

MEASURING PROVIDER PERFORMANCE: THE FOUNDATION OF PURCHASER STRATEGIES TO REFORM AMERICA'S HEALTH CARE SYSTEM



PubM 7564 - Online 2013

1

Measuring Provider Performance

- ▶ Who constructs measures?
- ▶ Common measurement issues
- ▶ Measuring quality
- ▶ Measuring cost/efficiency
- ▶ Measuring overuse



PubM 7564 - Online 2013

2

Who Constructs Measures?

- ▶ Health plans
 - Responding to employer demands for quality and cost information to share with their enrollees to inform choice of providers.
 - Used internally in reimbursement negotiations, setting of reimbursement rates, and constructing tiered networks.
- ▶ Community/purchaser coalitions
 - Used in public reports of provider performance.
- ▶ Independent, non-profit organizations
 - Generate national consensus measures that can be used by health plans and community coalitions.
- ▶ Physician specialty and hospital organizations
 - Generate measures acceptable to providers with intention that the measures be used by health plans and community coalitions.



PubM 7564 - Online 2013

3

Considerations in Selecting Measures*

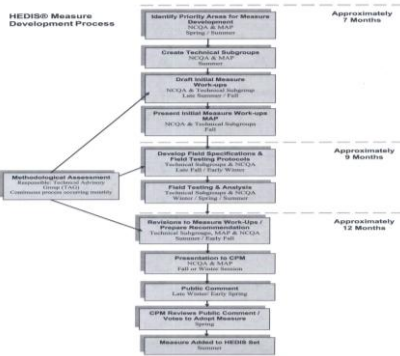
► Feasibility

- Data are available for constructing measure
- Clear specifications for data sources and for methods of data collection and reporting
- Data collection/provision does not impose unreasonable burden on providers
- Data collection does not violate confidentiality standards
- Measure can be audited to prevent manipulation by providers

*<http://www.ncqa.org/tabid/425/Default.aspx>

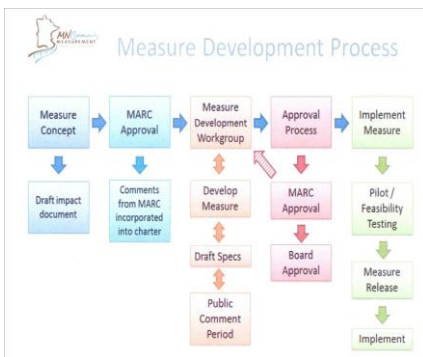
PubH 7564 - Online 2013

7



PubH 7564 - Online 2013

8



Minnesota Community Measurement. *The Measurement Minute*, August 2012

PubH 7564 - Online 2013

9

Measure Concepts In Development Pediatric Preventive Care Total Cost of Care Care Coordination (Health Care Homes Only)	Measure Concepts on Hold (data source) High Tech Diagnostic Imaging 30 Day All-Cause Readmission
Measures in Pilot/ Implementation Total Knee Replacement Lumbar Spine Surgery Colonoscopy Quality	New Measures for Submission Maternity Care- C-Section

Minnesota Community Measurement. *The Measurement Minute*, August 2012

Common Measurement Issues

- ▶ Attribution
- ▶ Composite vs. specific
- ▶ Risk Adjustment
- ▶ Validity/Reliability (small numbers problem)

Attribution

- ▶ Problem is more important for measurement of physician performance, as opposed to hospital performance.
- ▶ Attribution becomes an issue when patient sees multiple providers during measurement time period, as is often the case for patients with chronic conditions.
 - Whose quality or cost performance is being measured?

Attribution

- ▶ Options for attribution:
 - Assign responsibility for quality or cost outcomes to physician with most visits from patient during measurement period.
 - Assign responsibility to physician who accounted for largest share of costs (for cost measures)
 - Assign responsibility to primary care physician
 - Assign measure to all providers seen by patient during measurement period.



13

Attribution

- ▶ How to attribute people who don't use care in measurement period.
- ▶ In analysis of physician cost profiling, RAND concluded that attribution rule (tested 12 variations) strongly affected cost category physician was placed in — using different rules resulted in reassignment of up to 61 percent of physicians (RAND Research Highlights: Is Physician Cost Profiling Ready for Prime Time?)



14

Composite vs. Specific Measures

- ▶ As number of performance measures grow, it becomes more confusing for purchasers and payers to make judgments about high versus low performing providers.
 - Providers may be high performers on some measures and low on others.
 - Relative importance of different measures can be hard for consumers to assess.
 - Consumers may be confused, not use information in making decisions.



15

Composite vs. Specific Measures

- ▶ Composite, summary, aggregate, or “all or nothing” measures collapse a large number of measures into a single, or smaller number, of measures.
 - Easier for consumers to reach judgments about providers
 - but involve value judgments (that may not be clear) in collapsing measures.



16



Minnesota Community Measurement, 2009. http://www.thed5.org/index.php?r=about_the_d5



17

Creating Composite Measures

- ▶ Denominator-based weighted measures
 - Simple average of all indicators
 - Weighted average, based on number of eligible patients
 - “All or nothing”
- ▶ Presentation
 - Composite measure only
 - Composite measures with measures on specific indicators
 - Ranking of providers
 - Star rankings (e.g. quartile in which provider’s score falls)



18

Risk Adjustment

- ▶ **Problem:** Patients are not randomly distributed across providers; some providers may attract more severely ill patients than others, even within a specific condition group (e.g. diabetes).
 - Performance measures may indicate that a provider is more costly, or of lower quality, if the provider attracts sicker patients and/or patients with other characteristics that make them more difficult or costly to treat.
 - Could discourage providers from accepting these patients.
 - Could unfairly rank some providers higher than others, or pay them more, simply because they treated less complex, or less "compliant," patients.
- ▶ **Goal of "risk-adjustment":** Take into account inherent differences in patients when measuring performance.
 - In quality measurement, more important for patient outcome than treatment process measures.



Risk Adjustment

- ▶ **Approaches to problem:**
 - Group patients into "cells" with separate scores computed for each cell (e.g. gender); sometimes called "segmentation".
 - Use statistical methods to "control for" differences in patient characteristics; large number of statistical approaches are available (some are proprietary).



Risk Adjustment

- ▶ **Factors included in risk-adjustment models:**
 - Demographic information
 - Self-reported health status
 - Administrative data (medical and pharmacy claims) indicating patient condition, diagnosis, past expenditures
- ▶ **Demographic information:** easy to collect, accurate, but explains relatively little variation.
- ▶ **Self-reported health status:** expensive to collect as it requires surveys, adds slightly to explanatory power of demographic information.
- ▶ **Administrative data:** adds considerably to predictive power, especially for expenditures; may create undesirable incentives for providers; diagnosis information not always complete



Validity/Reliability

- ▶ Question: How accurately does measure reflect actual performance?
 - In any measure, there is likely to be some error.
 - One criterion for measure selection is the reliability of the measure.
- ▶ Problem of small numbers
 - Generally accepted that a minimum of 30 patients is needed per physician or physician practice to construct reliable measures of performance
 - The number may need to be much higher when there is large variation in performance measure
 - It is unusual for physicians to have enough patients with a specific condition enrolled in any single plan to construct reliable measures for plan enrollees.
 - True even for relatively common conditions like diabetes
 - When the number of patients for whom performance is measured (assigned to a single physician) is low, the risk of misclassifying physicians as "high" or "low" performers is great.

PubM 7564 - Online 2013

22

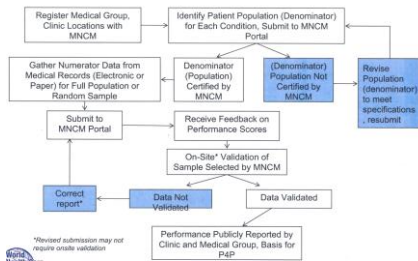
Measuring Quality

- ▶ Process versus outcome measures
- ▶ Process measures
 - Often can be calculated with claims data, so are relatively cheap to construct.
 - Can address preventive or acute care
 - Actual measures are typically based on HEDIS measures, so are familiar to health plans and providers.
 - Physicians dispute their accuracy (don't trust accuracy of claims)
- ▶ Outcomes measures
 - Typically cannot be constructed using claims data, so require chart review or abstraction of data from electronic medical records.
 - Closer to ultimate goal of measuring health of patients of different providers.
 - Providers worry that some outcomes are influenced by patient behaviors as much as the medical treatment they provide; don't want the "blame" for patient's bad lifestyle choices.

PubM 7564 - Online 2013

23

Process Flow



PubM 7564 - Online 2013

24

Measuring Overuse

- ▶ Traditional quality measures have focused on documenting underuse of recommended services; for instance, percent of physician patients receiving recommended service, with smaller percentage indicating worse quality.
- ▶ More recently, "overuse" of services has become a focus for measurement
 - Overuse: "The use of a service that is unlikely to improve patient outcomes or for which potential harms exceed likely benefits."⁴
 - Better quality requires that physician reduce use of particular procedure or service.
 - Reaching professional consensus on measures of overuse is difficult
 - Examples include some types of CT scans, chemotherapy for some types of tumors.
- ▶ Reducing overuse has the potential to be "win-win" – improving quality while lowering costs – for the health care system but a revenue loss for providers under fee-for-service payment.

⁴Mathias and Baker, "Developing Quality Measures to Address Overuse," *JAMA*, May 8, 2013, p. 1897.

PubH 7564 – Online 2013

31

Table. Potential Unintended Consequences of Overuse Measurement

Potential Unintended Consequences of Overuse Measurement	Example
Underuse of the service when it is actually indicated	Measuring overuse of bone scanning for staging low-risk patients may unintentionally lead to underuse in higher-risk patients for whom bone scan is indicated.
Underuse of other related services	Measuring overuse of cervical cancer screening in low-risk women may unintentionally lead to underuse of other preventive services, such as colonizing, mammography.
Patient selection	Measuring overuse of spine imaging may unintentionally lead primary care physicians to avoid caring for patients with low back pain by mislabeling them all as urgent.
Conservation shift	Measuring overuse of imaging in physicians' offices may unintentionally lead to physicians sending patients to the emergency room for imaging.
Increasing use of alternative tests or treatments	Measuring overuse of lumbar spine radiographs may unintentionally lead to increased use of lumbar spine magnetic resonance imaging.
Change to the patient-physician relationship	Measuring overuse of antibiotics for bronchitis may unintentionally change a physician's relationship with a patient because the physician did not order the treatment the patient desired.
Caricature dissatisfaction with quality measurement	Excessive measurement burden (eg, additional documentation to confirm that service use is not excessive) may lead to clinician dissatisfaction with quality measurement on the front lines, gaining the system, or both, to improve performance on the measure without improving patient care.
Adverse public health effects	Measuring overuse of blood cultures may unintentionally lead to decreased availability of data necessary to track antibiotic resistance over time; such difficult situations must be anticipated and addressed prior to implementation of outcome measure.

Mathias and Baker, "Developing Quality Measures to Address Overuse," *JAMA*, May 8, 2013, p. 1897.

PubH 7564 – Online 2013

32