Lessons from Medicaid Pay-for-Performance in Nursing Homes

R. Tamara Konetzka, PhD

Based on work with Rachel M. Werner, Daniel Polsky, Meghan Skira
Funded by National Institute of Aging (R01 AG034182, PI: Werner)

March 2015
Overview

• Background on nursing homes and efforts to improve quality

• Research questions
  – Does Medicaid P4P improve nursing home quality?
  – Which providers improve?
  – How does the design of P4P incentives matter?

• Lessons for policy
BACKGROUND
Nursing Home Care

- 2 distinct populations
  - Post-acute care (financed predominately by Medicare)
  - Long-term care (financed predominately by Medicaid)

- Many people, high cost
  - 1.5 million people
  - Costs $120 billion per year

- Medicaid is the majority payer
  - 50% of all expenditures for NH
  - Cover 65% of all bed-days
  - Reimburses 10-30% less than private pay rate
Persistent Concerns about Quality

• 1986 IOM report calling for major revisions in monitoring nursing home quality
• 1987 Nursing Home Reform Act (OBRA)
  • Regular inspections
  • Resident care plans
• Quality improved
• Follow up IOM report (2000)
  – Significant problems remained
• Public Reporting (Nursing Home Compare) 2002
• Medicaid Pay-for-Performance in some states
Insights from Research on Public Reporting

• Small, somewhat inconsistent improvement in nursing home quality; some evidence of “gaming”

• Heterogeneous consumer response: non-Medicaid respond more than Medicaid
  – Distance
  – Medicaid bed availability

• Which nursing homes improve depends on:
  – Type of quality measure
  – Market structure

• Need to consider costs of quality improvement
  – For different types of quality
  – For different types of providers
Early Nursing Home P4P (Norton, JHE 1990)

• 1980 experiment in San Diego
• 36 nursing homes randomized to receive financial incentives

• Three types of incentives
  – Admission (case-mix reimbursement) to improve access
  – Case outcomes (lump sum bonus for improved resident health)
  – Discharge (lump sum bonus when resident discharged home or lower-level facility)

• Results
  – Increase case mix
  – Decrease length of stay
  – Decrease in hospitalization or death
Medicare P4P Demonstration for Post-Acute Care

• Voluntary demonstration July 2009-2012
  – New York (randomized); Arizona and Wisconsin (matched controls)

• Based on performance and improvement for:
  – Staffing
  – Potentially avoidable hospitalizations
  – Survey deficiencies
  – Resident outcomes

• Financial rewards tied to Medicare payment, 80/20 shared savings; complex design

• Results: little savings, little improvement
MEDICAID P4P IN NURSING HOMES: THE LANDSCAPE
Data From State Medicaid Agencies

• Telephone survey of 50 state Medicaid agencies in 2008-2009

• In 14 states with planned or existing nursing home P4P programs, conducted in-depth interviews
  — Described P4P program features
States With Planned or Existing P4P
Werner Konetzka Liang (2010) MCRR

Existing P4P
- 3,050 nursing homes
  - 20% of nursing homes
  - 17% of residents

Planned P4P
- 2,259 nursing homes
  - 15% of nursing homes
  - 14% of residents
# Clinical Quality Measures Used

<table>
<thead>
<tr>
<th>Dates of P4P program</th>
<th>% of residents</th>
<th>Bladder catheter</th>
<th>Restraints</th>
<th>Pain</th>
<th>Falls</th>
<th>Pressure sores</th>
<th>Weight loss</th>
<th>Deficiencies</th>
<th>Staffing ratios</th>
</tr>
</thead>
<tbody>
<tr>
<td>Colorado (7/2009 to present)</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>Georgia (7/2007 to present)</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Iowa (7/2002 to present)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Kansas (7/2005 to present)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>Minnesota (10/2006 to 9/2008)</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Ohio (7/2006 to present)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Oklahoma (7/2007 to present)</td>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Utah (7/2003 to present)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>X</td>
<td></td>
</tr>
</tbody>
</table>
### Other Quality Measures Used

<table>
<thead>
<tr>
<th></th>
<th>Dates of P4P program</th>
<th>Consumer Satisfaction</th>
<th>Occupancy</th>
<th>Efficiency</th>
<th>Medicaid Use</th>
<th>Culture Change</th>
</tr>
</thead>
<tbody>
<tr>
<td>Colorado</td>
<td>(7/2009 to present)</td>
<td>X</td>
<td></td>
<td></td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Georgia</td>
<td>(7/2003 to present)</td>
<td>X</td>
<td></td>
<td></td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Iowa</td>
<td>(7/2002 to present)</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Kansas</td>
<td>(7/2005 to present)</td>
<td></td>
<td>X</td>
<td>X</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>Minnesota</td>
<td>(10/2006 to 9/2008)</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ohio</td>
<td>(7/2006 to present)</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>Oklahoma</td>
<td>(7/2007 to present)</td>
<td>X</td>
<td></td>
<td></td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Utah</td>
<td>(7/2003 to present)</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td>X</td>
</tr>
</tbody>
</table>
Tying Measures to Incentives

• Performance on each measure translates into points
  – Relative rank
  – Achieving target-level performance

• Points are summed across measures

• Translate to per diem add-on to all Medicaid resident-days
# Size of Incentives

Table 3. Summary of the Size of Financial Incentives Used in State Nursing Home Pay-for-Performance (P4P) Programs

<table>
<thead>
<tr>
<th></th>
<th>Maximum Per Diem Add-On</th>
<th>Average Per Diem Rate</th>
<th>Total Paid in P4P Bonuses (in millions)</th>
<th>P4P Bonuses as Percentage of Nursing Home Budget</th>
</tr>
</thead>
<tbody>
<tr>
<td>Colorado</td>
<td>$4.00</td>
<td>$143.75</td>
<td>___b</td>
<td>___b</td>
</tr>
<tr>
<td>Georgia</td>
<td>1.0%&lt;sup&gt;c&lt;/sup&gt;</td>
<td>$119.51</td>
<td>$5.0</td>
<td>0.4</td>
</tr>
<tr>
<td>Iowa</td>
<td>$3.68</td>
<td>$102.56</td>
<td>$6.7</td>
<td>1.4</td>
</tr>
<tr>
<td>Kansas</td>
<td>$3.00</td>
<td>$101.81</td>
<td>$2.4</td>
<td>0.7</td>
</tr>
<tr>
<td>Minnesota</td>
<td>2.4%&lt;sup&gt;c&lt;/sup&gt;</td>
<td>$137.01</td>
<td>$12.0</td>
<td>1.4</td>
</tr>
<tr>
<td>Ohio&lt;sup&gt;d&lt;/sup&gt;</td>
<td>$6.16</td>
<td>$157.00</td>
<td>$18.4</td>
<td>0.6</td>
</tr>
<tr>
<td>Oklahoma</td>
<td>$5.45</td>
<td>$96.20</td>
<td>$12.7</td>
<td>1.8</td>
</tr>
<tr>
<td>Utah</td>
<td>$0.60</td>
<td>$105.55</td>
<td>$1.0</td>
<td>0.7</td>
</tr>
<tr>
<td>Vermont</td>
<td>___e</td>
<td>$147.24</td>
<td>$0.1</td>
<td>0.1</td>
</tr>
</tbody>
</table>
MEDICAID P4P IN NURSING HOMES: DOES IT IMPROVE PERFORMANCE ON AVERAGE?
Empirical Approach

• Test for differences in nursing home performance after P4P implementation

• Difference-in-difference model
  – Pre-post in 8 nursing home states
    • Variation in timing of P4P across states
  – Use 42 control states plus DC as contemporaneous controls
Data

• Minimum Data Set (2001-2009)
  – Includes all nursing home admissions
  – Detailed clinical data collected quarterly (at least)
  – Source to calculate quality score for P4P in some states

• OSCAR
  – Survey of all certified nursing homes
  – Source of staffing and deficiency measures
  – Facility covariates

• State Medicaid agency survey (though 2009)
  – P4P implementation data
Does nursing home P4P improve nursing home performance on average?
(Werner Konetzka Polsky, HSR 2013)

• No.
  – **Improved**: restraints, pain
  – **Worsened**: catheters, falls, weight loss, deficiencies, RN+LPN staffing
  – **Neutral**: pressure sores, total staffing
Why not?

- Incentives small, potentially not noticeable
- Targeted toward NHs least able to respond
- Heterogeneity across components and facilities
MOVING BEYOND THE AVERAGE: WHICH NURSING HOMES IMPROVED?
Threshold-Based Incentives: Theory

• Threshold-based incentives differ from continuous incentives
  – The marginal benefit of improved performance is zero unless you cross the threshold
  – Effect of incentive vary non-monotonically and discontinuously around the threshold

• As a result
  – Those farthest below the threshold put forth little effort
  – Those just below put forth most effort
  – Those above the threshold put forth little effort
Our Objective

• To investigate the effect of using performance thresholds in P4P on provider response in the setting of nursing homes
  – Do low-performing providers improve their performance?
  – Do providers above the threshold improve their performance?
• Compare changes in performance related to how far NH was from threshold in prior period
Medicaid-based P4P in 2009
Werner et al (2010) MCRR

Per-diem add-on:
- Clinical quality (4)
- Staffing (8)
- Deficiencies (7)
Setting the Thresholds

NHs with clinical performance equal to or above the threshold earn points toward a P4P bonus payment

Colorado
  – Sets two pre-specified thresholds for each quality measure

Georgia
  – Uses average performance on each measure

Oklahoma
  – Uses a composite measure of performance
Results

• Little evidence of predicted strategic response to threshold incentives

  – Largest improvements in performance among nursing homes farthest below the threshold
Why no Strategic Response?

• Low cost of improving performance
  – True improvement vs. changes in coding
  – Examined clinical quality measures only

• Uncertainty of threshold

• Uncertainty of the relationship between effort and receiving the incentive
  – Complexity of NH P4P point system

• But good news that lowest-quality nursing homes improved.
MOVING BEYOND THE AVERAGE: HOW DOES PROGRAM DESIGN MATTER?
Objective

• Examine specific elements of P4P design and their effects on performance
  – Weights—do quality measures weighted more heavily see more improvement?
  – Qualifiers—do requirements for participation in P4P lead to improvement in achieving the requirement?
• Which providers seem to respond more to these incentives?
Small vs Large Weights: Clinical Quality

The bar chart compares Physically Restrained, Pressure Sores, Pain, Weight Loss, Catheter Inserted, and Falls under the smallest and largest clinical weight scenarios.
Small vs Large Weights: Staffing

![Bar chart showing comparison between smallest and largest staffing weight in total staffing and RN+LPN categories.](chart.png)
Small vs Large Weights vs Qualifier: Deficiencies

![Chart showing the comparison between Smallest Deficiency Weight, Largest Deficiency Weight, and Deficiencies as Qualifier in terms of Any Deficiencies and Any Serious Deficiencies.]
If Deficiencies Used as Qualifier: Effect on Any Deficiencies

- Chain: yes, no
- For-Profit: yes, no
- High Medicaid: yes, no

The chart shows the effect of using deficiencies as qualifiers on the outcome of Medicaid P4P in nursing homes for chain, for-profit, and high Medicaid facilities.
Deficiencies Used as Qualifier: Effect on Serious Deficiencies

<table>
<thead>
<tr>
<th>Chain</th>
<th>For-Profit</th>
<th>High Medicaid</th>
</tr>
</thead>
<tbody>
<tr>
<td>yes</td>
<td>yes</td>
<td>yes</td>
</tr>
<tr>
<td>no</td>
<td>no</td>
<td>no</td>
</tr>
</tbody>
</table>

![Graph showing the effect of different qualifications on serious deficiencies.](image)
Conclusions

• Use of weights in bonus formulae had mixed effects
  – Larger weights were only sometimes associated with more improvement
  – Smaller weights sometimes associated with worsening
• Simple requirement for participation – no deficiencies – was more effective
• Well-resourced nursing homes more likely to improve on average
• But we see important improvement among nursing homes considered lowest quality
Lessons Learned from Medicaid P4P in Nursing Homes

• Structure of P4P incentives matters
  – Using weights may have unintended consequences
  – Simple rules for participation may incent larger improvement

• Context (market, competing incentives) matters

• Heterogeneity is key
  – Looking for average effects of a multi-faceted P4P program may not be fruitful
  – “Personalized” P4P – facility-specific targets – may be more effective than one-size-fits-all