeons and what experts are saying about it

Find Near Me | Find a Surgeon | Find a Hospital

Surgeons and Hospitals Near My Location

ex: 155 6th Avenue, New York, New York

Knee Replacement

Use My Location

...or jump straight to your state: Pick a state

Eight Elective Procedures

We focused on procedures done thousands of times a day, mostly without incident. They are scheduled in advance and generally performed on patients in stable health. We excluded patients who

Surgeons, Not Hospitals

Conventional wisdom tells patients to simply choose a good hospital when they need surgery. But ProPublica has found that even within "good" hospitals, performance between surgeons can

Public Reporting of Surgical Outcomes

- Routine expectation of stakeholders
  - Patients, payers, legislators, policy makers
- Ethical responsibility
  - Accountability
  - Affirm patient autonomy
- Goals
  - Education of consumers
  - Improve patient care (2 mechanisms)
  - Improve resource utilization

Public Reporting of Surgical Outcomes: Improving Quality

NY State Public Reporting for CABG

Year	**1989**	**1992**
In-hospital mortality	3.5%	2.7%
Expected mortality (risk-adjusted)	2.9%	3.2%

Hannan EL et al. J Am Coll Cardiol. 2012;59

Public Reporting of Surgical Outcomes: Improving Quality

Society of Thoracic Surgeons Database : “A Decade of Change – Risk Profiles and Outcomes for CABG”

Year	**1990**	**1999**
Operative mortality	3.9%	3.0%
Expected mortality (risk-adjusted)	2.6%	3.4%

Ferguson BT et al. Ann Thorac Surg. 2002;73

Public Reporting of Surgical Outcomes

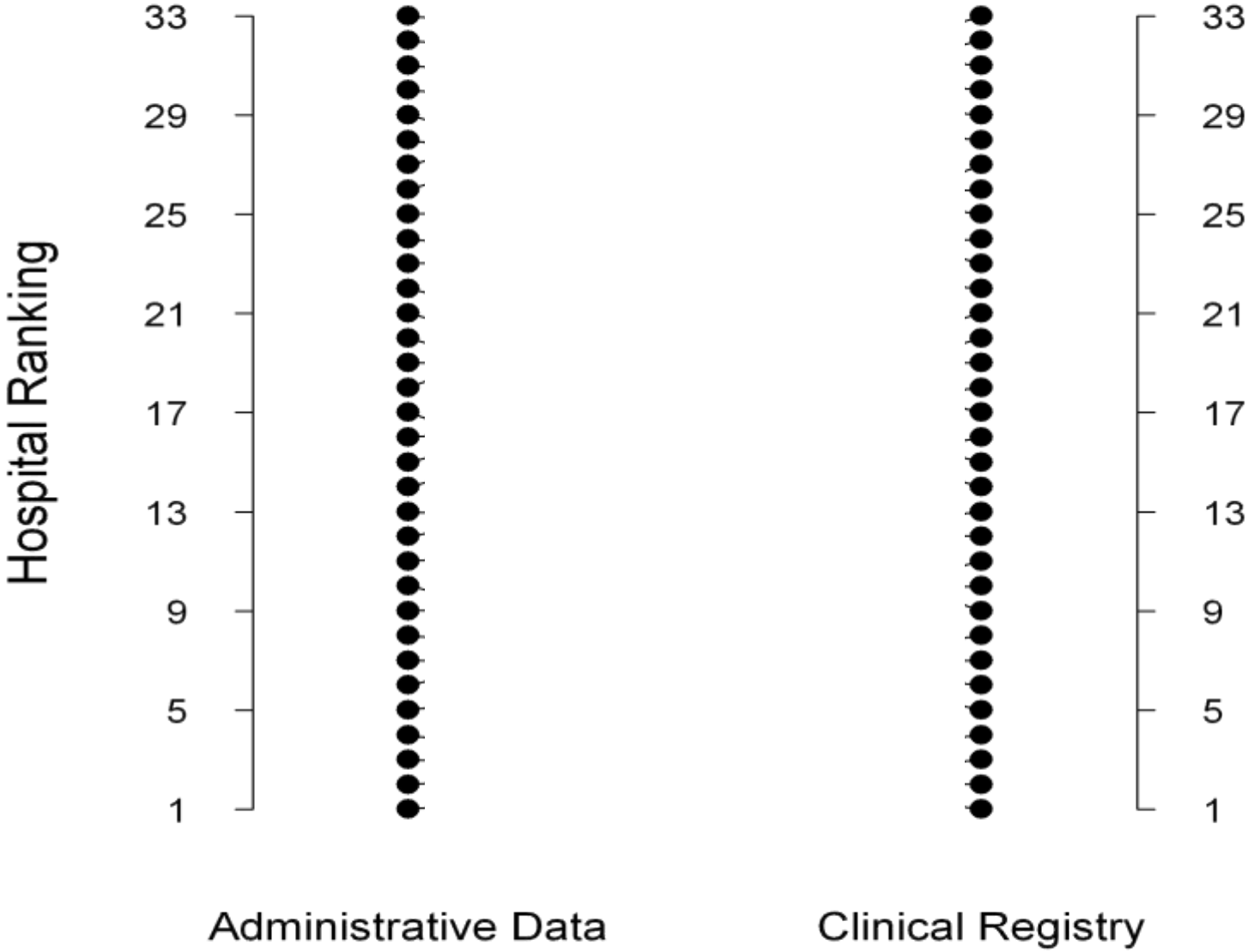
- Opportunity
  - Currently being done with poor methodology and administrative data
  - Need more relevant, patient-centered and long-term outcomes
  - Need robust methodologies to compare outcomes, “quality”, and performance
- Need for transparency

Barriers to Public Reporting:
Requirements for Comparing Institutions

- We don’t have robust, accurate data
- We don’t have relevant outcomes measures
- We need sufficient statistical power to identify meaningful differences
- We don’t know how to communicate effectively to consumers

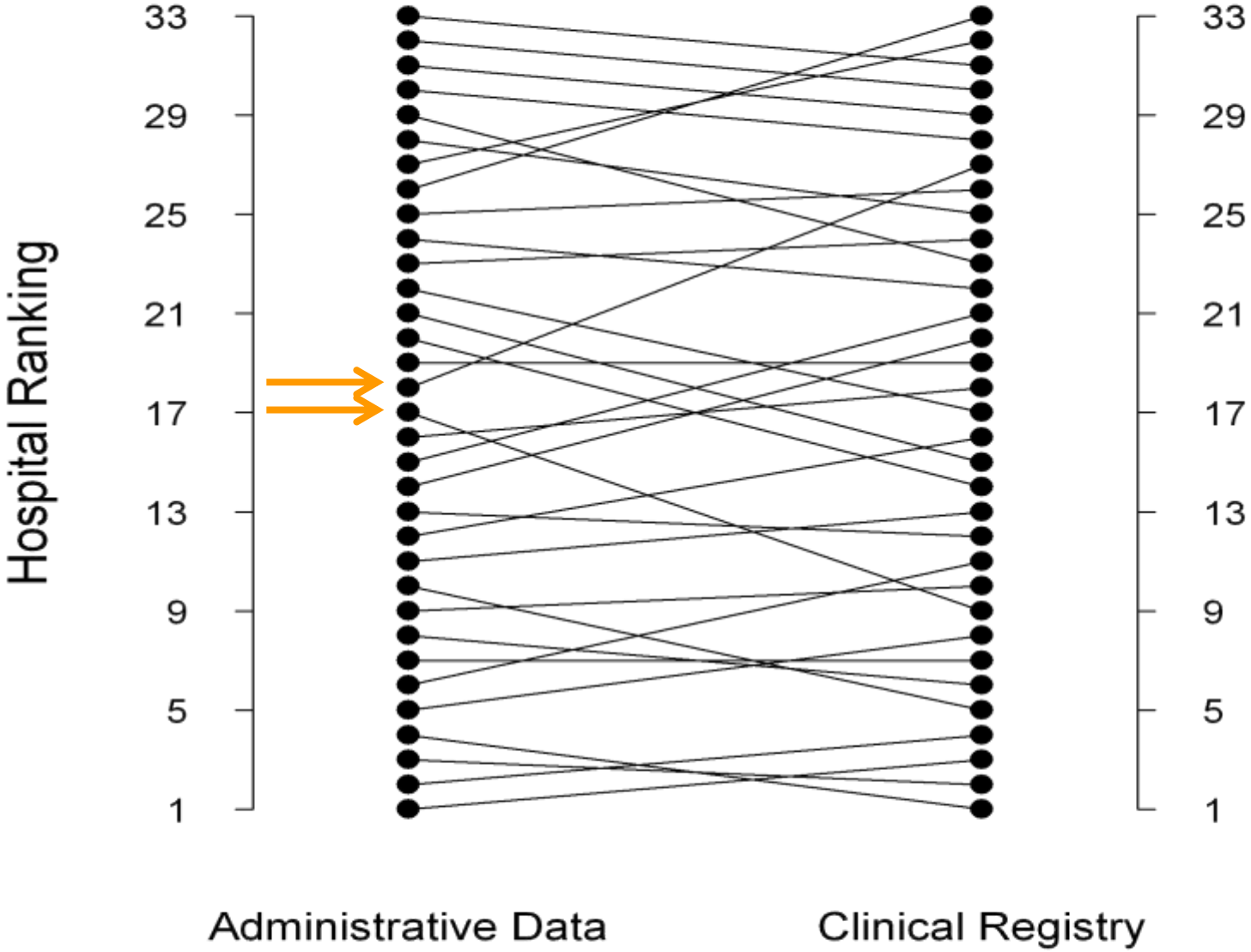
Lansky D, Health Affairs April, 2012
Gaynor JW, With Permission 2015

Impact of Case Ascertainment Methodology on Hospital Rankings



Pasquali SK et Ann Thorac Surg. 2013

Impact of Case Ascertainment Methodology on Hospital Rankings



Pasquali SK et Ann Thorac Surg. 2013

Data Source: Administrative vs. Clinical

ORIGINAL SCIENTIFIC ARTICLES

Predicted Risk of Mortality Models: Surgeons Need to Understand Limitations of the University HealthSystem Consortium Models

Benjamin D Kozower, MD, MPH, FACS, Gorav Ailawadi, MD, FACS, David R Jones, MD, FACS, Robert D Pates, PhD, Christine L Lau, MD, FACS, Irving L Kron, MD, FACS, George J Stukenborg, PhD

BACKGROUND: The University HealthSystem Consortium (UHC) mortality risk adjustment models are increasingly being used as benchmarks for quality assessment. But these administrative database models may include postoperative complications in their adjustments for preoperative risk. The purpose of this study was to compare the performance of the UHC with the Society of Thoracic Surgeons (STS) risk-adjusted mortality models for adult cardiac surgery and evaluate the contribution of postoperative complications on model performance.

STUDY DESIGN: We identified adult cardiac surgery patients with mortality risk estimates in both the UHC and Society of Thoracic Surgeons databases. We compared the predictive performance and calibration of estimates from both models. We then reestimated both models using only patients without any postoperative complications to determine the relative contribution of adjustments for postoperative events on model performance.

RESULTS: In the study population of 2,171 patients, the UHC model explained more variability (27% versus 13%, $p < 0.001$) and achieved better discrimination (C statistic = 0.88 versus 0.81, $p < 0.001$). But when applied in the population of patients without complications, the UHC model performance declined severely. The C statistic decreased from 0.88 to 0.49, a level of discrimination equivalent to random chance. The discrimination of the Society of Thoracic Surgeons model was unchanged (C statistic of 0.79 versus 0.81).

CONCLUSIONS: Although the UHC model demonstrated better performance in the total study population, this difference in performance reflects adjustments for conditions that are postoperative complications. The current UHC models should not be used for quality benchmarks. (J Am Coll Surg 2009;209:551–556. © 2009 by the American College of Surgeons)

Comparing Administrative & Clinical Data: Apples & Oranges

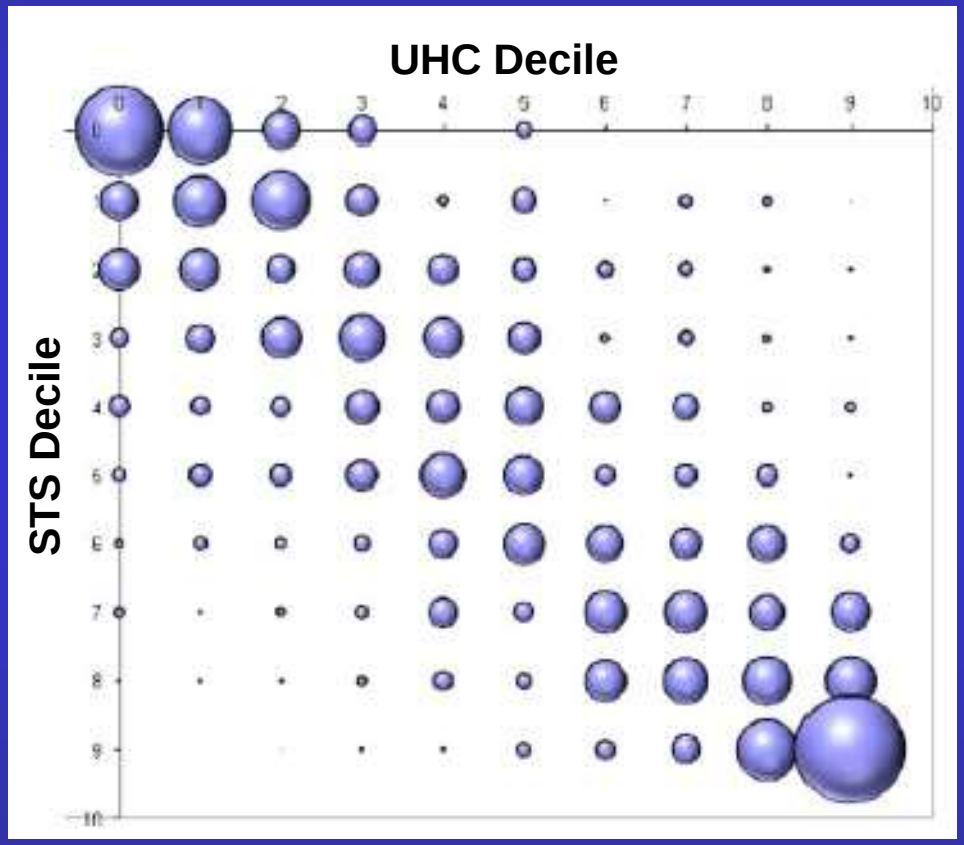
- The University HealthSystem Consortium (UHC)
  - Includes all patients in the cardiovascular service line
  - Decile rankings (no statistical reporting)

- The Society of Thoracic Surgeons (STS)
  - Includes patients operated on by cardiac surgeons
  - Risk-adjustment for common procedures
  - Tercile rankings (1-3 stars)

Kozower BD et al. JACS. 2009

Comparing Administrative & Clinical Data

- 28% difference in cardiac surgery patient populations
- STS appropriately uses preoperative risk factors
- UHC includes all diagnoses, including postoperative complications



The Society of Thoracic Surgeons

Google™ Custom Search | Search

Login

Surgeons | Residents & Students | Allied Health | Patients | Media

- Home
- About STS
- Membership
- Education & Meetings
- **STS National Database**
  - Database Participants
  - Data Managers
  - AC Anesthesia Module
  - STS/ACC TVT Registry
- Quality, Research & Patient Safety
- Advocacy
- Publications & Resources

STS National Database

The STS National Database was established in 1989 as an initiative for quality improvement and patient safety among cardiothoracic surgeons. There are three components to the STS National Database, each focusing on a different area of cardiothoracic surgery—Adult Cardiac, General Thoracic, and Congenital Heart Surgery, with the availability of Anesthesiology participation within the Congenital Heart Surgery Database. The Database has grown exponentially over the years, both in terms of participation and stature. Learn more in the STS National Database Brochure.

- Gold standard clinical database
- Adult Cardiac, General Thoracic, Congenital
- 95% data accuracy, externally audited
- Predictive risk models for common procedures

Barriers to Public Reporting:
Requirements for Comparing Institutions

- We don’t have robust, accurate data
- We don’t have relevant outcomes measures
- We need sufficient statistical power to identify meaningful differences
- We don’t know how to communicate effectively to consumers

Lansky D, Health Affairs April, 2012
Gaynor JW, With Permission 2015

The Swedish Approach to Hip Replacement

- Reported health gain (EQ-5D index) after one year
- Patient satisfaction at one year
- Complications at two years
- Ten-year implant survival

- None of these measures are reported in the US
- > 90% follow up publicly reported in Sweden

Swedish Hip Arthroplasty Register 2015
http://www.shpr.se/en/

Thoracic Surgery Outcomes: Measuring Hospital Performance

J. MAXWELL CHAMBERLAIN MEMORIAL PAPER FOR GENERAL THORACIC SURGERY

STS Database Risk Models: Predictors of Mortality and Major Morbidity for Lung Cancer Resection

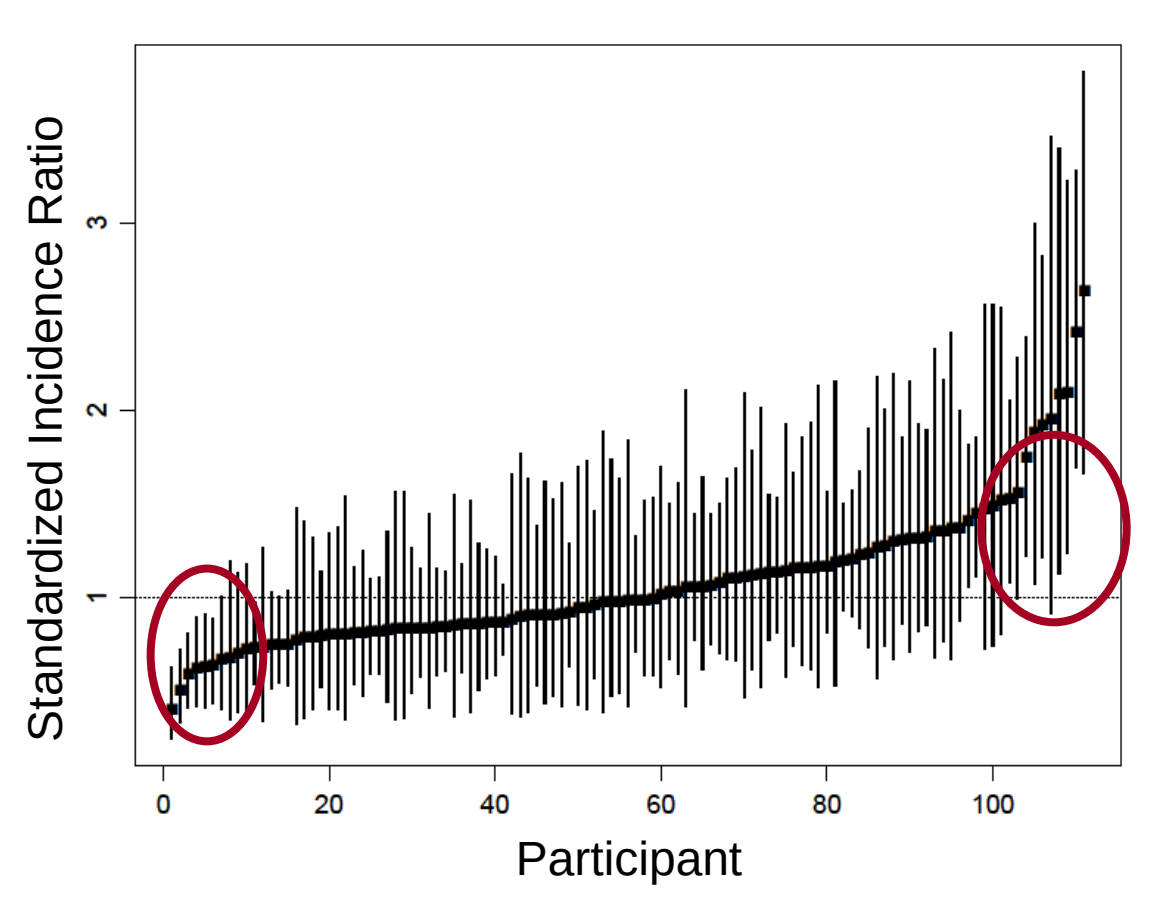
Benjamin D. Kozower, MD, MPH, Shubin Sheng, PhD, Sean M. O'Brien, PhD, Michael J. Liptay, MD, Christine L. Lau, MD, David R. Jones, MD, David M. Shahian, MD, and Cameron D. Wright, MD

Departments of Surgery & Public Health Sciences, University of Virginia Health System, Charlottesville, Virginia; Duke Clinical Research Institute, Duke University, Durham, North Carolina; Department of Cardiovascular and Thoracic Surgery, Rush University, Chicago, Illinois; and Division of Thoracic Surgery and Center for Quality and Safety, Massachusetts General Hospital, Boston, Massachusetts

Kozower. Ann Thorac Surg. 2010;90

Hospital Performance Variation:

Mortality or Major Morbidity – STS Database
18,800 lung cancer resections from 111 hospitals



Kozower et al. Ann Thorac Surg. 2010;90

Thoracic Surgery Outcomes: Measuring Hospital Performance

- Successfully measured and compared 30 day composite outcomes

- Need long term survival for cancer care
- Need patient centered outcomes

Kozower. Ann Thorac Surg. 2010;90

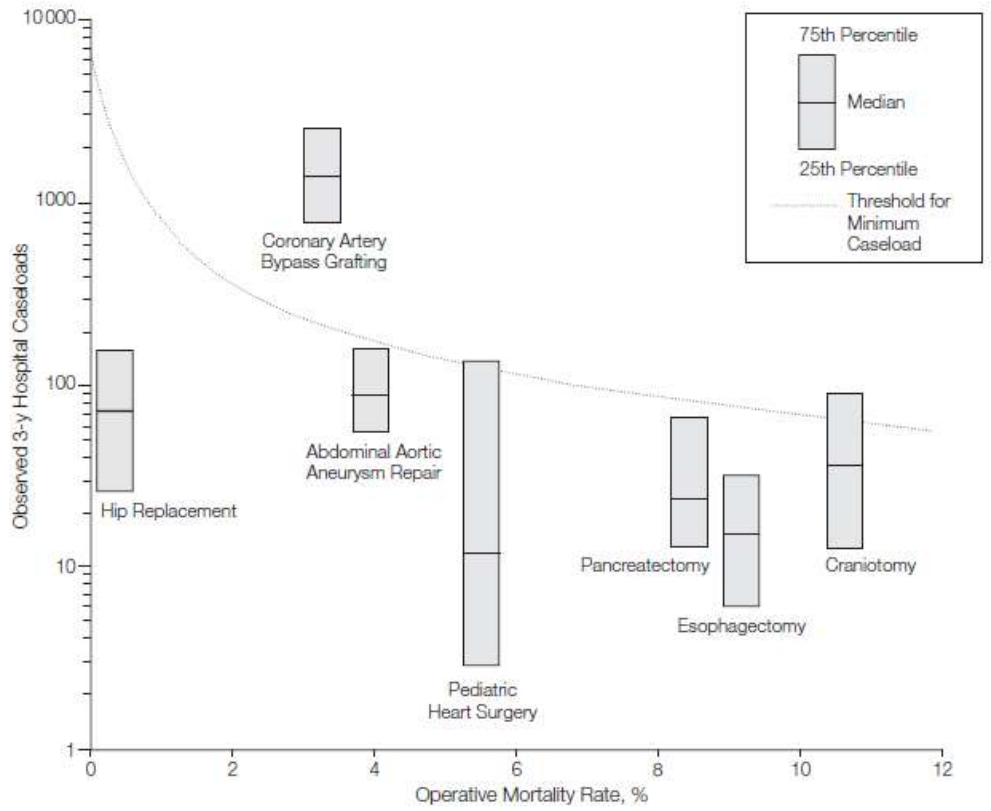
Barriers to Public Reporting:
Requirements for Comparing Institutions

- We don’t have robust, accurate data
- We don’t have relevant outcomes measures
- We need sufficient statistical power to identify meaningful differences
- We don’t know how to communicate effectively to consumers

Lansky D, Health Affairs April, 2012
Gaynor JW, With Permission 2015

Surgical Mortality as an Indicator of Hospital Quality:
The Problem with Sample Size

Figure. Distribution of Actual Hospital Caseloads and the Minimum Caseload Needed to Detect a Doubling of the Mortality Rate



Area above the dotted line indicates the minimum caseloads exceeded for a given mortality rate.

Only CABG has the case volumes to reach the threshold of 138 cases to detect a doubling of the mortality rate.

Dimick et al JAMA 2004;292:847

SURGEONRATINGS.ORG by CONSUMERS' CHECKBOOK

Surgeons with Good Results for
Pulmonary Surgery

Don't see your doctor?
Here's why

Sort by Distance

How confidently we say the surgeon's results are better-than-average: Very confidently = ★★★★★ Confidently = ★★★★ Probably = ★★★

- Medicare data 2009-2012
- 90-day mortality, readmission, complications
- Risk-adjustment methodology is unknown
- 5 categories, only report top three for now

Postoperative Mortality Is an Inadequate Quality Indicator for Lung Cancer Resection

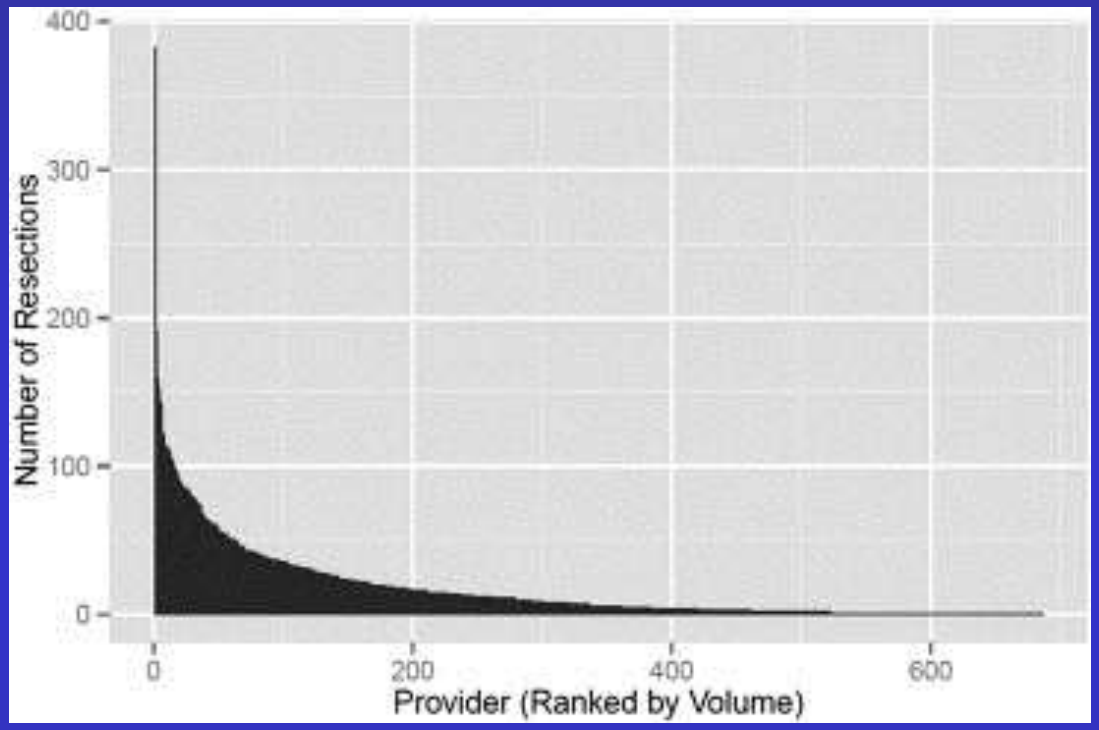
Yinin Hu, MD, Timothy L. McMurry, PhD, Kristen M. Wells, PhD, James M. Isbell, MD, MSCI, George J. Stukenborg, PhD, and Benjamin D. Kozower, MD, MPH

Department of Surgery, Division of Thoracic Surgery, and Department of Public Health Sciences, University of Virginia Health System, Charlottesville, Virginia

- Objectives
  - To compare operative mortality definitions
  - To determine if 90-day mortality can be used as a quality metric
- Methods
  - SEER-Medicare data 2006-2010
  - Hierarchical GLE models to estimate 30 and 90-day mortality

(Ann Thorac Surg 2014;97:973–9)

Results



-11,787 patients from 686 hospitals
- Provider volumes range from 1-383
- 33% of hospitals perform fewer than 5 resections (Medicare)

Results

Table 2. Postoperative Mortality Measures

Mortality	n (%)
In-hospital	355 (3.01)
30 day	435 (3.69)
Perioperative	482 (4.09)
90 day	812 (6.89)

Results : Provider Rankings

- Calculated observed/expected 90-day mortality estimates for 686 hospitals

- Hospital performance variation
  - 3% of hospitals (20/686) had an O/E 90-day mortality different than 1
    - 18 hospitals had O/E > 1
    - 2 hospitals had O/E < 1

Study Conclusions

- Currently reported measures of in-hospital and 30-day postoperative mortality do not adequately represent a patient's true mortality risk
  - Mortality almost doubles by 90 days

- 90-day mortality <u>is not</u> a useful metric for evaluating hospitals following lung cancer resection

SURGEONRATINGS.ORG by CONSUMERS' CHECKBOOK

Surgeons with Good Results for
Pulmonary Surgery

Don't see your doctor?
Here's why

Sort by Distance

How confidently we say the surgeon's results are better-than-average: Very confidently = ★★★★★ Confidently = ★★★★ Probably = ★★★

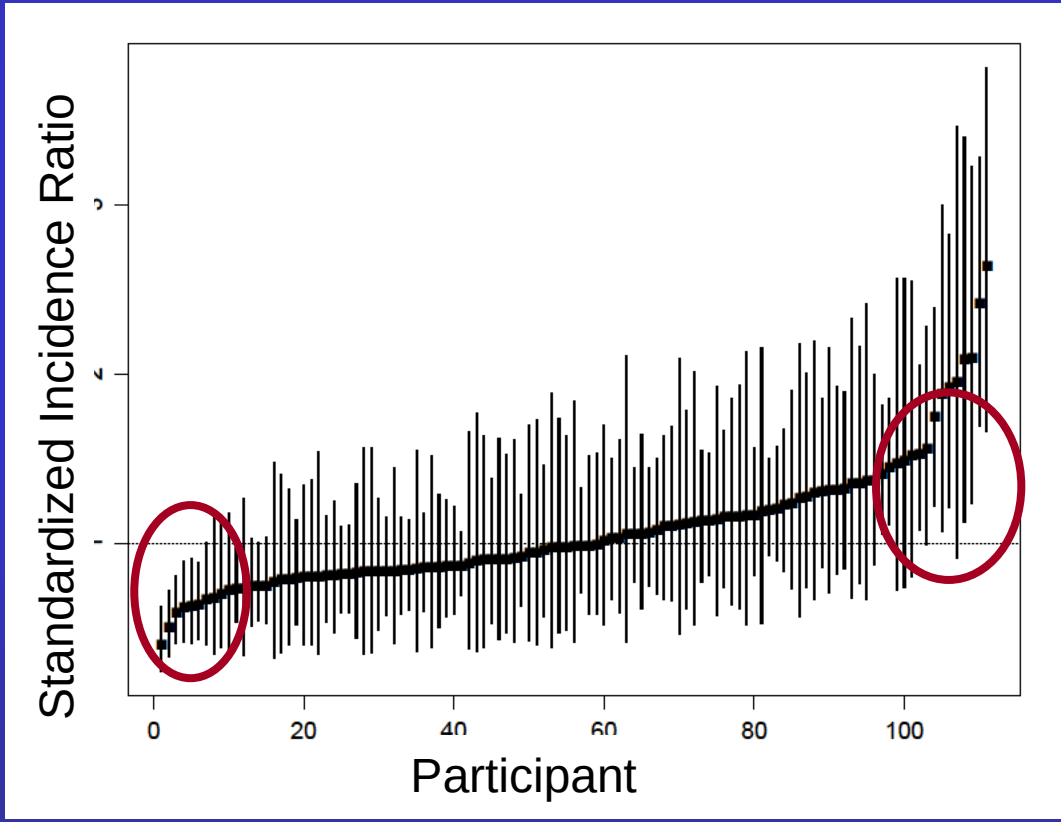
How accurate are surgeon level ratings for 90-day mortality when hospital level ratings only identify 3% of programs as different than average?

How is “Probably” defined and is it sufficient for public reporting?

Increasing Sample Size: Composite Measures:

Mortality or Major Morbidity – STS Database

18,800 lung cancer resections from 111 hospitals



Kozower et al. Ann Thorac Surg. 2010;90

Barriers to Public Reporting:
Requirements for Comparing Institutions

- We don’t have robust, accurate data
- We don’t have relevant outcomes measures
- We need sufficient statistical power to identify meaningful differences
- We don’t know how to communicate effectively to consumers

Lansky D, Health Affairs April, 2012
Gaynor JW, With Permission 2015

Communicating Outcomes with Consumers

- 60 million Americans use Yelp and dry cleaner ratings every month
- Healthcare is a very common search, but far from the top 10
  - 66% of internet users performed a healthcare search in the past year

Freudenheim M. New York Times. February 11, 2014

PATIENT ENGAGEMENT

By J. Matthew Austin, Ashish K. Jha, Patrick S. Romano, Sara J. Singer, Timothy J. Vogus, Robert M. Wachter, and Peter J. Pronovost

DOI: 10.1377/hlthaff.2014.0201
HEALTH AFFAIRS 34,
NO. 3 (2015): 423–430

National Hospital Ratings Systems Share Few Common Scores And May Generate Confusion Instead Of Clarity

Leapfrog Group, Consumer Reports Health, US News & World Report's, HealthGrades

- Examined 844 hospitals
- Four Rating Systems: Leapfrog, Consumer Reports, US News and World Report's, HealthGrades
- Objectives were to compare the criteria used in ratings systems and to determine the consistency of hospital ratings

EXHIBIT 4

Overlap Of Hospitals Rated As High Performers And Low Performers Across The Four Hospital Rating Systems

No. of ratings that overlap	No. of hospitals rated as high performers	Details of the high-performer overlap
HIGH PERFORMERS		
On all four ranking systems	0	—
On three of four ranking systems	3	2 rated by LF, CR, and HG; 1 rated by LF, US, and HG
On two of four ranking systems	85	45 rated by LF and HG; 7 rated by LF and US; 29 rated by LF and CR; 0 rated by CR and US; 3 rated by CR and HG; 1 rated by US and HG
LOW PERFORMERS[a]		
On all three ranking systems	0	—
On two of three ranking systems	15	7 rated by LF and CR; 5 rated by LF and US; 3 rated by US and CR

SOURCE Leapfrog Group, *Consumer Reports* Health, *U.S. News & World Report*'s Best Hospitals 2013–14, and HealthGrades America's Best Hospitals 2013 Report. **NOTE** LF is Leapfrog; CR is *Consumer Reports*; US is *U.S. News & World Report*; and HG is HealthGrades. [a]HealthGrades was excluded from the low performers because it lists only top hospitals and makes no distinctions among the lower-performing hospitals.

- No hospital was rated as a high performer by all 4 rating systems
- Each system has its own rating methods, a different focus to its ratings, and stresses different measures

Consumer Comprehension of Surgeon Performance Data for Coronary Bypass Procedures

Karen Donelan, ScD, Robert S. Rogers, BA, Andy Eisenhauer, MD, Elizabeth Mort, MD, MPH, and Arvind K. Agnihotri, MD

Mongan Institute for Health Policy, Massachusetts General Hospital, Boston; Division of Cardiology, Brigham and Women's Hospital, Boston; Department of Surgery, Division of Cardiac Surgery, Massachusetts General Hospital Heart Center, Boston; Harvard Medical School, Boston; and Department of Quality and Safety, Massachusetts General Hospital, Boston, Massachusetts

- Examined 4 different public reporting displays
  - 3 of 4 are used by publically available state reports
- Displays shown to 337 adults
- Respondents asked to identify which surgeon they were most likely to choose and surveyed on the importance of public reporting

Donelan K et al. Ann Thorac Surg. 2011;91

This is the "Best" State Display:
Which Surgeon Would You Choose?

		Mortality		Readmissions		
Surgeon	Cases	In-Hospital	30-Day	7-Day	30-Day	Length of Stay
1	92	◉	◉	◉	◉	5.8
2	100	◉	◉	◉	◉	6.8
3	89	○	◉	●	●	5.7
4	91	●	◉	◉	○	4.9
5	100	○	○	◉	◉	6.9

○ Lower than expected (meaning that the hospital or surgeon had fewer deaths than expected after accounting for how sick the patients were in that hospital)

◉ Same as expected (meaning that the hospital or surgeon had as many deaths as expected after accounting for how sick the patients were in that hospital)

● Higher than expected (meaning that the hospital or surgeon had more deaths than expected after for how sick the patients were in that hospital)

Donelan K et al. Ann Thorac Surg. 2011;91

Personal Importance of Public Ratings

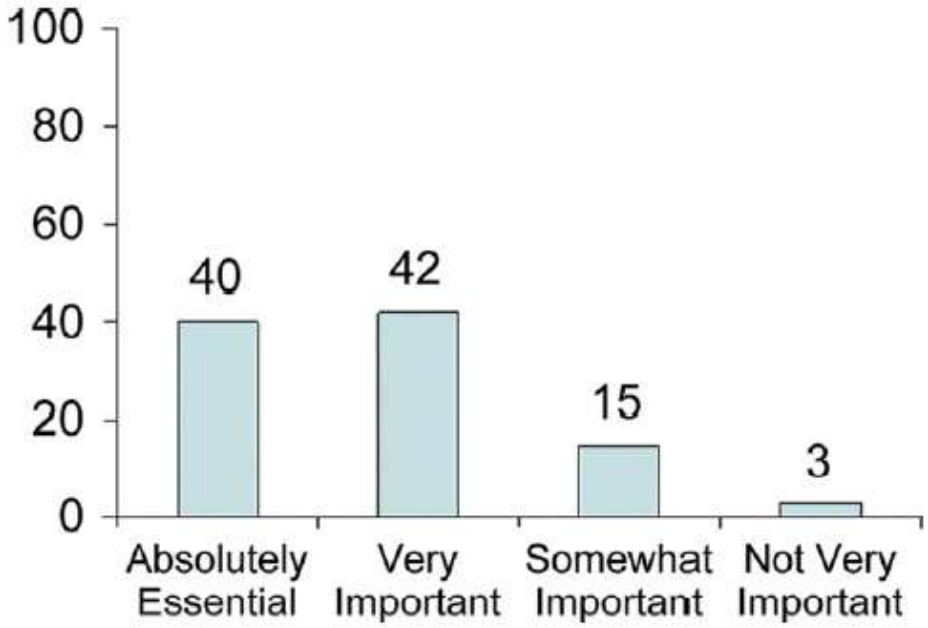


Fig 3. Personal importance of public ratings: "If you have to select a surgeon, how important is it to you to have data like these available to you?"

Donelan K et al. Ann Thorac Surg. 2011;91

Surgeon Preference by Display Format

Table 4. Surgery Results: Surgeon Preference for Display Type

	Surgeon 1 (%)	Surgeon 2 (%)	Surgeon 3 (%)	Surgeon 4 (%)	Surgeon 5 (%)	Need More Information (%)
Display A	14	44	16[a]	N/A	N/A	26
Display B	53[a]	15	1	1	6	25
Display C	7	3	3	4	66[a]	17
Display D	32	10	2	22[a]	13	21

Surgeon preference for each display is shown ("which surgeon would you choose given the information shown"). *Indicates the surgeon with the lowest risk-adjusted mortality in the display.

N/A = not available.

- The public does not understand current displays of reported outcomes
- The "best surgeon" was correctly identified between 13% and 66% of the time

Donelan K et al. Ann Thorac Surg. 2011;91

Public Reporting of Surgical Outcomes

The Society of Thoracic Surgeons

Google™ Custom Search | Search

Login

Surgeons | Residents & Students | Allied Health | Patients | Media

- Home
- About STS
- Membership
- Education & Meetings
- STS National Database
- **Quality, Research & Patient Safety**
  - Quality
  - Research
  - Patient Safety
  - STS Public Reporting Online
  - Guidelines
- Advocacy

Home / Quality, Research & Patient Safety / STS Public Reporting Online

STS Public Reporting Online

Heart Surgery Outcomes - Public Access

STS believes the public has a right to know the quality of surgical outcomes and considers public reporting an ethical responsibility of the specialty. STS Public Reporting Online enables STS National Database participants to voluntarily report to each other and the public their heart surgery composite star ratings and the component ratings from which those are derived.

Isolated CABG:

Search or browse star ratings for surgery groups

Search or browse star ratings for hospitals

Isolated AVR:

Search or browse star ratings for surgery groups

Search or browse star ratings for hospitals

In this section

- About STS Public Reporting
- History of the STS National Database
- The STS National Database Today
- Rationale for Public Reporting
- Explanation of Quality Rating, Composite Score, and Star Ratings
- Potential Unintended Consequences of Public Reporting
- STS Public Reporting Consent/Release Forms

The Society of Thoracic Surgeons Composite Score for Lobectomy for Lung Cancer

Benjamin D. Kozower, Mitchell J. Magee,
Sean M O'Brien, Andrzej S. Kosinski,
Rachel Dokholyan, Jeffrey P. Jacobs,
David M. Shahian, Cameron D. Wright,
Felix G. Fernandez

STS Quality Measurement Task Force
STS General Thoracic Surgery Database Task Force
Duke Clinical Research Institute

Kozower BD et al. Ann Thorac Surg. 2016;In press

Public Reporting of Surgical Outcomes

- Develop a platform for thoracic surgery
  - Start with lobectomy for lung cancer
  - Most common procedure
  - Use 3 years of rolling data
- Two domain composite
  - Operative mortality
  - Major complications

Methods: Estimation of Composite Scores & Star Ratings

- Composite score = (1-risk adjusted mortality rate) + (1-risk adjusted complication rate)
- Outcomes weighted inversely by their respective standard deviations
- Star ratings (3 categories)
  - High (3-star) and low (1-star) performers
    - 95% credible intervals outside average scores
  - Easy for the public to comprehend

Shahian DM et al. Ann Thorac Surg. 2014;97(5)

Results: Patient Population

22,362 lobectomies at 234 programs

Non-elective status (1,211)

Missing data (n=281)

Benign disease (n=122)

20,657 at 231 programs (92.4%)

Composite Outcomes

Operative mortality	1.5%
Major complications	9.6%
Pneumonia	4.3%
Unexpected return to OR	3.9%
Reintubation	3.4%
Pulmonary embolus	0.5%
Initial Vent Support >48 Hours	0.5%
Bronchopleural fistula	0.4%
Tracheostomy	1.0%
ARDS	0.7%
Myocardial infarction	0.4%
Length of stay (median)	4 days, IQR (3,7)

Composite Score Weights

	Mortality	Major Complications
Standard Deviation	0.84	3.42
Weight	0.80	0.20

Mortality is weighted approximately 4 times that of a major complication in the quality measure

Reliability of Star Ratings (Signal / Noise Ratio)

- Reliability is a key measure for profiling
- Describes the confidence in distinguishing performance between providers
- Signal is the proportion of variability in measured performance explained by real differences in performance
- Noise is the random statistical variation

Adams J. 2009. RAND – www.rand.org)

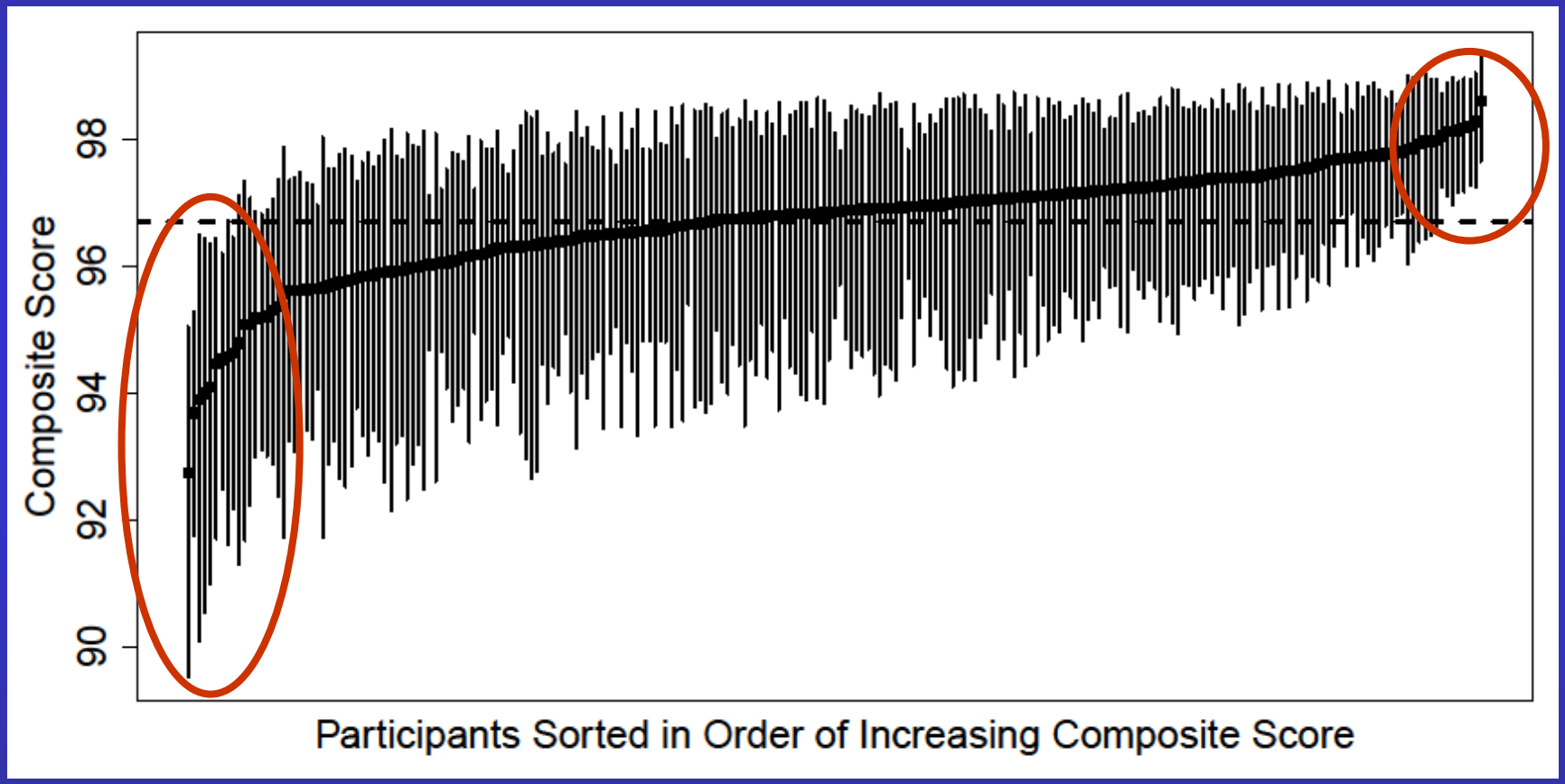
Reliability of Star Ratings Based on Program Volume

	No Minimum	**≥30 cases**	≥50 cases	≥100 cases	≥150 cases
Number of participants	231	**172**	136	80	40
Reliability (95% CrI)	42% [33%-50%]	**51% [41% 60%]**	55% [45%-64%]	60% [48%-70%]	65% [50%-78%]

Number of STS Participants Identified as Different form STS Average

Criterion for categorizing participants	Below STS-Average (1-star)	STS-Average (2-star)	Above STS-Average (3-star)
98% Credible interval	5 (2.9%)	160 (93.0%)	7 (4.1%)
95% Credible interval	**8 (4.7%)**	**152 (88.4%)**	**12 (7.0%)**
90% Credible interval	12 (7.0%)	142 (82.6%)	18 (10.5%)
80% Credible interval	16 (9.3%)	128 (74.4%)	28 (16.3%)

Distribution of Composite Scores



Construct Validity:

Complication Rates Vary Across Star Ratings

	1- Star	**2 - Star**	**3 - Star**
Operative Mortality (95% CrI)	3.2% [1.6% - 5.9%]	1.6% [0.6% - 3.5%]	0.9% [0.4% - 1.6%]
Major Complications (95% CrI)	17.1% [11.3% - 24.2%]	10.1% [5.1% - 16.9%]	6.5% [3.7% - 9.6%]

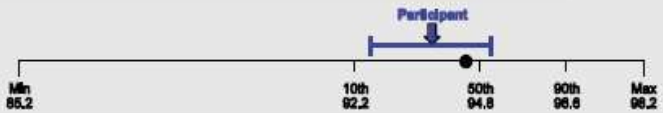
Risk Adjusted Outcomes: Star Ratings

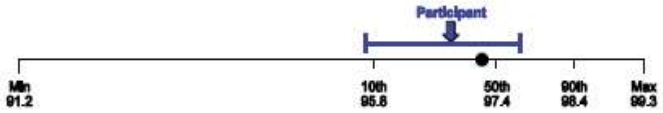
STS General Thoracic Surgery Database
Composite Quality Rating

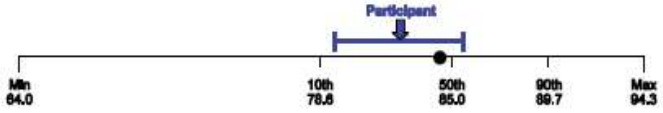
Participant 99999
STS Period Ending 12/31/2013

Duke Clinical Research Institute

Quality Domain	Participant Score (95% CI)	STS Mean Participant Score	Participant Rating[1]
Jan 2011 - Dec 2013 Overall	93.8% (92.5 , 95.0)	94.5%	★★
Jan 2011 - Dec 2013 Absence of Mortality	96.8% (95.7 , 97.7)	97.2%	★★
Jan 2011 - Dec 2013 Absence of Major Complication[2]	82.5% (79.3 , 85.5)	84.4%	★★







Quality Domain	Eligible Procedures	Detail	Count	Percent[1]
Jan 2011 - Dec 2013 Absence of Mortality	400	Mortality	13	
Jan 2011 - Dec 2013 Absence of Major Complication	400	Major complication	70	
		Pneumonia	18	25.7%
		ARDS	2	2.9%
		Bronchopleural fistula	1	1.4%
		Pulmonary Embolus	1	1.4%
		Initial Vent Support >48hrs	2	2.9%
		Reintubation	10	14.3%
		Tracheostomy	6	8.6%
		Myocardial infarction	1	1.4%
		Unexpected return to OR	16	22.9%
		Multiple complications (more than one of the above)	13	18.6%

[1]Percentages represent the proportion that the specific complication contributed to the total number of patients with major complication.

This information is intended to facilitate and focus process and quality improvement initiatives by providers.

Limitations

- Generalizability
  - STS outcomes are better than national benchmarks
  - Captures < 50% of lobectomies in the US
- Short term outcome measure
  - Vital status out to 5-years added to STS Database in 2015
  - Lack patient reported outcomes

STS Study Conclusions

- Developed a two-domain quality measure for lobectomy for lung cancer
- Identified 12% of participating programs as outliers
- Will present risk-adjusted outcomes and star ratings in the public reporting effort
- Encourage participation in voluntary public reporting

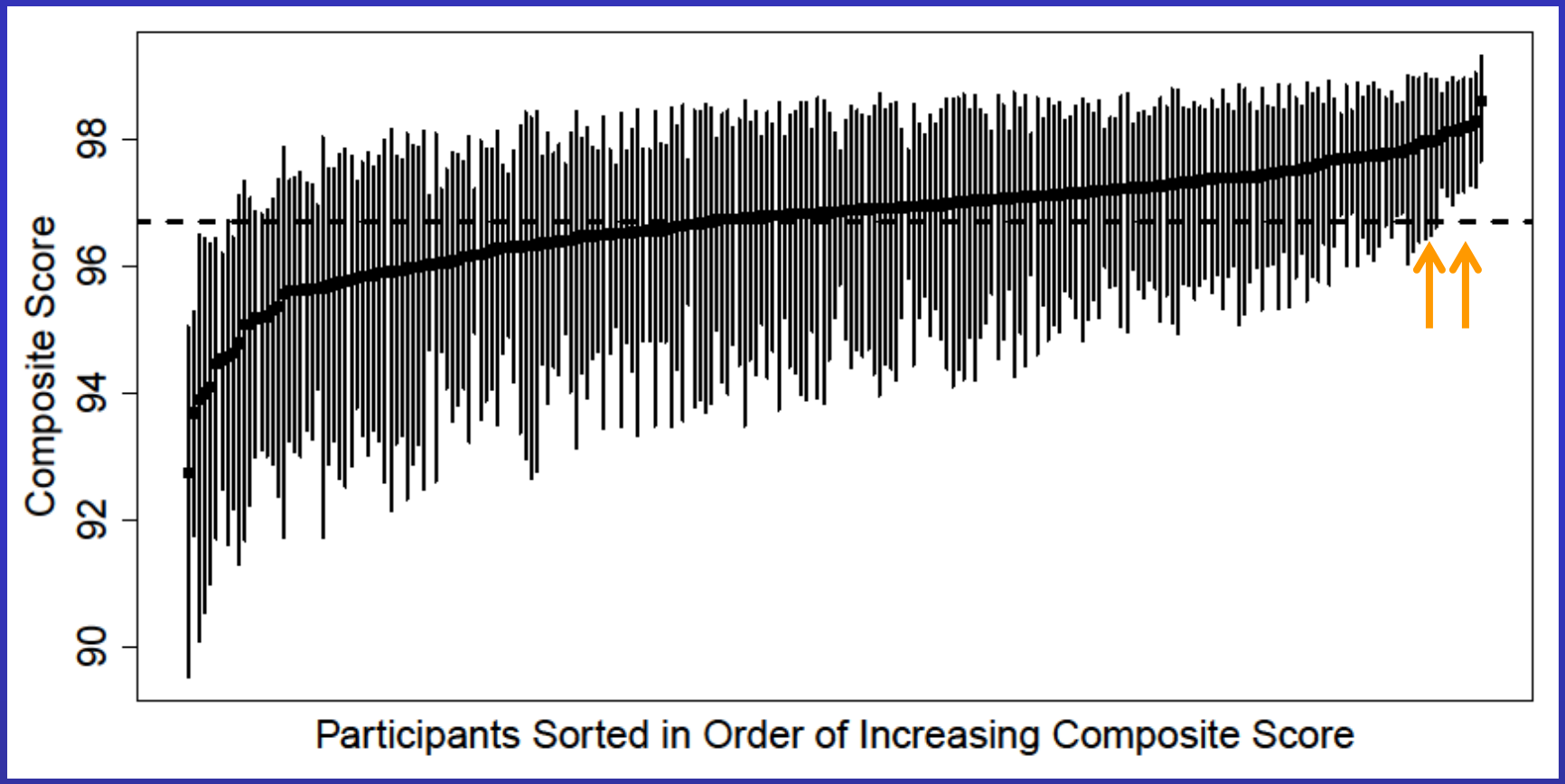
Unintended Consequences of Public Reporting

- Misclassification of surgeons or hospitals
- Risk aversion for high risk patients with reduced access to care
- Confusion for the public and payers
- Potential for serious financial consequences

Distribution of Composite Scores:

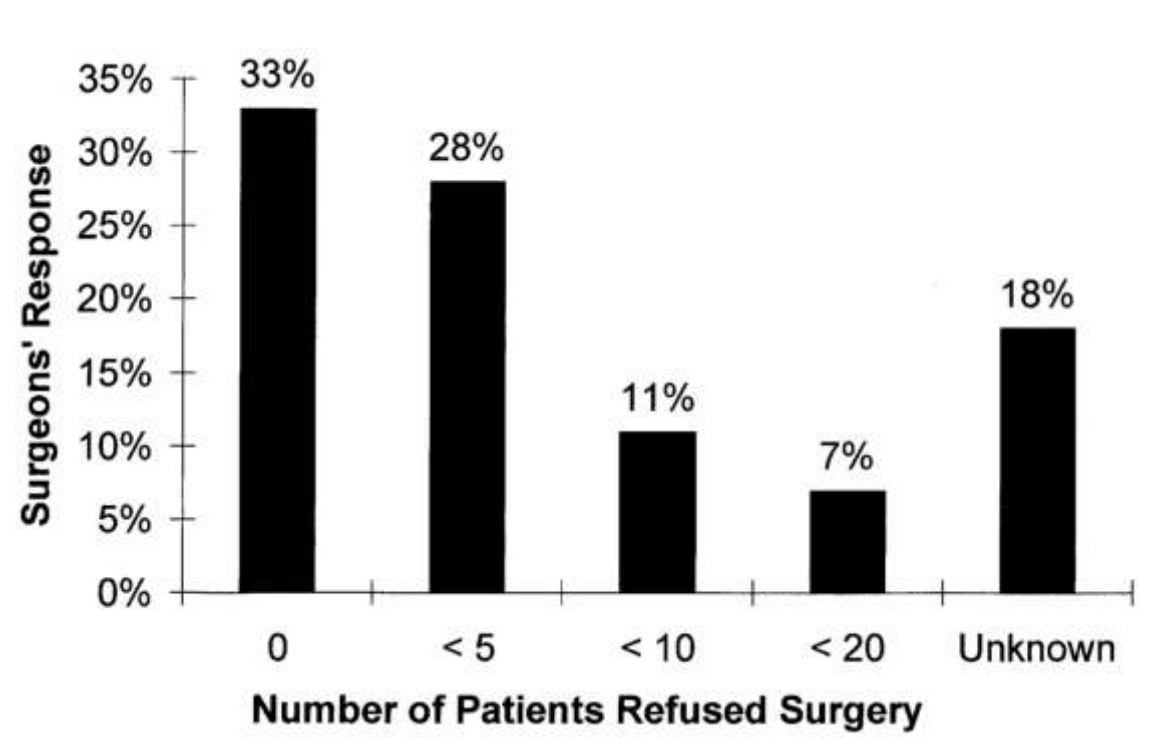
Potential for misclassification

Comparison to the mean, not to nearest neighbor



Public Reporting of Surgical Mortality: A Survey of New York State Cardiothoracic Surgeons

Joshua H. Burack, MD, Paul Impellizzeri, BA, Peter Homel, PhD, and Joseph N. Cunningham, Jr, MD



Most surgeons (62%) refused to operate on at least one high-risk CABG patient over the prior year, primarily because of public reporting.

Burack JH et al. Ann Thorac Surg. 1999;68(4)

REVIEW AND COMMENTARY

The Public Health Hazards of Risk Avoidance Associated With Public Reporting of Risk-Adjusted Outcomes in Coronary Intervention

Frederic S. Resnic, MD, MSc, Frederick G. P. Welt, MD, MSc
Boston, Massachusetts

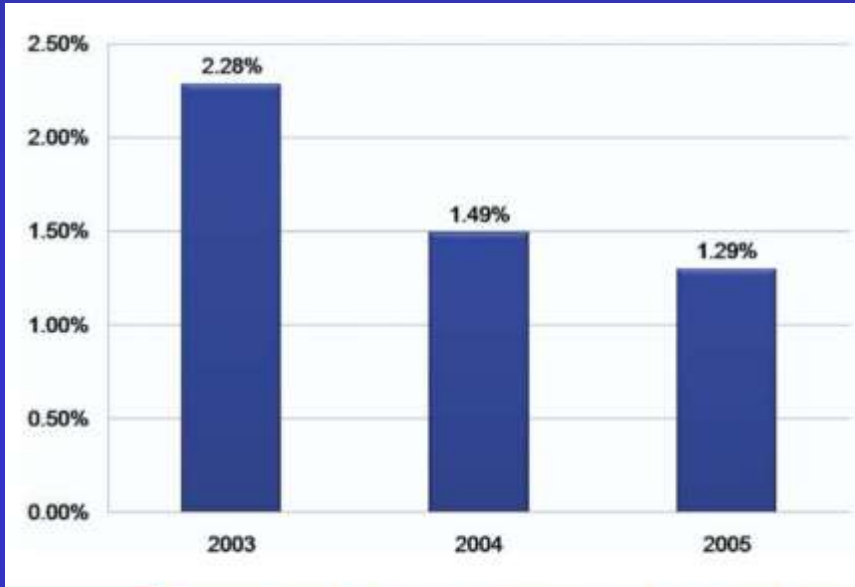


Figure 1 Proportion of Patients Undergoing PCI in Massachusetts for Cardiogenic Shock, 2003 Through 2005

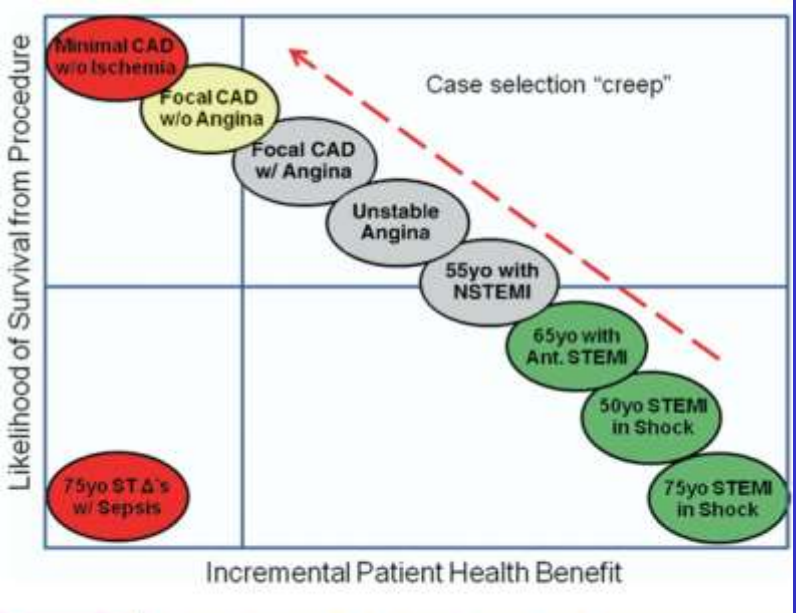


Figure 3 Potential for "Risk Avoidance Creep"

"We must address the underlying incentives for "case-selection creep" by improving risk adjustment methods for the highest-risk patients and by highlighting centers and physicians who undertake high-risk procedures in appropriate patients."

Resnic FS. JACC. 2009;53(10

Public Reporting of Surgical Outcomes: Conclusions

- The primary goals are:
  - Inform patients and stakeholders
  - Improve the quality of care
- Current public reporting initiatives have serious limitations
  - Opportunity for surgeons to lead
- Need accurate data
- Important and relevant outcomes

Public Reporting of Surgical Outcomes: Conclusions

- Need collaboration between key stakeholders
- Need to minimize potential consequences
- Reliable and transparent methodology
  - Iterative process

The Search for Quality

Acknowledgements

- University of Virginia Surgical Outcomes Research Center
  - George Stukenborg, PhD
  - Tim McMurry, PhD
- Society of Thoracic Surgeons
  - David Shahian, MD
  - J William Gaynor, MD