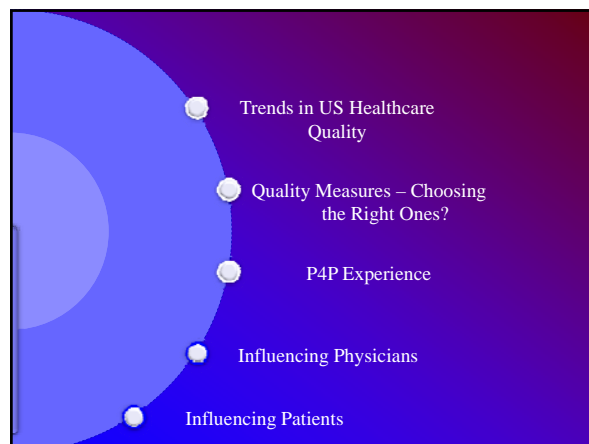


Measuring the Right Thing – Key to Successful Engagement of Physicians and Patients

Paul C. Tang, MD

Vice President, Chief Medical Information Officer
Palo Alto Medical Foundation
and
Consulting Associate Professor
Stanford University



Palo Alto Medical Foundation Sutter Health Affiliate



San Francisco Bay Area

- ~1000-physician multi-specialty group practice
- ~2 M outpatient visits/yr
- IHA P4P: 90th percentile for 9 of 11 clinical metrics

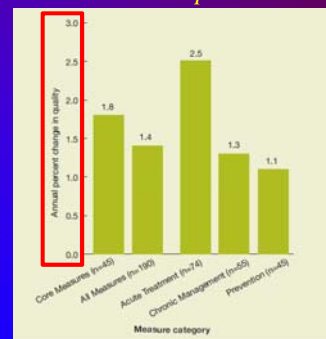
Quality: an Elusive Goal?

Role of P4P

US National Healthcare Quality Report

- Agency for Healthcare Research and Quality (AHRQ) generates annual report since 2003, mandated by Congress
- Originally built on 220 measure, now focused on 45 core measures reflecting
 - Effectiveness
 - Patient safety
 - Timeliness
 - Patient-centeredness

US National Healthcare Quality Report Median Annual Improvement Rate



Why hasn't performance reporting improved quality in the US?

- a) Are we measuring the right things?
- b) Are the incentives adequate?
- c) Do physicians care?
- d) Do we provide system support for physicians to do the right thing?
- e) Do we incent the wrong things?

Are we measuring the right things?

Deriving Quality Measures *US Historical Perspective*

1. What data are available?
2. Which are standardized and combinable?
3. What important quality questions can you answer with the data you have?
4. What quality goals can you set based on the data?



Impact of Using Administrative Data for Clinical Quality Reporting

Comparing Claims-Based Methods with EHR-Based Methods

Funded by US Centers for Medicare and Medicaid Services Tang PC, et al. J Am Med Inform Assoc. 2007;14:10-15.
<http://www.jamia.org/cgi/reprint/14/1/10>

Methods

- Randomly selected charts of Medicare patients reviewed for presence of diabetes by 3 methods
 - Gold standard chart review (to identify 125 diabetics)
 - Claims-based definitions used in CMS DOQ project (2 visits with encounter diagnosis of diabetes)
 - Query of coded information in EHR
 - Problem list, medication list, lab results (and not progress notes)
- Apply DOQ quality measures using standard definition vs. clinical definition

Results

- 98% of gold-standard diabetics identified using EHR coded data (sens=97.6%, spec=99.6%)
 - 94% identified using problem list alone
- 25% of gold-standard confirmed diabetics “missed” by *administrative claims-based* definition
- Statistically significant difference for 50% of diabetic performance measures when comparing those identified using administrative definition vs. those missed by administrative definition

Tang PC, et al. J Am Med Inform Assoc. 2007;14:10-15.
<http://www.jamia.org/cgi/reprint/14/1/10>

Results

Performance Measure Differences in Subgroups

Table 5 ■ DOQ Diabetes Measures Calculated From Expert Review Data for all Patients Identified as Having Diabetes: Comparison of Patients With Two Visits for Diabetes Vs. Patients With Zero or One Visit

Measure (Probability, Fisher's Exact Test)	Zero or One Visit For Diabetes		Two visits For Diabetes		P
	N (%)	D	N (%)	D	
DM1: HbA1c Management (p<.001)	21 (67.2%)	31	91 (86.5%)	94	
DM2: HbA1c Management Control (measure of poor control) (p=0.27)	0 (0.0%)	21	6 (6.6%)	91	
DM3: Blood Pressure Management (p=0.05)	14 (43.2%)	31	57 (66.6%)	84	
DM4: Lipid Measurement (p=0.06)	22 (71.0%)	31	78 (83.9%)	94	
DM5: LDL Cholesterol Level (p=0.23)	21 (95.5%)	22	69 (88.5%)	78	
DM6: Urine Protein Testing (p<.001)	17 (54.8%)	31	80 (85.1%)	94	
DM7: Eye Exam (p=0.03)	12 (41.4%)	29	55 (61.8%)	89	
DM8: Foot Exam (p=0.13)	2 (7.1%)	28	15 (16.5%)	91	

Implications

Claims-Based Measures

- Underestimates target population (denominator)
- Biased toward spuriously higher scores (self-fulfilling prophecy)
- Potential to misdirect quality-improvement efforts
- Subject to "gaming" (no clinical downside)

EHR-Based Measures

- Accurately identifies target population (subject to policies)
- More accurate, though lower scores may disincent EHR adoption
- More accurate tool to manage clinical QI initiatives
- Clinical record less subject to "gaming" due to clinical reuse

Physician Engagement in Reports Alignment with Clinical Guidelines Key

- Diabetes Guideline: Check HbA1c every 3-6 mo
 - Reporting: 1 yr (??) – 92%
- Guideline: Lower HbA1c < 7.0%
 - Performance (2004 NCHS, survey): 49%
 - Reporting: HbA1c > 9% (??) – 12%
- Guideline: Maintain BP < 130/80
 - Performance (2000 NHANES): 36%
 - Reporting (2004 NCHS, 140/80 (??)): 57%

JAMA 2004; 291:335, AHRQ MEPS, 2004, NCHS, 2004

Deriving Quality Measures that Drive Care

1. Can we make them exist in EHRs?
2. Are they standardized and combinable?
3. What data do you need to measure important quality outcomes?
4. What data are needed to influence care?
5. What are the important quality goals?

Retooling Quality Measures for an EHR-Enabled World

National Quality Forum's HIT Expert Panel (HITEP)

HITEP Charge

- Accelerate current efforts to **identify** a set of **common data elements** to be **standardized** in order to **enable automation** of a prioritized set of AQA and HQA measures **through EHRs**

Measure Development Framework Data Quality Criteria

- 1. Authoritative source/accuracy:** Is the entry in the EHR from an authoritative data source? What is the accuracy of the data element in EHRs? [Weight 5]
- 2. Use of data standards:** Does the data element use standardized data elements for coding? [Weight 5]
- 3. Fit workflow:** Does capture of the data element by the most appropriate healthcare professional fit the typical EHR workflow for that user? [Weight 4]
- 4. Availability in EHRs:** Is the data element currently available within EHRs? [Weight 4]
- 5. Auditable:** Can the data be tracked over time to assess accuracy? [Weight 2]

Scale: 1-5; Weight (out of 5)

Measure Development Framework Analysis of Quality Measure Clusters (DM)

example: Diabetes

Data class-type name	Quality Score	Freq %	measures							
			HbA1c, checked	HbA1c, at goal	BP, checked	Lipids, checked	Lipids, at goal	Eye exam, screen	Eye exam, severity	
history-birth_date	100	53	●	●	●	●	●	●	●	
physical_exam-vitals	95	1			▲					
laboratory-result	91	13	▲	▲		▲	▲	▲	▲	
diagnosis-inpatient	86	71	● x	● x	● x	● x	● x	● x	● x	
medication-outpatient_order_filled	76	10							▲	
medication-outpatient_order	72	19	●	●	●	●	●	●	▲	
diagnostic_study-result	59	33							▲	
procedure-consult_result	55	6							▲	
diagnosis-outpatient (billing)	47	26	● x	● x	● x	● x	● x	● x	●	
optout-other_reason	20	38				x			x	

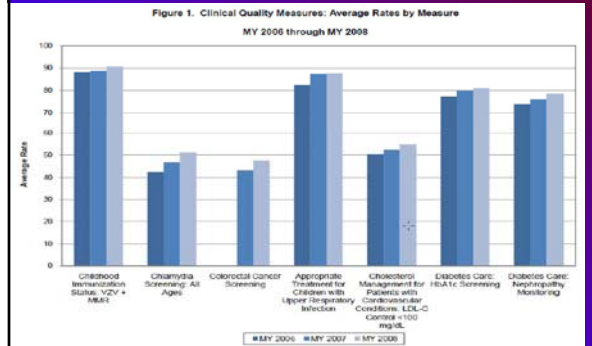
● - denominator
x - exclusion
▲ - numerator

Are we providing adequate incentives?

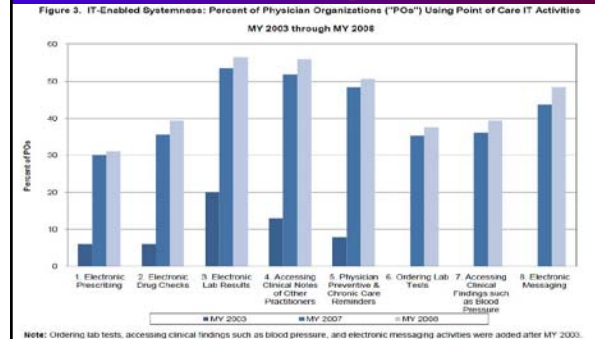
California IHA P4P HMO P4P

- Began in 2003: 6 health plans, 200 medical groups (35K MDs), 7M HMO lives
 - Single report card, 25 measures (clinical, patient experience, HIT)
- 2008: 8 health plans, 11.5M HMO lives
 - 68 measures (clinical, diabetes, patient experience, HIT-enabled systemness, efficiency)

Impact of IHA P4P Clinical

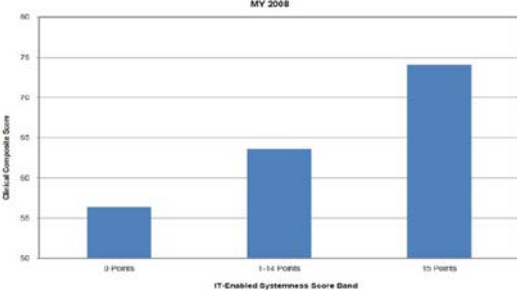


Impact of IHA P4P HIT Adoption



Impact of IHA P4P Impact of HIT

Figure 4. IT-Enabled Systemness: Clinical Composite Scores Distribution by IT-Enabled Systemness Score Band



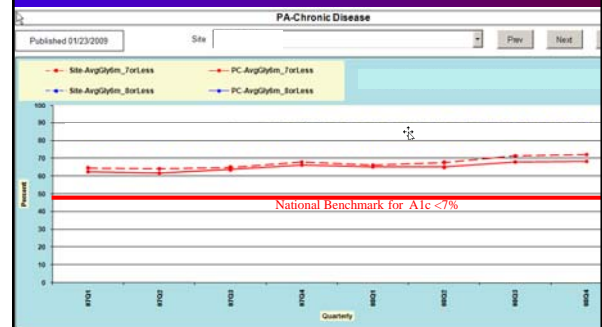
Note: Clinical composite score averages are calculated for PCIs with reportable rates for at least half of the clinical quality measures.

Do physicians care?
Do they have the right tools to
change their behavior?

*Engaging Physicians: Aligning
Performance Measures with Clinical
Guidelines*

Clinical Decision Support Diabetes Example

Improvement in Diabetes Control HbA1c Control



Barriers to Quality Improvement? "The Answers"

Questions

- Do physicians care?
- Measuring right things?
- Incentives adequate?
- Do we provide physicians system support to do the right thing?
- Incent the right things?

"Answers"

- Yes, physicians drive towards clinical goals they believe in
- No, need clinical measures aligned with clinical guidelines
- No, need >10% at risk
- No, need EHRs with clinical decision support – make the right thing trivial to do
- No, in US perverse volume rewards

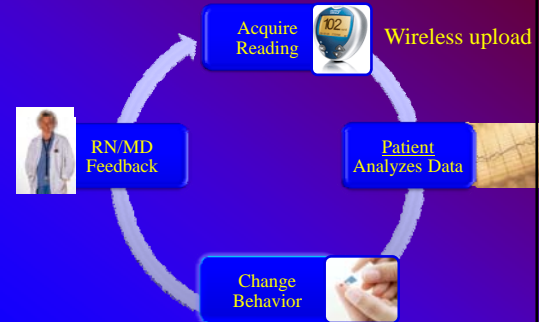
What about patients?

Enabling Reports for Patients

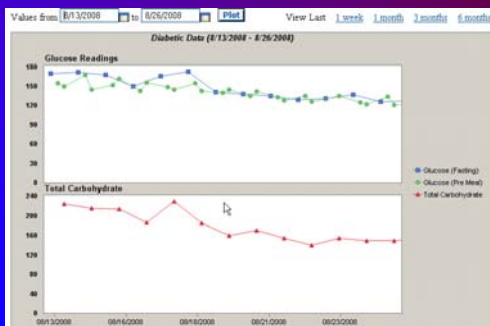
Providing Tools for Realtime Feedback

Helping to 'Connect the Dots'

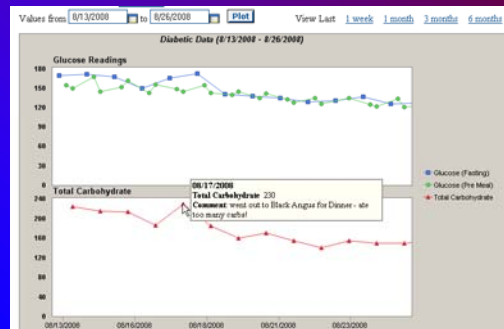
Disease Management Self Management



Providing Feedback



"Connecting the Dots" Oops



"Connecting the Dots" Changing Behavior



Feedback from Beta Group Mar 20, 2008

- **Doing it for us:**
 - "Being in the [online disease management] program means people are interested in you."
 - "Kelly was watching" "Knowing information will get to Kelly"
 - "Fill the gaps"
- **Learning from data:**
 - "Eating made a big difference in readings..."
 - "...also found out that what I eat affects the readings."
 - "My father lost his legs, so I have already been concerned."
- **Doing it for themselves:**
 - "If I'm going to eat something, I think about what my reading will be, so I don't eat it."
 - "It makes denial more difficult."
 - "I've incorporated the tools into my daily life."

Summary

- Using P4P to improve care *requires measuring* the “*right thing*” for the *right people* (clinicians and patients) at the *right time* (teachable moment)
- To engage physicians, measures must be:
 - Clinically valid (consistent with guidelines)
 - Based on clinical data (not administrative data)
 - Reported efficiently (as a byproduct of care)
- To effect change, motivated physicians must have tools to change behavior (EHR + decision support)
- Patients need the same support (tools and feedback)