



Continuing Conflict Over Use of Imperfect Healthcare Performance Measures

December 6, 2017
National Healthcare Ratings Summit



Arnold Milstein, MD
Professor of Medicine
Stanford University
Clinical Excellence Research Center

© 2017 A. Milstein/Stanford





A Parable for Our Time

**The first emperor of China
fatally poisoned himself by
eating mortality-preventing pills**



Qin Shi Huang
Died September 10, 210 BC

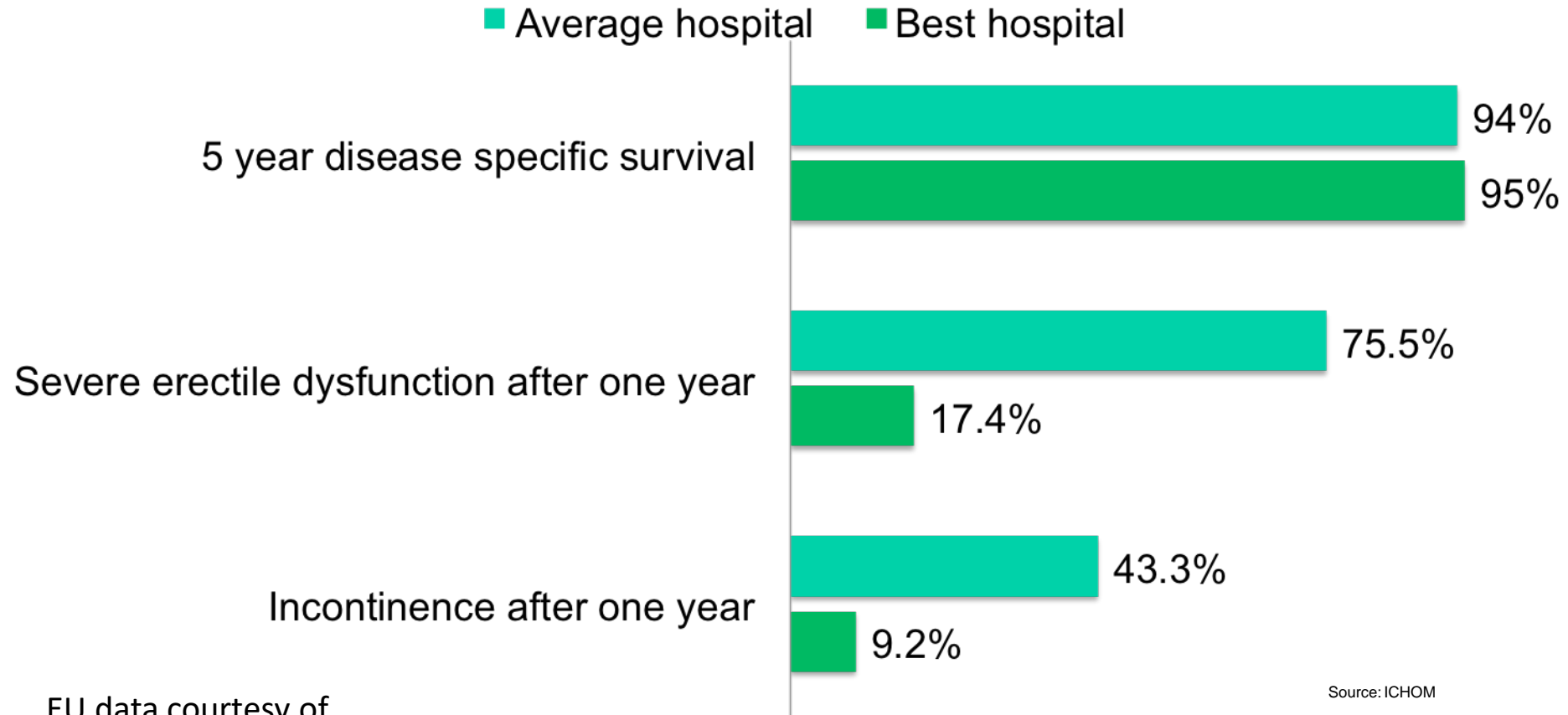


Thoughtful People Disagree About Minimum Validity Cutoffs





More Granular Outcomes Measures Best



EU data courtesy of Elizabeth Teisberg

Source: ICHOM



Research on “Social Loafing” Also Favors Granular Performance Reporting

Net Number of Pro-Social Suggestions

<u>Individual identifiability</u>	<u>Group</u>	<u>No Group</u>
Identifiable	21	15
Anonymous	15	15

Condensed from: Guerin, B. (1999). Social behaviors as determined by different social consequences: social loafing, social facilitation, deindividuation, and a modified social loafing. *The Psychological Record*, 49, 565-578.



If Granular Outcomes Measures → Smaller Sample Sizes → Less Validity, Why Not Limit Reports to Large Orgs?

Variability in the effectiveness of care is driven by medical group **and** individual physician performance.

–Jeffery Levin-Scherz MD and Thomas H. Lee MD
(emphasis added)

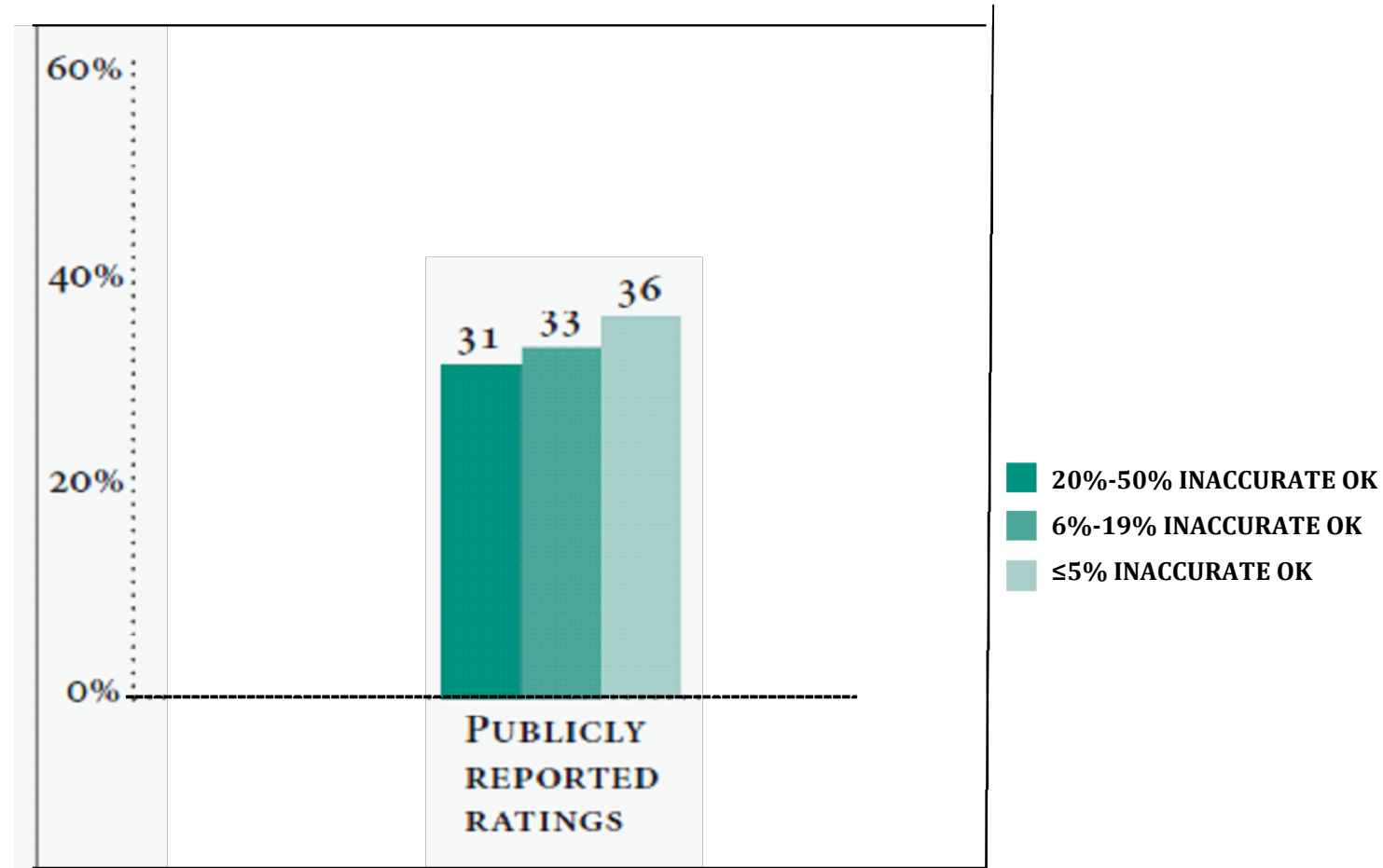
Improving the quality of care ultimately requires changes in the behavior of individual physicians, even if systems to improve the quality of care play an important role.

–Arnold Epstein MD, Thomas H. Lee MD & Mary Beth Hamel MD



Relax Minimum Validity Cutoffs ? American Public's Perspective

Percent of
American
Public



* Difference from public ratings statistically significant at $p < .05$.

Difference from public ratings and for ratings to encourage use of highly rated physicians statistically significant at $p < .05$.

Source: Center for Studying Health System Change, national survey conducted by Knowledge Networks, Inc., December 2006

Davis, Hibbard, Milstein; Consumer Tolerance for Inaccuracy in Physician Performance Ratings; HSC; 2007



AHRQ 2012 Evidence Report on Current Imperfect Public Reporting

- **In general, public reporting has a positive impact on quality measures**
- **More evidence of no harm than evidence of harm**
- **Both individual clinicians and provider organizations respond to public reports with positive behavior changes**
- **Little to no impact on selection of providers by patients and families because (1) information not available when needed or (2) not presented comprehensibly**

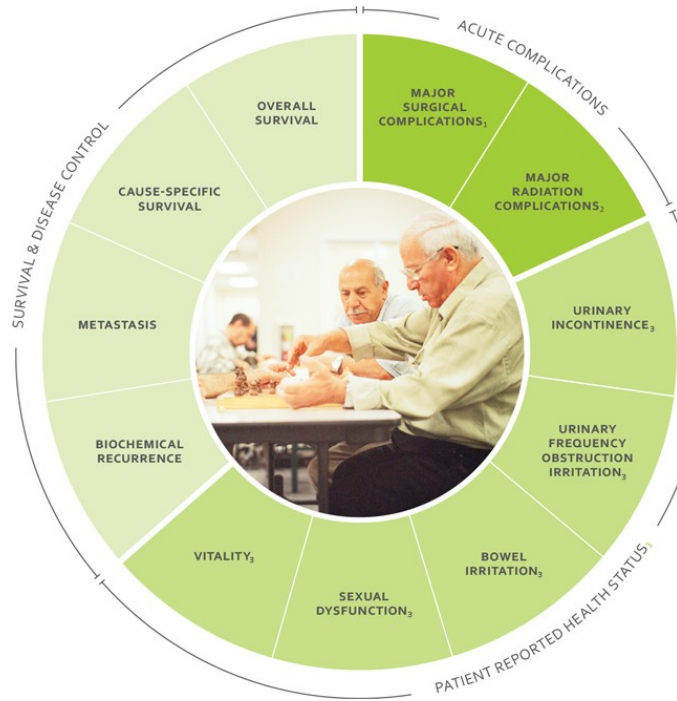


An Ethical Perspective

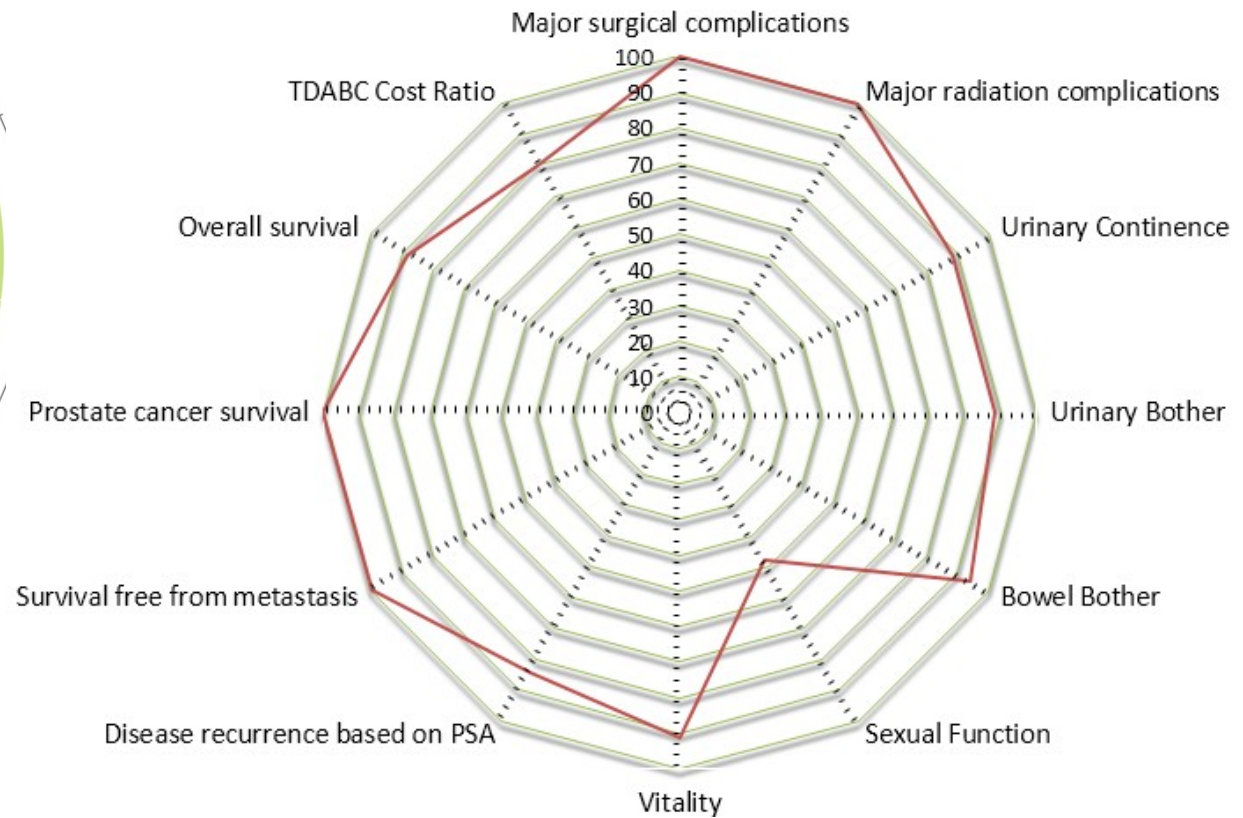




A Less Conflictful Waystation?



Example: prostate cancer.
Data are illustrative only.
Courtesy of Elizabeth Teisberg





Our 21st Century Destination?

- **Using conflict constructively**
- **Linking short and long-term PROMs with data from connected devices, EHRs, & insurance claims**
- **Applying AI-based similarity analytics for patient-tailored guidance in real time**

