RAND HEALTH

University of California Berkeley Haas School of Business

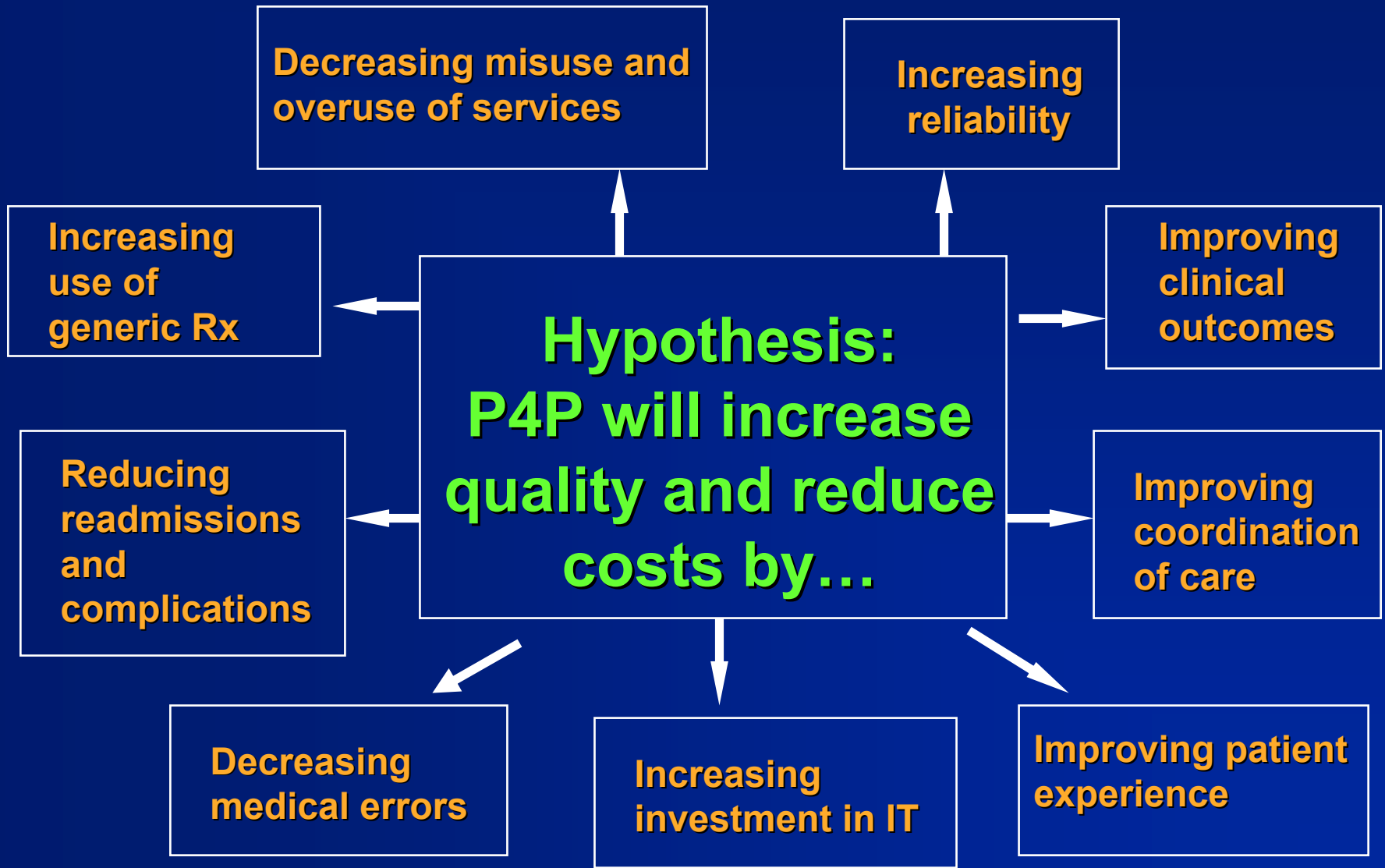
The Pay for Performance Experiment: Have we Reached the Promised Land?

Cheryl L. Damberg, PhD (RAND)
Kristiana Raube, PhD (UC Berkeley)

May 28, 2008
15th Annual Council on Health Economics and Policy Princeton Conference

Financial support provided by the California Healthcare Foundation

What is the Promise of P4P?

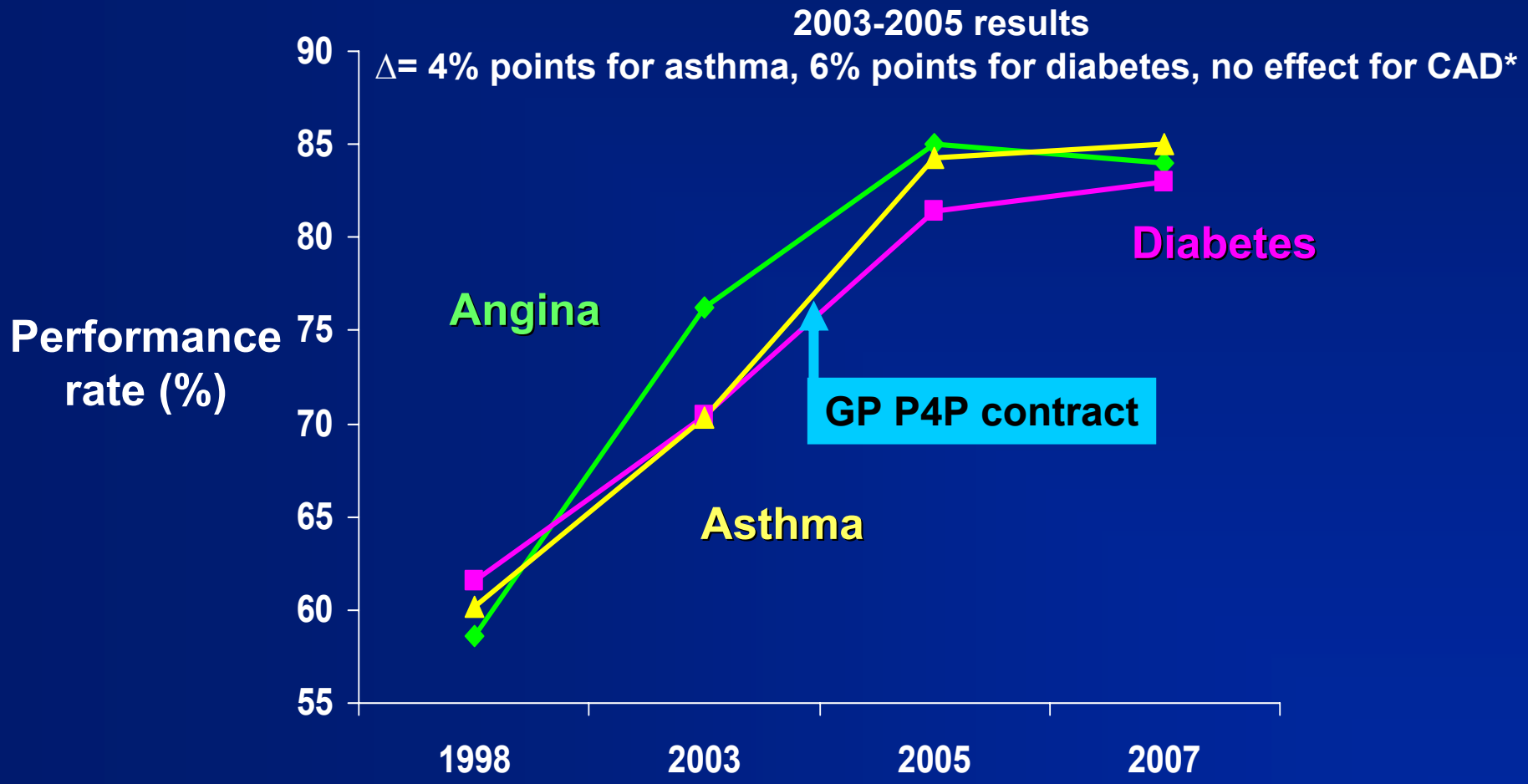


What is the Evidence to Date?

- Little evaluation of pay for performance (P4P) has occurred to assess impact
- The handful of published studies show mixed results or modest positive results
- P4P program context and design factors matter in terms of program impact
- P4P alone is unlikely to solve quality and cost problems, but may be useful when combined with other policy leavers

The UK Experience

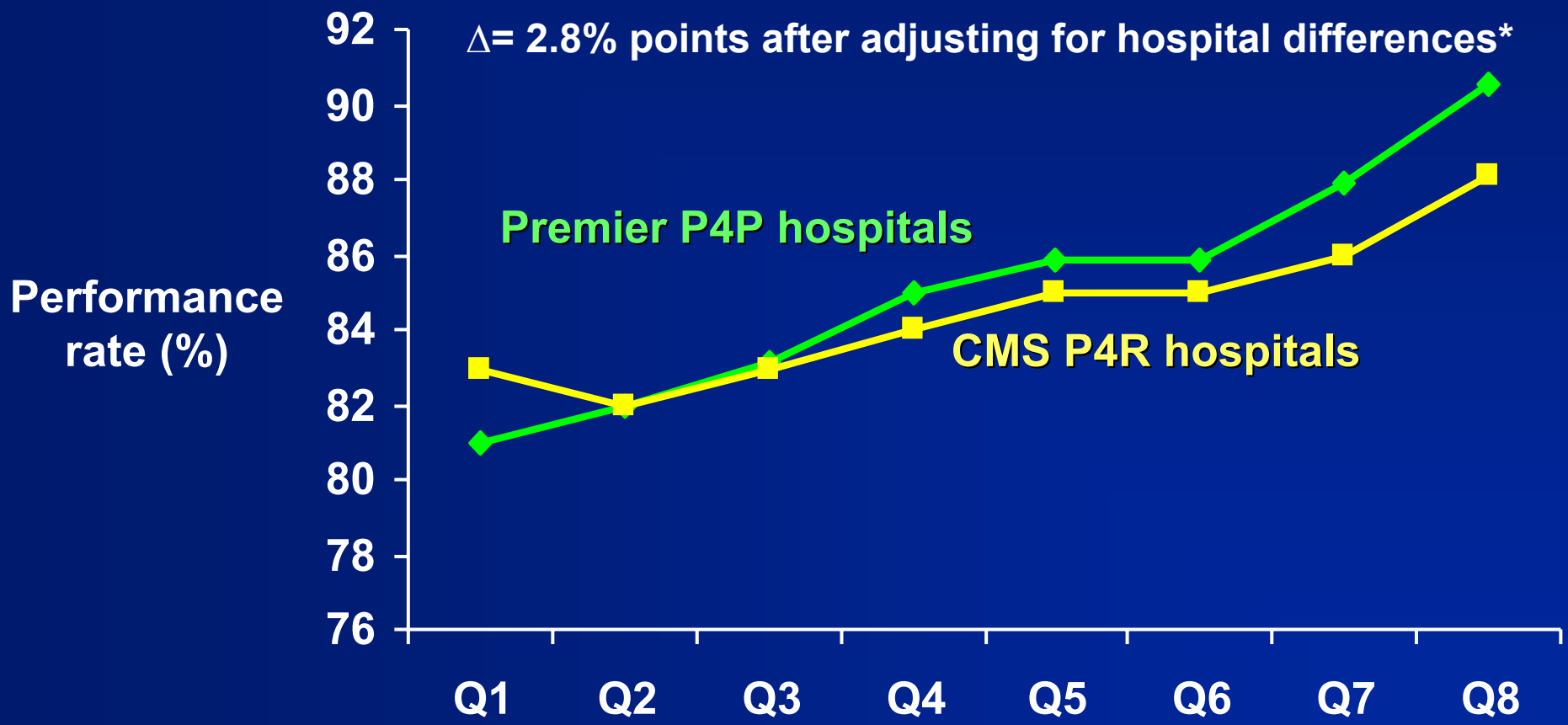
Nationally representative sample of 42 GP practices
Composite quality scores reflecting 48 indicators
Max score for each condition = 100



*Study by Campbell et al., July 12, 2007 (*New England Journal of Medicine*), and personal communication with Martin Roland, 2008

Premier CMS P4P Demonstration

Comparison of Performance on Composite of 10 Measures: Q4 2003-Q3 2005



*Study by Lindenaur et al., 2007 (*New England Journal of Medicine*)

Integrated Healthcare Association P4P Experiment

- Largest P4P program in the U.S.
  - Started in 2003, with first payout in 2004
- 225 capitated physician organizations (POs) located in California
  - >40,000 physicians
- 7 major health plans
- Addresses 3 domains
  - Clinical (13 measures)
  - Patient experience (5 measures)
  - IT capability (2 domains)
- More than $203 Million in payouts (2004-2007)
  - 2007 incentive payments represented a small fraction of total payments (2-3% of total capitation for average PO)

Small Improvements in Patient Experience, Similar to pre-P4P Public Reporting Trends

Measure	MY 2005	MY 2006	Mean Difference
Rating of Health Care	83.2	83.2	0.06
Rating of Doctor	86.2	86.6	0.39
Rating of Specialist	84.2	84.7	0.56
Doctor Communication	87.2	87.8	0.59
Timely Care and Access	73.8	73.9	0.07
No Problem Seeing Specialist	71.7	72.3	0.56

Note: Mean scoring, all items converted to a 100 point scale

Absolute Change in Performance 2005 vs. 2006*

Composite Measure	+/- 1 pt	+/- 2 pts	+/- 3 pts	+/- 4 pts	+/- 5 pts	+/- 10 pts
Access	42.9%	25.5%	18.0%	6.8%	5.6%	1.2%
Coordination	29.8%	25.5%	17.4%	9.3%	8.1%	9.9%
MD Interaction	50.3%	29.2%	15.5%	3.1%	0.6%	1.2%
All Care	44.9%	38.9%	12.0%	4.2%	0.0%	0.0%

*Each row sums to 100%

Upward Improvement, but not Breakthrough

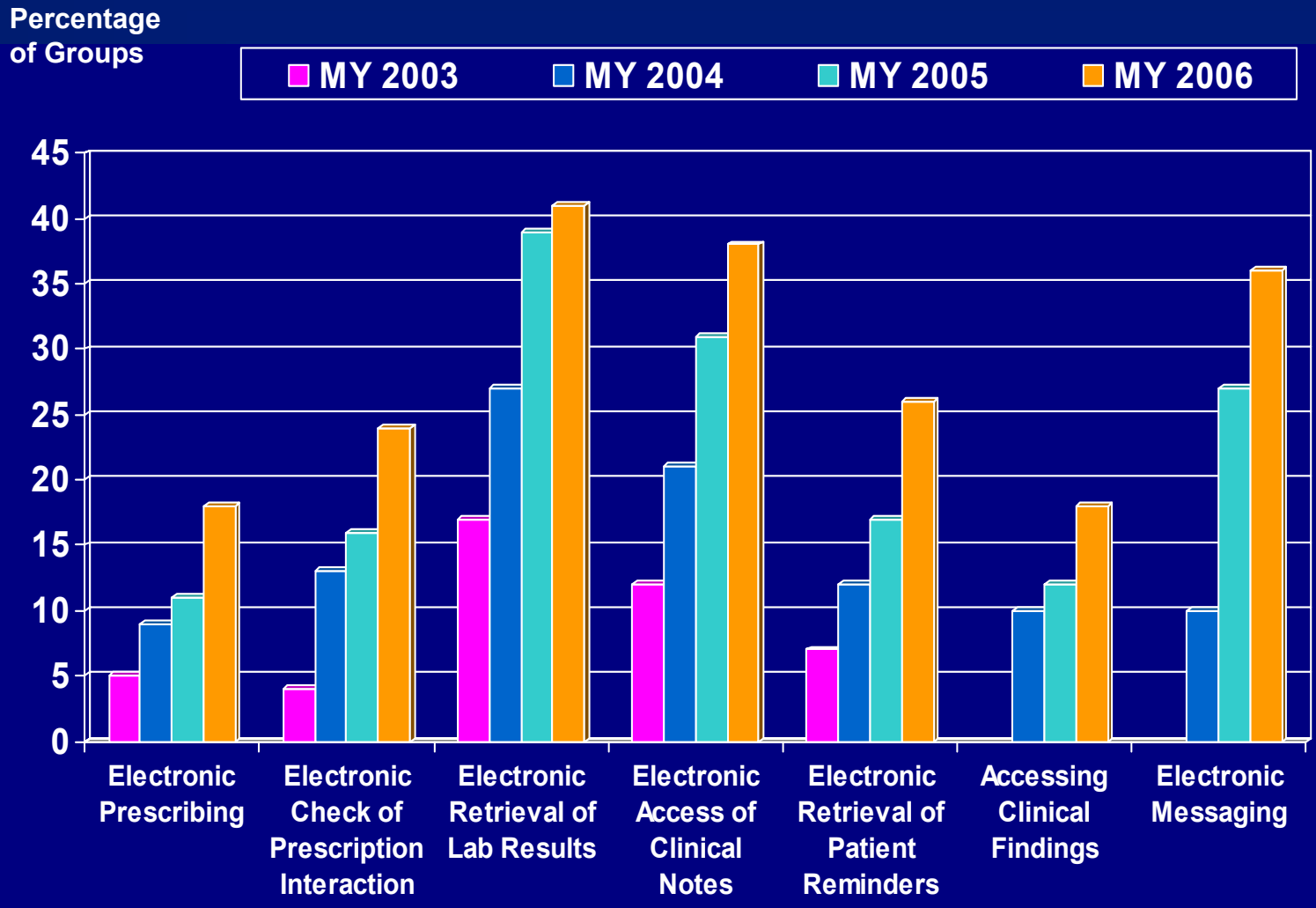
Clinical Measure	2005-2006 % Pt. Change	Trend Change (2003 or 2004 to 2006)	Change in Spread (10-90th %)
Asthma	+1.54	NA	+0.57
Breast Cancer	+1.54	+5.09	-0.01
Cervical Cancer	+1.75	+9.08	-11.46
Chlamydia	+1.97	+6.36	+5.42
MMR	+2.11	+3.75	-6.03
VZV	+2.47	+4.79	+1.48
HbA1c Screen	+2.11	+9.62	-20.15
LDL Screen	+2.48	NA	-4.88
URI treatment	-0.14	NA	+1.40

All figures are percentage point changes

Physician Organization Percentage Point Gains, by Performance Quartile

Clinical Measure	Lowest 1st quartile	2nd	3rd	Highest 4th quartile
Asthma	2.6	1.3	1.2	1.2
Breast Cancer	6.0	2.9	5.3	6.0
Cervical Cancer	16.8	8.5	6.3	4.6
Chlamydia	4.0	5.6	8.2	7.2
MMR	4.6	5.7	3.3	1.5
VZV	6.1	6.2	5.0	2.4
HbA1c Screen	19.4	6.1	7.4	5.8
LDL Screen	6.4	1.2	1.6	0.9
URI treatment	-0.6	-0.1	0.0	-0.1

An Increase in IT Investment: Will this Lead to Improvements in the Future?



Unintended Consequences

- Absence of empirical evidence, though much speculation
- At this stage given low power of current incentives, unlikely to be causing the disparities gap to widen further
  - Physician incentives of $500-$5000 are not leading to much behavior change, positive or negative
  - IHA results show reduced variation and greatest improvement for lowest performers
  - UK program saw a decrease in disparities gap
- In California, likely biggest "negative" impact has been causing providers to ignore those areas that are not measured and incentivized
  - Need for broader set of measures to mitigate this effect
  - Requires R&D investment to "feed the measures pipeline"

Mixed Perspectives on Results to Date

- POs value P4P and note the importance of the program has grown over past 3 years
  - Increased MD attention to the program
  - Improvements in data capabilities and QI support
  - Leadership attention
  - POs face challenges in monitoring their own performance
- Health plans success metrics: rating 2.5 (scale of 1-5)
  - Year-over-year changes in performance (*"Marginal improvements")*
  - Improvement in health of plan's members relative to HEDIS 90th percentile benchmarks (*"Still lagging significantly")*
  - Net savings accrued (*"No evidence of any savings to justify increased investment")*

Balancing Differing IHA Stakeholder Priorities Moving Forward

	Mean Score		
Priority Areas (1=low priority, 5=high priority)	**Medical Groups (n=35)**	**Health Plans (n=7)**	**Purchaser (n=2)**
Increasing incentive amount	**4.2**	2.0	**5.0**
Expanding clinical measures set	2.5	**3.9**	**4.5**
Providing technical assistance on how to improve	3.3	3.7	2.0
Retiring measures that have topped out	2.6	**4.2**	1.0
Adding specialty care measures	3.2	**4.2**	**3.5**
Uniform measures for all health plans	**4.6**	3.8	3.0
Aligning IHA measures with national measures	3.7	3.8	2.0
Addition of efficiency measures	2.9	**4.7**	**4.5**
Adding in other measures used by plans	2.5	3.0	2.0
Expanding to include Medicaid	2.1	2.0	2.0
Expanding to include Medicare Risk	**4.3**	2.8	**4.5**
Expanding to include PPO business	2.9	2.8	2.5

P4P Not the Sole Solution to the Quality and Costs Problems

- Modest performance gains from P4P Version 1.0
- P4P is a small fix to larger toxic payment system
  - *"We're working at the margins"*
- P4P can support efforts to reign in costs and improve quality
- Other policy levers are needed in conjunction with P4P to align and strengthen signals to providers, which will increase the likelihood for impact
  - Measurement and accountability
  - Quality improvement support
  - Public reporting or transparency
  - Investments in information systems
  - More fundamental payment reform

Looking Ahead…

- What will be the impact of redesigned P4P experiments (Version 2.0)?
- Need for independent evaluations to assess impact
  - Absent investments in evaluation, we will continue to make policy without good evidence
  - Change takes time, so assessments of impact need to look over time
- Providers struggling to know which changes to make to drive improvements


RAND
HEALTH