The Design of Pay-For-Performance Programs for Reducing Disparities


Joel S. Weissman, PhD
Deputy Director/Chief Scientific Officer, Center for Surgery and Public Health
Brigham and Women’s Hospital
IDEA: Incent providers’ performance not just on overall quality, but also on *differences* in quality between majority and minority patients. But…. 

- A “rising tide lifts all boats”, so why bother to think about anything else? 
- The rich get richer: Hospitals starting out with fewer resources may not perform as well as financially robust providers. 
- Cherry-picking (and lemon-dropping): Might incent caregivers to give preference to patients who are more likely to generate improved quality scores
**Medicare Study: Most Hospitals Do Not have the Minimum # of Cases for Disparity P4P**

<table>
<thead>
<tr>
<th>Condition</th>
<th>Number of Hospitals</th>
<th>% of Total</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Hospitals w/ Any Cases</td>
<td>Hospitals w/ 30/30 Cases*</td>
</tr>
<tr>
<td>AMI</td>
<td>3999</td>
<td>1443</td>
</tr>
<tr>
<td>CHF</td>
<td>4355</td>
<td>1946</td>
</tr>
<tr>
<td>PNE</td>
<td>4425</td>
<td>2353</td>
</tr>
</tbody>
</table>

* -- at least 30 white and 30 minority cases

But... Hospitals that Met the 30/30 Threshold Treated Most of the Minority Patients

<table>
<thead>
<tr>
<th>Condition</th>
<th>Number of Minority Patients Treated…</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>In Hospitals w/ Any Cases</td>
<td>In Hospitals w/ 30/30 Cases</td>
<td>% of Total</td>
</tr>
<tr>
<td>AMI</td>
<td>92,000</td>
<td>84,000</td>
<td>91%</td>
</tr>
<tr>
<td>CHF</td>
<td>241,000</td>
<td>223,000</td>
<td>93%</td>
</tr>
<tr>
<td>PNE</td>
<td>159,000</td>
<td>152,000</td>
<td>96%</td>
</tr>
</tbody>
</table>

Our Medicare simulations showed that a P4P program using quality scores to rank performance could at most have a modest impact on disparities; a program using disparity scores to rank performance could have a larger impact on disparities.

<table>
<thead>
<tr>
<th>Change in Score After Simulation</th>
<th>Decrease in National Disparity Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rank Using Quality Scores</td>
<td>-3.7%</td>
</tr>
<tr>
<td>Rank Using Disparity Scores</td>
<td>-5.2%</td>
</tr>
</tbody>
</table>

Weissman JS, et al 2012
Hasn’t this been tried anywhere?

Yes, in Massachusetts!

What happened?
The Massachusetts Universal Health Care Law Also Mandated Hospital P4P

- Section 25 of Chapter 58 (2006)
  - MassHealth hospital rate increases contingent on quality standards, including the reduction of racial and ethnic disparities.
As with national Medicare data, care of minority patients in Massachusetts Medicaid is concentrated among a few hospitals

4 most voluminous hospitals (of 64) see 20.6% of whites and 33.2% of minorities

Blustein J, Weissman JS, Ryan AM, Doran T, Hasnain-Wynia R. Analysis raises questions on whether pay-for-performance in Medicaid can efficiently reduce racial and ethnic disparities Health Affairs 2011; 30(6):1165-1175
Differences among hospitals were small and could not be reliably estimated due to small numbers of minorities, yet bonuses and penalties were applied anyway.

Blustein J, Weissman JS, Ryan AM, Doran T, Hasnain-Wynia R. Analysis raises questions on whether pay-for-performance in Medicaid can efficiently reduce racial and ethnic disparities Health Affairs 2011; 30(6):1165-1175
Mixed Reactions from Hospital Community

- Strong stated support for the program’s goal
- Participation required extraordinary effort
- Frustration with the effort required to adapt to the clinical reporting system
- Some measures felt to be “ambiguous”
- Perceived focus on documentation at the expense of quality improvement

MASS HEALTH LAW,  MASS GENERAL

Report: State program to boost care across races is a good effort but comes up short

June 13, 2011 12:21 PM

By Chelsea Conaboy, Globe Staff
In Sum - no simple story

Despite growing interest in using P4P to reduce disparities in healthcare, we may not be quite ready yet to implement the idea in a high stakes program.

Not only do we need to know more about measures that are “disparities-sensitive”, but how to select measures that are ready to have an impact on clinical practice, have sufficient numbers of cases, how to represent differences in a statistically meaningful and policy-relevant way, and that address the “between” as well as the “within” problem.
End of Slides
OUR STUDY: P4P to reduce disparities – Design Considerations

- CMS National Hospital Quality Alliance data
  - AMI, HF, PN
  - All payer
  - CY 2005
- 2.3 million discharges from 4,450 non-federal hospitals.

Study Methods – II

Simulation Methods

• Calculate scores for each race-ethnicity within each hospital, and nationally, using:
  • Quality Scores
  • Disparity Scores

• Rank the hospitals – once using quality scores and once using disparity scores

• Simulate “success”
  • Make the bottom half look like the top
  • Re-calculate national quality scores and national disparity scores
    • Increase in overall quality scores?
    • Reduction in disparities scores?