New York Perspective: Pay for Performance in Medicaid Managed Care

3rd Annual Pay for Performance Summit
February 28, 2008

Joseph Anarella, MPH
Director, Quality Measurement and Improvement, NYSDOH

Thomas Foels, MD,
Medical Director, Independent Health Association, Inc.

Robert Berenson, MD,
Senior Fellow, The Urban Institute
Medicaid in New York State

- $49 billion program (40% of state budget); 4.1 million beneficiaries;

- Enrollment in MAMC is over 2.57 million (62% of total), served by 23 health plans;

- SSI roll-out complete in late 2008, will add an additional 200,000;

- On deck? HIV, MC/MA duals (600,000)
How We Reward Quality?

- Public reporting of Quality Assurance Reporting Requirements (QARR) - web, consumer guides, annual report

- The DOH has legislative authority to direct beneficiaries who do not choose a plan to high performing plans. This began in 2000.

- Bonus premium payments began in fall of ’02. Plans initially could earn up to 1% in additional premium. That amount was increased to 3% in 2004.
What are our goals for Incentive?

- Accelerate improvement; reduce, eliminate disparities;
- ‘Business case’ for investing in quality
  - Empower medical directors/QI staff with CFOs, COOs, CEOs
- Align with other P4P initiatives
  - Health plan initiated
  - Private payors (Bridges to Excellence)
Methodology

- 150 Points
  - HEDIS/NYS-specific data = 100 points
    Benchmark = 75th percentile from 2 years prior.
  - CAHPS data = 30 points
    Benchmark = At or above statewide average
  - Compliance (2 measures - fiscal and provider network reports)
    Benchmark = No statements of deficiency.
Methodology

- Plans can earn 3%, 2.25%, 1.5%, .75% or no additional premium depending on their overall score.

- Plans that earn no incentive get no autoassignment.

- Measures change annually with NCQA rotation/DOH priorities.

- Typically 2/3 of plans qualify for some level of award.
Results

- Issues looking at changes over time due to:
  - rotation of measures
  - Changes in specifications (e.g. asthma)
  - Old measures dropped, new measures added
  - Measures dropped during the year by NCQA

- Measured improvement by examining:
  - Year a measure was introduced
  - Next time that measure was included
  - Last time the measure was included
Performance Improvement

- Immunization 0-2: 64, 62, 73
- WC 0-15: 58, 64, 65
- WC 3-6: 71, 73, 76
- Adolescent WC: 45, 45, 49
Performance Improvement

Asthma 5-56: 1st 60, 2nd 62, last 70
Postpartum: 1st 59, 2nd 63, last 70
Diabetes poor control: 1st 42, 2nd 37, last 35
Shrinking Disparities

MH 30 day Medicaid

MH 30 day Commercial

diabetes poor Medicaid

diabetes poor Commercial
Shrinking Disparities
Shrinking Disparities

Graph showing data for postpartum Medicaid and postpartum Commercial.
Satisfaction with Care

Commercial

Medicaid

- Rating of Health Plan
  - Comm 2005: 62
  - Comm 2006: 67
  - Medicaid 2005: 75
  - Medicaid 2003: 73

- Rating of Specialist
  - Comm 2005: 80
  - Comm 2006: 78
  - Medicaid 2005: 74
  - Medicaid 2003: 74

- Rating of Doctor
  - Comm 2005: 79
  - Comm 2006: 77
  - Medicaid 2005: 79
  - Medicaid 2003: 79

- Rating of Health Care
  - Comm 2005: 75
  - Comm 2006: 79
  - Medicaid 2005: 77
  - Medicaid 2003: 74
Incentive Payments to Date

<table>
<thead>
<tr>
<th>Year</th>
<th>Amount</th>
</tr>
</thead>
<tbody>
<tr>
<td>2002*</td>
<td>$6.4</td>
</tr>
<tr>
<td>2003*</td>
<td>$6.9</td>
</tr>
<tr>
<td>2004**</td>
<td>$20.9</td>
</tr>
<tr>
<td>2005**</td>
<td>$43.7</td>
</tr>
<tr>
<td>2006**</td>
<td>$50.0</td>
</tr>
<tr>
<td>2007**</td>
<td>$62.3</td>
</tr>
</tbody>
</table>

*1% incentive **3% incentive
Observations

- We’ve got the plans attention.
  - Rates are increasing
  - Disparities between payers shrinking;
- We see more:
  - Experimentation
  - Physician incentives
  - IT investment
  - Case management
Issues

- Reward improvement or good quality?
- Are the best really the best?
- Studying for the test
- Sustainability
  - From both a state and plan perspective
- Purity (competition for P4P measures)
  - (e.g. reg. compliance; retention measure being considered for 2009)
What is Ahead?

- Beyond P4P.....supporting improvement
- Focused approach?
- Incenting use of HIT
- ‘No-pay’ for ‘no-performance’?
Questions?

Joe Anarella
jpa02@health.state.ny.us
(518) 486-9012
Independent Health:
The Health Plan Perspective

Thomas Foels, MD MMM
Medical Director
drfoels@independenthealth.com
Independent Health

Upstate (Western) New York
8 counties (2 urban: Buffalo, Niagara Falls)

380,000 covered lives
  25,000 Medicaid
  45,000 Medicare
  310,000 Commercial

Physicians
  Many solo / small group (15-20% EHR)
  1,200 PCP
  2,400 SCP

“Medicaid Provider Network” vs. Commercial Network
2007 NYS Medicaid Incentive Results
(2006 dates of service)

NCQA Clinical measures (40)  8-10 Selected

5 CAHPS measures  3 Selected

3 Compliance  3 Selected
<table>
<thead>
<tr>
<th>Service</th>
<th>2007 SWA</th>
<th>2007 IHA</th>
<th>Difference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Smoking advice</td>
<td>72%</td>
<td>86%</td>
<td>14%</td>
</tr>
<tr>
<td>DM BP &lt; 130/80</td>
<td>30%</td>
<td>39%</td>
<td>9%</td>
</tr>
<tr>
<td>Antidepressants acute phase</td>
<td>42%</td>
<td>51%</td>
<td>9%</td>
</tr>
<tr>
<td>Appropriate asthma Rx 5-56</td>
<td>89%</td>
<td>97%</td>
<td>8%</td>
</tr>
<tr>
<td>Antidepressants continuation</td>
<td>27%</td>
<td>34%</td>
<td>7%</td>
</tr>
<tr>
<td>Adolescent well</td>
<td>49%</td>
<td>55%</td>
<td>6%</td>
</tr>
<tr>
<td>F/U in pt admission behavioral 7 days</td>
<td>60%</td>
<td>66%</td>
<td>6%</td>
</tr>
<tr>
<td>DM BP &lt; 130/80</td>
<td>30%</td>
<td>39%</td>
<td>9%</td>
</tr>
<tr>
<td>DM DRE</td>
<td>57%</td>
<td>60%</td>
<td>3%</td>
</tr>
<tr>
<td>DM nephropathy</td>
<td>80%</td>
<td>81%</td>
<td>1%</td>
</tr>
<tr>
<td>Control BP</td>
<td>60%</td>
<td>60%</td>
<td>0%</td>
</tr>
<tr>
<td>F/U in pt admission behavioral 30 days</td>
<td>76%</td>
<td>76%</td>
<td>0%</td>
</tr>
<tr>
<td>Breast cancer screen</td>
<td>62%</td>
<td>60%</td>
<td>-2%</td>
</tr>
<tr>
<td>A1C good control</td>
<td>35%</td>
<td>32%</td>
<td>-3%</td>
</tr>
<tr>
<td>Ongoing prenatal care</td>
<td>70%</td>
<td>65%</td>
<td>-5%</td>
</tr>
<tr>
<td>Postpartum</td>
<td>70%</td>
<td>65%</td>
<td>-5%</td>
</tr>
<tr>
<td>Imaging LBP</td>
<td>82%</td>
<td>77%</td>
<td>-5%</td>
</tr>
<tr>
<td>DM LDL &lt; 100</td>
<td>39%</td>
<td>34%</td>
<td>-5%</td>
</tr>
<tr>
<td>LBW at level II-IV facility</td>
<td>80%</td>
<td>75%</td>
<td>-5%</td>
</tr>
<tr>
<td>DM A1C test</td>
<td>86%</td>
<td>80%</td>
<td>-6%</td>
</tr>
<tr>
<td>Antidepressant optimal contact</td>
<td>29%</td>
<td>23%</td>
<td>-6%</td>
</tr>
<tr>
<td>DMRD therapy rheum arthritis</td>
<td>72%</td>
<td>65%</td>
<td>-7%</td>
</tr>
<tr>
<td>Annual monitoring rheum arth Rx</td>
<td>72%</td>
<td>65%</td>
<td>-7%</td>
</tr>
<tr>
<td>DM lipid test</td>
<td>85%</td>
<td>77%</td>
<td>-8%</td>
</tr>
<tr>
<td>Poor A1C control</td>
<td>65%</td>
<td>56%</td>
<td>-9%</td>
</tr>
<tr>
<td>Inappropriate BP bronchitis</td>
<td>28%</td>
<td>18%</td>
<td>-10%</td>
</tr>
<tr>
<td>F/U ADHD initiation</td>
<td>39%</td>
<td>28%</td>
<td>-11%</td>
</tr>
<tr>
<td>Lead testing</td>
<td>86%</td>
<td>74%</td>
<td>-12%</td>
</tr>
</tbody>
</table>
### CAHPS member survey

<table>
<thead>
<tr>
<th>Service</th>
<th>2007 SWA</th>
<th>2007 IHA</th>
</tr>
</thead>
<tbody>
<tr>
<td>Overall sat health plan</td>
<td>75%</td>
<td>87%</td>
</tr>
<tr>
<td>Getting care</td>
<td>70%</td>
<td>77%</td>
</tr>
<tr>
<td>Customer service</td>
<td>75%</td>
<td>78%</td>
</tr>
<tr>
<td>Services quickly</td>
<td>74%</td>
<td>77%</td>
</tr>
<tr>
<td>Rating doctor</td>
<td>79%</td>
<td>79%</td>
</tr>
</tbody>
</table>
2007 Health Plan Performance vs State Wide Ave

Percent Variable from State Wide Ave

-15% -10% -5% 0% 5% 10% 15% 20%
Independent Health Medicaid Incentive

<table>
<thead>
<tr>
<th>Year</th>
<th>Award Value (percentage of premium)</th>
</tr>
</thead>
<tbody>
<tr>
<td>2003</td>
<td>3%</td>
</tr>
<tr>
<td>2004</td>
<td>3%</td>
</tr>
<tr>
<td>2005</td>
<td>3%</td>
</tr>
<tr>
<td>2006</td>
<td>2%</td>
</tr>
<tr>
<td>2007</td>
<td>1%</td>
</tr>
</tbody>
</table>

Performance Score

Incentive Award

Performance Level
Possible Paths to Declining Award Performance

“Fall Behind”
\begin{itemize}
  \item Actual performance deteriorates
\end{itemize}

“Others Gain Ground”
\begin{itemize}
  \item Relative performance deteriorates
\end{itemize}

“Luck of the Draw”
\begin{itemize}
  \item Favorable metric selection followed by unfavorable metric rotation
\end{itemize}
Lessons Learned #1

Do incentives promote quality improvement?

“Yes, but…”

“Phased approach: prefer beginning with limited focus and introduce new measures over time.”

“It did cause us to focus on areas that were otherwise not a high priority.”
Lessons Learned #2

Does a monetary incentive matter to health plans?

“Public reporting is an equally strong driver”

“The total award value at stake is more than sufficient to get our attention.”

“Award money was not directly reinvested in programs initially. We may have become complacent during the first 3 years because of our success.”

“Temptation to ‘study to the test’ ”
Lessons Learned #3:

**Improvement is difficult: Physician Network Perspective**

Provider network distinct from commercial network

Aligned physician incentives less effective
* Salaried physicians
* Unionized staff
* Rotating metric selection

Physician attribution is difficult
  Auto-assignment of Medicaid members

Actions of one provider can drive metrics
  (ex. strep screening with one pediatrician)

Learning collaborative (systems improvement) an option
Lessons Learned #4

Improvement is difficult: Member Perspective

- Locating the member
- Lack of perceived “medical home”
- Auto-assignment of members
- Effectiveness of Outreach Workers
- Member incentives
Evaluation of the NYS DOH Quality Incentive Program

Robert A. Berenson, M.D.
Senior Fellow, The Urban Institute
Study Questions

• How do senior managers of health plans view and respond to the QI initiative?
• What impact has the QI program had on health plan performance?
• Do trends in performance differ between Medicaid plan enrollees and commercial?
• Is there evidence of an impact of the Q.I. Program on Medicaid enrollees?
Qualitative Study Method

• On site, 60 minute interviews using a respondent-specific protocol with narrow and open-ended questions, conducted in 2006
• Respondents – CEO, CFO, CMO, QD
• Some answers analyzed at the plan level, others at the respondent level
The Priority of the QI Program to Plans

- 65% of 89 respondents said “very important” and 31% “somewhat important”

- The importance relates to staff and provider network, to the state, to general reputation and, importantly, to the opportunity to obtain bonuses – not to competition for members
Approaches Targeted to Enrollees

- Direct member outreach through mailings and phone calls (12 plans thought very successful)
- Build on home visits/disease mgt. for patients with asthma, diabetes – geared to increasing compliance on QARR measures
- Financial incentives – gift certificates to movies, hair salons, drug stores, toy stores
- Direct member outreach was also most common unsuccessful approach
Approaches Targeting Providers

• 9 plans thought this quite useful
• Used outreach and education generally
• Some plans used direct financial incentives, esp. “bill above” in plans paying on capitation
Priority Setting Among Measures

- Broad consensus that QARR measures reasonable and appropriate for measurement of plan performance
- Some respondents thought that plans cannot affect patient perceptions, i.e., CAHPS scores
- Practical problems with some measures
Priorities (cont.)

- Plans first focus on measures on which doing relatively poorly – “we don’t want to be an outlier.”
- P4P does not take place in isolation to other quality-related reporting
- 24% say measures they are most able to affect; 20% say focus on those with most clinical importance -- related to better outcomes
Priorities (cont.)

- There was some strategic behavior, but less than one might have thought, i.e. not focusing on measures where far from target (6 plans) or compatibility with other corporate goals (5 plans)
Constraints

- Difficulty getting requisite data – 14 plans (from both successful and unsuccessful ones)
  - Specific issues – problem of being part of larger systems, use of capitation, out-of-network providers
Constraints (cont.)

- 8 plans cited limited resources to be able to respond adequately
- Getting members to available services
  - Problem for preventive services
  - “churning” within Medicaid population
Plan-specific Constraints

- Almost all plans thought there were some.
- Most common was whether a plan was provider-owned.
  - Those not provider-owned but contracting with a provider thought they lacked influence.
  - But some provider-owned thought their provider owner might have a larger agenda, ignoring plan issues.
Plan-specific constraints (cont.)

• Type of provider network
• Small plans thought they were at disadvantage – limited resources for HIT and provider incentives, to “turn on a dime,” when measures announced, to get provider attention
• But some larger plans thought size and broader book of business obscured focus on QI program
• Recent mergers and acquisitions
Views of P4P Generally

- 89% of 82 respondents think that “having purchasers use financial incentives to health plans is a good strategy for improving quality”
- Only 3 thought that P4P was a bad idea
Reservoir of Skepticism about Measures Themselves

- 21 of 44 thought that measures used were an accurate reflection of quality provided to members. “They are as good as any”
- 21 of 44 thought that measures did not reflect quality – mostly negative about CAHPS – a “crap shoot”
- 23 comments on specific problems, but rarely consensus on which measures produce problems
Does Performance Reflect Quality or Ability to Report?

- 53% -- better data; 23% better care; 24% a mixture
- CEOs more likely to answer “better data”
- But many go on to assert the two are linked – need better data to improve care; some think linked temporally – first, need data, which permits improvement in care
- But, “Our plan does not provide health care, providers do... It’s all a number’s game.”
Perceived Strengths of the Program

- 80% identify basic strength of central purpose of providing incentives to have plans focus on quality
- Data-driven and relies on good measures
- Efficiency of using established measures
- Measures relevant to population served
- Here, identify lots of other
Perceived Weaknesses

- Only 10 of 90 without criticisms
- The three major ones:
  - Plans do not know measures until late in year
  - Some plans unfairly disadvantaged by size, location or type of network
  - Particular metrics are flawed
Variation Based on Respondents’ Success in Getting Bonuses

- Unsuccessful plans had an average of 9 criticisms per plan, and successful plans had 3 per plan.
- But had similar rates of complaints about metrics used and timing of release of measures.
- More from successful plans thought that some plans had unfair advantage.
Quantitative Study Approach

- QARR outcomes result from interactions of enrollees, providers and plan managers as well as market forces and state policies.
- Difference-in-differences framework: Medicaid versus commercial-only measures.
- Despite phases to the QI program, we use a simpler pre-post analysis that recognizes data constraints imposed by the small number of plans and the short time period.
QARR Measures

- **Women’s Health Care:** breast cancer screening (mammography), and postpartum care
- **Mental Health Care:** ambulatory follow-up visits within 30 days of a hospitalization; effective antidepressant medication management (for 84 or 180 days)
- **Preventive Health Care:** lead testing in children, visits to primary care physicians for children of different ages; and
- **Chronic Disease:** diabetes HbA1c testing and poor control of diabetes
Quantitative Conclusions

- QI had limited positive effects, and these were more likely among plans with a high Medicaid share.
- But Medicaid performance had not yet reached commercial performance.
- Medicaid was improving before the QI program (state studies) and may have had no place to go but up.
A Real Evaluation of P4P Would:

- Create a payer-specific control group that does not get the incentive payment
  - Possibly, from another state
- Keep the QARR/HEDIS measures defined consistently over time
- Acquire more comprehensive plan-level data on enrollees and providers
- Try P4P without other policies that could affect outcomes
  - Is this possible given market pressures?