Pay for Performance in High-Medicaid Practices

Tenth National Pay for Performance Summit

March 2 2014

San Francisco, CA
Disclosures

Naomi Bardach and co-authors have documented that they have no financial relationships to disclose or Conflicts of Interest (COIs) to resolve.
Objectives

- Brief overview of what’s known about P4P in Medicaid populations in the outpatient setting
- Describe the results of a P4P program in clinics serving a high proportion of Medicaid patients
  - Designed to address known limitations
- Discuss potential implications for P4P program design and future research
Background

- It remains unclear whether pay for performance programs are effective, particularly with small group safety net providers
Background

- The programs have the potential to increase health care disparities
  - Rewards go to highly-resourced providers who can achieve benchmarks while low-resourced providers cannot achieve them and do not receive rewards
  - Focus on the “low hanging fruit” healthier patients
  - “Creaming” (decreased access for high-risk patients)

Lindenauer, et al. NEJM 2007
Werner, et al. JAMA 2005
Background

• In a New York Medicaid-focused managed care P4P program focused on Diabetes there was no change in incentivized practices on process and outcome measures
  – Authors suggest that this may have been due to lack of infrastructure

• For outpatient providers caring for commercially insured children in MA, P4P had a small though statistically significant effect on process measures

Objectives

- Brief overview of what’s known about P4P in Medicaid populations in the outpatient setting
- **Describe the results of a P4P program in clinics serving a high proportion of Medicaid patients**
  - *Designed to address limitations*
- Discuss potential implications for P4P program design and future research
Design focused on:

- Different design from benchmarking approach in order to avoid penalizing under-resourced providers and discouraging poor performers from participating
- Pay more for achieving a metric in sicker patients or patients with socio-economic stressors
- Infrastructure is in place to support improvement
- Include outcomes as well as processes
Larger context of program: PCIP

- **Primary Care Information Project**
  - *Focus on bringing EHRs to providers for NYC underserved*
  - *Same EHR with clinical decision support*
  - *Technical assistance or support for quality improvement, meaningful use, patient centered medical home*
  - *Funding: DOHMH NYC*

- **Pay for Performance program within PCIP**
  - *Health e-Hearts*
  - *Funded by the Robin Hood Foundation, interested in improving health for low income NYC communities*
Talk outline: 4 studies from PCIP

- P4P year 1
- P4P year 2 (new cohort enters)
- Survey data from years 1 and 2—potential mechanisms to explain control vs incentive differences in performance
- Unintended consequences

- Strengths
  - Pragmatic implementation with ongoing data stream
  - Longitudinal data with varying incentives and different levels of exposure
Pay-for-Performance in High-Medicaid Practices: Implications from a Cluster-Randomized Trial in New York City

JAMA, 2013 Sep 11;310(10):1051-9
Research question

• What is the effect of a piece-rate, graduated pay for performance program in small, EHR-enabled practice performance on cardiovascular outcomes and processes?
Study Design

- A cluster-randomized, controlled trial of incentives
  - *Clustered at the clinic level for randomization*
  - *Incentives also paid at the clinic level*
- Patients: > 18 years old
- Two program years, with the design of the program changing between year 1 and year 2
Population Year 1

- 84 small (1-2 providers) practices in New York City
- All practices were participants in Primary Care Improvement Project (PCIP)
  - Electronic Medical Record (EMR) with clinical decision support reminders for measures
  - Ongoing quality improvement site visits
## Incentive Structure

<table>
<thead>
<tr>
<th></th>
<th>Base Payment</th>
<th>Payment for High-Risk Patients</th>
<th>Total Possible Payment per Patient</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Insurance:</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Commercial Insurance:</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No IVD or DM</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Qualifying Insurance:</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Uninsured Medicaid</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Qualifying Co-Morbidities:</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>IVD or DM</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Combination of qualifying insurance and co-morbidity:</td>
<td>$20</td>
<td>$20</td>
<td>$20</td>
</tr>
<tr>
<td>Uninsured/Medicaid and IVD/DM</td>
<td>$20</td>
<td>$20</td>
<td>$20</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>Aspirin</th>
<th>BP Control</th>
<th>Cholesterol Control</th>
<th>Smoking Cessation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Aspirin</td>
<td>-</td>
<td>-</td>
<td></td>
<td></td>
</tr>
<tr>
<td>BP Control</td>
<td>$20</td>
<td>$40</td>
<td>$40</td>
<td>$80</td>
</tr>
<tr>
<td>Cholesterol Control</td>
<td>$20</td>
<td>$40</td>
<td>$40</td>
<td>$80</td>
</tr>
<tr>
<td>Smoking Cessation</td>
<td>$20</td>
<td>$20</td>
<td>$20</td>
<td>$20</td>
</tr>
</tbody>
</table>

**Maximums:** $200 per patient. $100,000 per practice

IVD: Ischemic Vascular Disease; DM: Diabetes Mellitus
# Quality measures

| A | Antithrombotic Rx | Antithrombotic prescribed  
|   | Patients with Diabetes or IVD* |
| B | Blood pressure control("BP") | BP controlled (<140/90 or <130/80)  
|   | Patients with hypertension |
| S | Smoking Cessation Intervention | Intervention delivered  
|   | Patients who smoke |

*IVD: Ischemic Vascular Disease; TC: Total Cholesterol; LDL: Low Density Lipoprotein
Study Timeline

**YEAR 1**
- April 2009: Intervention starts Cohort 1 (n=84 clinics)
- April 2010: End of year 1

**YEAR 2**
- October 2010: Start of year 2 with Cohort 2 (n=60 new clinics, Total=140 clinics)
- October 2011: End of year 2
Study Timeline

April 2009
- Intervention starts Cohort 1 (n=84 clinics)

April 2010
- End of year 1

October 2010
- Start of year 2 with Cohort 2 (Total=140 clinics)

October 2011
- No intervention
Study Timeline

Quarterly performance reports for all clinics

April 2009
- Intervention starts Cohort 1 (n=84 clinics)

April 2010
- End of year 1

October 2010
- Start of year 2 with Cohort 2 (Total=140 clinics)

October 2011
- End of year 2

Ongoing quality improvement support site visits
Study Timeline

- **April 2009**: Intervention starts Cohort 1 (n=84 clinics)
- **April 2010**: End of year 1
- **October 2010**: Start of year 2 with 2\textsuperscript{nd} cohort (n=60 new clinics, Total=140 clinics)
- **October 2011**: End of year 2

**Payments**
- **Lump payment**
- **Quarterly payments**
Analysis

• Difference-in-differences approach to quantify the effect size in each cohort
  • *Compares the difference in performance change over time between intervention and control clinics*

• Mixed effects logistic regression to account for clustering of patients

• A treatment by time interaction term assessed the statistical significance of the effect
RESULTS
YEAR 1
## Baseline Characteristics of Intervention and Control Patients

<table>
<thead>
<tr>
<th>Patient Characteristics</th>
<th>Incentive</th>
<th>Control</th>
<th>P value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age, y</td>
<td>45.8 (6.7)</td>
<td>46.6 (4.8)</td>
<td>0.62</td>
</tr>
<tr>
<td>Male, %</td>
<td>42.0 (8.6)</td>
<td>39.8 (10.5)</td>
<td>0.48</td>
</tr>
</tbody>
</table>

**Year 1**
## Baseline Characteristics of Intervention and Control Clinics

<table>
<thead>
<tr>
<th>Clinic Characteristics</th>
<th>Incentive</th>
<th>Control</th>
<th>P value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Clinicians, median (IQR)</td>
<td>1 (1-2)</td>
<td>1 (1-2)</td>
<td>0.77</td>
</tr>
<tr>
<td>Patients, median (IQR)</td>
<td>2500 (1200-4607)</td>
<td>2000 (1100-3500)</td>
<td>0.45</td>
</tr>
<tr>
<td>Time since EHR implementation, mo</td>
<td>9.93 (4.47)</td>
<td>9.57 (4.44)</td>
<td>0.81</td>
</tr>
<tr>
<td>QI specialist visits</td>
<td>5.17 (3.43)</td>
<td>4.24 (2.73)</td>
<td>0.25</td>
</tr>
<tr>
<td>Insurance, %</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Commercial</td>
<td>33.8 (23.9)</td>
<td>32.1 (21.6)</td>
<td>0.89</td>
</tr>
<tr>
<td>Medicare</td>
<td>25.6 (22.0)</td>
<td>26.8 (17.6)</td>
<td>0.32</td>
</tr>
<tr>
<td>Medicaid</td>
<td>35.3 (28.3)</td>
<td>35.7 (24.8)</td>
<td>0.88</td>
</tr>
<tr>
<td>Uninsured</td>
<td>4.3 (4.8)</td>
<td>4.7 (4.9)</td>
<td>0.60</td>
</tr>
</tbody>
</table>
## Results: Baseline Performance

<table>
<thead>
<tr>
<th>Measure</th>
<th>Control (%)</th>
<th>Incentive (%)</th>
<th>P value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Aspirin therapy, CAD or DM</td>
<td>54.4</td>
<td>52.6</td>
<td></td>
</tr>
<tr>
<td>BP control, no comorbidities</td>
<td>31.8</td>
<td>52.1</td>
<td>&lt;0.05</td>
</tr>
<tr>
<td>BP control, DM</td>
<td>10.4</td>
<td>16.8</td>
<td>&lt;0.05</td>
</tr>
<tr>
<td>Smoking cessation intervention</td>
<td>19.1</td>
<td>17.1</td>
<td></td>
</tr>
</tbody>
</table>

Year 1
## Results: Baseline Performance

<table>
<thead>
<tr>
<th>Measure</th>
<th>Control (%)</th>
<th>Incentive (%)</th>
<th>P value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Aspirin therapy, CAD or DM</td>
<td>54.4</td>
<td>52.6</td>
<td></td>
</tr>
<tr>
<td>BP control, no comorbidities</td>
<td>31.8</td>
<td>52.1</td>
<td>&lt;0.05</td>
</tr>
<tr>
<td>BP control, DM</td>
<td>10.4</td>
<td>16.8</td>
<td>&lt;0.05</td>
</tr>
<tr>
<td>Smoking cessation intervention</td>
<td>19.1</td>
<td>17.1</td>
<td></td>
</tr>
</tbody>
</table>

Year 1
Improvements in Performance

Aspirin therapy, CAD or DM
BP control, no comorbidities
BP control, DM
Smoking cessation intervention

Control (%)
Incentive (%)

6.0**
5.5*
7.8**
4.7**

*<0.05
**<0.01
Year 1 summary

- All groups improved
- Incentive group had greater improvements on processes and intermediate outcomes for patients with and without comorbidities
- Patients with hypertension and diabetes did not fare worse, with that population benefitting substantially given low baseline rates of BP control
YEAR 2
(UNPUBLISHED)
Objectives

• To assess the effects of the incentive in the second year of the program

• Program was modified:
  – Higher amounts
  – Quarterly payments rather than one lump sum at the end

• We compare the incentive effect between the clinics participating in their second year of the program (exposed) and a sample of clinics newly enrolled and randomized in the second year (naïve)
## Incentive Structure

<table>
<thead>
<tr>
<th></th>
<th>Base Payment</th>
<th>Payment for High-Risk Patients</th>
<th>Total Possible Payment per Patient</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Insurance:</strong> Commercial</td>
<td>Aspirin: -</td>
<td>Qualifying Insurance: Insured</td>
<td>Combination of qualifying insurance and co-morbidity: Uninsured/Medicaid and IVD/DM</td>
</tr>
<tr>
<td></td>
<td>BP Control: $50</td>
<td>co-morbidities: Uninsured/Medicaid and IVD/DM</td>
<td>Uninsured/Medicaid and IVD/DM</td>
</tr>
<tr>
<td></td>
<td>Smoking Cessation: $50</td>
<td>Qualifying Co-Morbidities: IVD or DM</td>
<td>Qualifying Co-Morbidities: IVD or DM</td>
</tr>
</tbody>
</table>

| Aspirin                  | -            | -                              | $50                              | $50 |
| BP Control               | $50          | $100                           | $100                             | $150 |
| Smoking Cessation        | $50          | $50                            | $50                              | $50 |

**Maximums**: $400 per patient. $100,000 per practice

IVD: Ischemic Vascular Disease; DM: Diabetes Mellitus
Population Year 2

- Small (1-2 providers) practices in New York City
- Exposed cohort, n=80 clinics (Year 1 participants)
- Naïve cohort, n=60 clinics
- All practices were participants in Primary Care Improvement Project (PCIP)
  - Electronic Medical Record (EMR) with clinical decision support reminders for measures
  - Ongoing quality improvement site visits
## Baseline year performance

<table>
<thead>
<tr>
<th>Measure</th>
<th>Naïve</th>
<th></th>
<th>Exposed</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Control (%)</td>
<td>Incentive (%)</td>
<td>Control (%)</td>
<td>Incentive (%)</td>
</tr>
<tr>
<td>Antithrombotic therapy, IVD or DM</td>
<td>64.2</td>
<td>57.7</td>
<td>64.7</td>
<td>70.5</td>
</tr>
<tr>
<td>Blood pressure control, no comorbidities</td>
<td>57.1</td>
<td>49.8</td>
<td>42.8</td>
<td>51.1</td>
</tr>
<tr>
<td>Blood pressure control, in DM</td>
<td>32.0</td>
<td>30.2</td>
<td>26.1</td>
<td>27.8</td>
</tr>
<tr>
<td>Smoking Cessation intervention</td>
<td>16.6</td>
<td>18.8</td>
<td>24.4</td>
<td>32.6</td>
</tr>
</tbody>
</table>
Performance Changes Year 2, Naive

- Aspirin Therapy: 4.9**
- BP, no comorbidities: 9.6**
- BP, DM: 9.2**
- Smoking Intervention: 9.5**

**<0.01
Performance Changes Year 2, Exposed

Aspirin Therapy BP, no comorbidities
BP, DM
Smoking Intervention

Incentive
Control

-7.3**
-3.1*
-1.9

4.6**

*<0.05
**<0.01
Conclusion

- P4P with a higher incentive and quarterly payments was effective in the naïve cohort, with apparently larger effect sizes than in the first year of the program.
- However, among exposed clinics, control clinics improved more than intervention clinics on blood pressure measures, though all clinics improved.
Implications

• Unclear why control clinics improved more rapidly than intervention clinics in the exposed cohort on BP control
  • Secular trends
  • “Low hanging fruit”
  • P4P program signals policy attention to measures, with incentivized clinics responding more rapidly
• Piece-rate, graduated incentives are effective in the first year of the program
• Larger incentives may lead to larger effect sizes in the first year of the program
Unintended consequences
The intended and unintended consequences of quality improvement interventions for small practices in a community-based electronic health record implementation project.

Med Care. 2014 Sep;52(9):826-32
Research Question

- What is the effect of incentives on non-incentivized measures in the Health e-Hearts program?
Methods

• All P4P practices
• Unincentivized and not reported on measures:
  – Documentation
    • BMI measurement
  – Processes
    • Appropriate asthma rx
    • Pneumococcal vaccine
  – Intermediate Outcomes
    • HbA1C control
Decreased performance over time on unincentivized measures

- Relative to predicted performance in the same set of practices
- Underpowered to look at specific measures or by cohort, due to limitations in data transfer from EHR

<table>
<thead>
<tr>
<th>Incentivized measures</th>
<th>Unincentivized measures</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intensity of exposure</td>
<td>Intensity of exposure</td>
</tr>
<tr>
<td>6 months</td>
<td>12 months</td>
</tr>
<tr>
<td>6.7**</td>
<td>10.1*</td>
</tr>
<tr>
<td>10.1*</td>
<td>-4.9**</td>
</tr>
<tr>
<td>6 months</td>
<td>12 months</td>
</tr>
<tr>
<td>18 months</td>
<td>18 months</td>
</tr>
<tr>
<td>-7.6**</td>
<td>-8.3*</td>
</tr>
</tbody>
</table>

*P<0.05; ** P<0.01
Results from Preliminary Analysis of Individual Measures by Cohort (unpublished)
## Baseline Year 2 Performance: Unincentivized Measures

<table>
<thead>
<tr>
<th>Measure</th>
<th>Naïve</th>
<th>Exposed</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Control (%)</td>
<td>Incentive (%)</td>
</tr>
<tr>
<td><strong>Documentation</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Smoking Status</td>
<td>57.0</td>
<td>68.7</td>
</tr>
<tr>
<td>Depression Screening</td>
<td>5.6</td>
<td>6.1</td>
</tr>
<tr>
<td>EtOH screening</td>
<td>34.6</td>
<td>41.5</td>
</tr>
<tr>
<td><strong>Processes</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>A1C testing</td>
<td>23.3</td>
<td>35.3</td>
</tr>
<tr>
<td><strong>Intermediate outcome</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>A1C control (&lt;7%)</td>
<td>0.3</td>
<td>2.7</td>
</tr>
</tbody>
</table>
Performance Changes Year 2, Naive

-5.3**

Smoking Status

-2.9**

Depression Screening

-6.3**

EtOH screening

-0.7

A1C testing

4.2**

A1C control (<7%)

**<0.01
Performance Changes Year 2, Exposed

<table>
<thead>
<tr>
<th>Measure</th>
<th>Control</th>
<th>Incentive</th>
<th>Significance</th>
</tr>
</thead>
<tbody>
<tr>
<td>Smoking Status</td>
<td>0.7</td>
<td>0.7</td>
<td></td>
</tr>
<tr>
<td>Depression Screening</td>
<td>-0.3</td>
<td>0.9**</td>
<td></td>
</tr>
<tr>
<td>EtOH screening</td>
<td>-1.3</td>
<td>-1.3</td>
<td></td>
</tr>
<tr>
<td>A1C testing</td>
<td>2.5**</td>
<td>2.5**</td>
<td>**&lt;0.01</td>
</tr>
<tr>
<td>A1C control (&lt;7%)</td>
<td></td>
<td></td>
<td>**&lt;0.01</td>
</tr>
</tbody>
</table>
Summary

• On individual non-incentivized measures, P4P program had differential effects

• For documentation measures, both groups improved, with the incentive group improving at a slower rate than the control group in the first year
  – The incentive group caught up by the end of the second year

• For DM process measure, incentive and control groups behaved similarly in both years

• For DM intermediate outcome measure, incentive groups improved at a faster rate than control groups, in both years
Discussion

- The summative P4P incentive effect on a combination of unincentivized measures found worse performance over time

- It may be that the summative approach obscured differences in the incentive effect by type of measures (documentation vs. processes vs. outcomes) and differences in effects over two different cohorts
  - Less focus in the incentive group on documentation in the first year
  - The same way the intended effects of the P4P incentive diminished over time, with the control group catching up, the incentive group caught up on the unincentivized measures
Discussion

- Population of focus for unincentivized measures may change the story
  - Design of incentive programs needs to consider effects on other populations (asthma, depression screening)
  - There may be a quality spillover effect on other measures of care within the same population (diabetics)
Objectives

• Brief overview of what’s known about P4P in Medicaid populations in the outpatient setting

• Describe the results of a P4P program in clinics serving a high proportion of Medicaid patients
  
  Designed to address known limitations

• Discuss potential implications for P4P program design and future research
Design implications

- Graduated incentive design
- Piece-rate design
- Advice to clinics and policy-makers: invest in using QI tools—decision support, registries
Design implications (cont.)

• Program successes may occur through multiple mechanisms—control clinics are potentially affected as well
  – Consider rotating clinic incentive eligibility
    • Enables clinics to build capacity
    • Renews focus in incentive years, potentially prolonging sustainability

• Assess for performance on unincentivized measures within the same population and different populations
  – Consider population when suites of incentivized measures and unincentivized measures
Acknowledgements

• New York DOHMH, PCIP
  – Jason Wang, PhD
  – Samantha F. DeLeon, PhD
  – Sarah C. Shih, MPH

• Phil R. Lee Institute for Health Policy Studies, UCSF
  – L. Elizabeth Goldman, MD
  – John Boscardin, PhD
  – R. Adams Dudley, MD MBA

Thomas Frieden, MD MPH—Director, Centers for Disease Control and Prevention
Farzad Mostashari, MD ScM—Former National Coordinator for Health Information Technology, at Brookings Institution
Questions?

bardachn@peds.ucsf.edu
Need for prolonged technical assistance to achieve improvements
Small Practices’ Experience With EHR, Quality Measurement, and Incentives


Authors:
Rohima Begum, MPH
Mandy Smith Ryan, PhD
Chloe H. Winther, BA
Jason J. Wang, PhD
Naomi S. Bardach, MD, MAS
Amanda H. Parsons, MD
Sarah C. Shih, MPH
R. Adams Dudley, MD, MBA
Study Objectives

• To assess clinician attitudes towards P4P measures and intervention
• To assess clinician use of the EHR
• To explore potential explanatory variables for differences in performance on P4P metrics between control and intervention clinics
Setting and Design

- Lead clinicians from each participating practice in the Health eHearts program (years 1 and 2)
- Survey administered at the end of the second year (October 2011)
RESULTS
Quality Reports

- Understood the information in the reports
- Prioritization of ABCS was appropriate
- Received and reviewed quality reports
- ABCS were clinically meaningful
- Reports had enough information
- Reports accurately reflected progress on ABCS

b<0.05
N=104 (74% response rate)
EHR Functionalities

- Clinical Decision Support System
- Smart forms
- Use registry to generate patient lists
- Order set (already within the EHR)
- Flow sheet (part of progress note)

Control
Incentive

b<0.05
EHR Functionalities

- Clinical Decision Support System
- Smart forms
- Use registry to generate patient lists
- Order set (already within the EHR)
- Flow sheet (part of progress note)

Control
Incentive

Percent
Summary

- Providers in incentive groups report getting “signal” more than control groups
- Providers in incentive groups report more buy-in than control groups re: report accuracy
- Providers in incentive groups report more comfort with functionality that supported success on the incentivized measures
- Data suggest more focus on documentation in the control providers
- These mechanisms may contribute to differences in performance