Pay-for-Performance in Health Care: Trends, Impact, and Policy Issues

Meredith B. Rosenthal, Ph.D.
Harvard School of Public Health
December 15, 2006
A Snapshot of Pay-for-Performance in the U.S.

- Inventories of programs across all types of payers document more than 100 extant pay-for-performance programs\(^1\)
- In a national survey, 52% of HMOs (covering 81% of enrollees) report using pay-for-performance\(^2\)

---

What Types of Health Plans Use Pay-for-Performance?

- HMO programs most common, particularly those with:
  - PCP gatekeeping
  - Capitation
- Anywhere but the South
- Those in markets where employers use performance-contracting with health plans
How Are Pay-for-Performance Programs Structured?

- Physicians (medical groups) about twice as likely as hospitals to be target
- Average of 5 performance measures: clinical process measures most common
- Maximum bonus 5-10% of pay for physicians, 1-2% for hospitals
- Rewards for reaching fixed threshold dominate; only 23% reward improvement
Pay-for-Performance and IT

- ~60% of HMOs that have physician pay-for-performance include IT measures in their programs
- ~30% of HMOs with hospital pay-for-performance include CPOE measures
- Standardized instruments have been developed for IT process measures (to capture effective use)
Increasing Inclusion of Specialists and Hospitals in Pay-for-Performance

Source: Rosenthal et al., Climbing Up the Pay-for-Performance Learning Curve, Manuscript, Harvard University 2006.
Increasing Emphasis on Outcomes, IT, Cost-Efficiency

Source: Rosenthal et al., Climbing Up the Pay-for-Performance Learning Curve, Manuscript, Harvard University 2006.
Overview of Impact Estimates

- Rigorous studies of pay-for-performance in health care are few (17 since 1980)
- Overall findings are mixed: many null results even for large dollar amounts
- But in many cases negative findings may be due to short-term nature, small incentives
- Evidence suggests pay-for-performance can work but also can fail
Case Study #1: The Integrated Healthcare Association (CA)

- Probably largest effort in U.S.
- Statewide in California
- Capitated, multispecialty medical groups targets
- Core measures common to 7 plans, coordinated data collection
- Public reporting of all-payer data
## 2004 IHA Measure Set

<table>
<thead>
<tr>
<th>Domain (Weight)</th>
<th>Measures</th>
</tr>
</thead>
<tbody>
<tr>
<td>Clinical (40%)</td>
<td>Mammography</td>
</tr>
<tr>
<td></td>
<td>Cervical cancer screening</td>
</tr>
<tr>
<td></td>
<td>Childhood immunization</td>
</tr>
<tr>
<td></td>
<td>HbA1c Testing</td>
</tr>
<tr>
<td></td>
<td>LDL Cholesterol Testing</td>
</tr>
<tr>
<td></td>
<td>Asthma medication management</td>
</tr>
<tr>
<td>Patient Experience (40%)</td>
<td>Various patient survey composites</td>
</tr>
<tr>
<td>IT (20%)</td>
<td>Integration of electronic data sets</td>
</tr>
<tr>
<td></td>
<td>Point of care decision support</td>
</tr>
</tbody>
</table>
IHA Reported Impact

- All targeted measures improved
- Average improvement ~ 3 percentage points (less for patient experience)
- Many measures had no valid baseline comparison
- IT measures showed strongest results
- No way to establish how much due to pay-for-performance
IHA Part II: PacifiCare Quality Incentive Program (QIP)

- Evaluation using one member plan’s trend and comparison data suggests effects on process measure improvement minimal (only cervical cancer shows impact).
- Also gives credence to concern that rewarding all providers who can meet a fixed performance target will not stimulate uniform improvement.
- Implication: pay-for-performance programs as now designed may be good screening devices but will yield little QI.
# Quality Improvement and Payments to Groups with High, Middle or Low Baseline Performance

<table>
<thead>
<tr>
<th>Quality Domain</th>
<th>Total PacifiCare Members</th>
<th>Pre-QIP Rate</th>
<th>Post-QIP Rate</th>
<th>Improvement (Post-Pre)</th>
<th>Bonuses Paid in Year 1</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cervical Cancer Screening</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Group 1</td>
<td>597,091</td>
<td>53.6%</td>
<td>56.0%</td>
<td>2.5% (0.8%)</td>
<td>$ 436,618</td>
</tr>
<tr>
<td>Group 2</td>
<td>287,610</td>
<td>40.8%</td>
<td>48.1%</td>
<td>7.4% (2.4%)</td>
<td>$ 127,632</td>
</tr>
<tr>
<td>Group 3</td>
<td>305,041</td>
<td>23.0%</td>
<td>34.1%</td>
<td>11.1% (3.9%)</td>
<td>$ 26,859</td>
</tr>
</tbody>
</table>
Case Study #2: National Health Service General Practitioner Contract

- 146 performance indicators (clinical, organizational, patient experience, additional services)
- Subsidies for equipment and staff
- Bonuses for performance up to 25% of pay
- Penalties built in for very low performance
Scoring in the NHS GP Contract

- Rewards under the GP contract are based on point system
- Total points vary by measure – reflecting both importance and usefulness of measure
- Within measures, there are population based thresholds: e.g., one point for screening at least 25% of patients; 2 points for screening at least 50%, etc.
- Exclusion of patients from denominator may be requested
GP Contract Initial Results

- Practices received on average 95.5% of available points
- Actual adherence to each of the clinical process indicators average 83.4% overall
- Median exception reporting was 6% but some practices excluded more than 15%
- Exception reporting largest factor predicting performance
Summary

- Pay-for-performance is theoretically defensible, but little is known about best practices
- Trends suggest increasing breadth, depth
- Design improvements needed to improve quality, minimize unintended consequences
Key Issues for the Future

1. Current pay-for-performance programs not consistent with incentive design principles
   - Need to align incentives with the true cost of delivering the care we want (including foregone revenues)
   - Incentives should reward all increments of high-value care, not just “best” providers

2. Pay-for-performance is likely to focus increasingly on ROI:
   - Quality improvement with savings (e.g., reducing complications)
   - Incorporation of efficiency measures (quality-adjusted cost per episode)

3. What CMS ultimately does will clearly matter for ultimate impact