Two-Year Impact of the Alternative Quality Contract on Pediatric Health Care Quality and Spending

WHAT'S KNOWN ON THIS SUBJECT: Payment arrangements that blend global budgets with pay-for-performance are proliferating. However, little is known about how these contracts affect pediatric health care quality and spending for children with and without special health care needs receiving care from large provider organizations.

WHAT THIS STUDY ADDS: A prototypical global budget contract significantly improved preventive care quality measures tied to pay-for-performance, especially for children with special health care needs. It did not alter trends for spending or for quality measures that were not tied to pay-for-performance.

abstract

**OBJECTIVE:** To examine the 2-year effect of Blue Cross Blue Shield of Massachusetts' global budget arrangement, the Alternative Quality Contract (AQC), on pediatric quality and spending for children with special health care needs (CSHCN) and non-CSHCN.

**METHODS:** Using a difference-in-differences approach, we compared quality and spending trends for 126,975 unique 0- to 21-year-olds receiving care from AQC groups with 415,331 propensity-matched patients receiving care from non-AQC groups; 23% of enrollees were CSHCN. We compared quality and spending pre (2006–2008) and post (2009–2010) AQC implementation, adjusting analyses for age, gender, health risk score, and secular trends. Pediatric outcome measures included 4 preventive and 2 acute care measures tied to pay-for-performance (P4P), 3 asthma and 2 attention-deficit/hyperactivity disorder quality measures not tied to P4P, and average total annual medical spending.

**RESULTS:** During the first 2 years of the AQC, pediatric care quality tied to P4P increased by +1.8% for CSHCN (P < .001) and +1.2% for non-CSHCN (P < .001) for AQC versus non-AQC groups; quality measures not tied to P4P showed no significant changes. Average total annual medical spending was ~5 times greater for CSHCN than non-CSHCN; there was no significant impact of the AQC on spending trends for children.

**CONCLUSIONS:** During the first 2 years of the contract, the AQC had a small but significant positive effect on pediatric preventive care quality tied to P4P; this effect was greater for CSHCN than non-CSHCN. However, it did not significantly influence (positively or negatively) CSHCN measures not tied to P4P or affect per capita spending for either group. *Pediatrics* 2014;133:96–104

**AUTHORS:** Alyna T. Chien, MD, MS,a,b Zirui Song, PhD,c Michael E. Chernew, PhD,c Bruce E. Landon, MD, MBA, MSc,d,e Barbara J. McNeil, MD, PhD,c,d Dana G. Safran, ScD,f,g and Mark A. Schuster, MD, PhDh,i

aDivision of General Pediatrics, Department of Medicine, Boston Children’s Hospital, Boston, Massachusetts; Departments of Pediatrics and Health Care Policy, Harvard Medical School, Boston, Massachusetts; bDivision of General Medicine and Primary Care, Department of Medicine, Beth Israel Deaconess Medical Center, Boston, Massachusetts; cDepartment of Radiology, Brigham and Women’s Hospital, Boston, Massachusetts; cBlue Cross Blue Shield of Massachusetts, Boston, Massachusetts; and cDepartment of Medicine, Tufts University School of Medicine, Boston, Massachusetts

**KEY WORDS**

pay-for-performance, capitation, payment, pediatrics, children with special health care needs, preventive care, acute care quality, chronic disease care quality, asthma, ADHD

**ABBREVIATIONS**

ADHD—attention-deficit/hyperactivity disorder
AQC—alternative quality contract
BCBSMA—Blue Cross Blue Shield of Massachusetts
CSHCN—children with special health care needs
P4P—pay-for-performance
PCP—primary care physician

Dr Chien conceptualized and designed the study, analyzed and interpreted the data, drafted the initial manuscript, critically revised the manuscript, performed the statistical analysis, and obtained funding; Dr Song conceptualized and designed the study, acquired data, analyzed and interpreted the data, critically revised the manuscript, and performed the statistical analysis; Dr Chernew conceptualized and designed the study, acquired data, analyzed and interpreted the data, critically revised the manuscript, performed the statistical analysis, obtained funding, and supervised the study; Dr Landon conceptualized and designed the study, acquired data, analyzed and interpreted the data, critically revised the manuscript, performed the statistical analysis, and obtained funding; and Dr Safran conceptualized and designed the study, acquired data, analyzed and interpreted the data, critically revised the manuscript, performed the statistical analysis, provided administrative, technical, or material support; Dr Schuster conceptualized and designed the study, acquired data, analyzed and interpreted the data, critically revised the manuscript, performed the statistical analysis, obtained funding, provided administrative, technical, or material support, and supervised the study.

(Continued on last page)
There is great interest in understanding whether combining incentives to reduce spending (via global budgets or risk) with those to improve quality (via pay-for-performance [P4P]) will be effective at slowing spending growth while improving or maintaining quality. The Centers for Medicare & Medicaid Services use this payment approach for accountable care organizations, and many public and private payers are embracing this payment model. In 2009, Blue Cross Blue Shield of Massachusetts (BCBSMA) established the Alternative Quality Contract (AQC), a prototypical global budget arrangement in which provider groups share in both upside (savings) and downside financial risk while being eligible to earn bonuses for improving care quality. BCBSMA initially engaged 7 provider groups to join 5-year contracts; 4 more groups joined in 2010. General results from the contract, focused mainly on adult care, showed significant but modest reductions in spending and small improvements in quality that began in Year 1 (2009) and continued to grow in Year 2 (2010).

In the commercial market, global budget arrangements such as the AQC often include children, but because of the size and spending levels of the adult population, the effects of such arrangements on the pediatric population are often an afterthought. It is important to understand the early impact of the AQC on pediatric health care to inform current efforts to adopt and spread global budget strategies. The 2010 Affordable Care Act promotes the use of global capitated payment models through Medicaid demonstration projects. Private health plans, which are a critical source of insurance for children generally and for children with special health care needs (CSHCN) in particular, are increasingly engaging providers in global contracts. Children’s hospitals are entering contracts such as the AQC, and pediatric providers are joining delivery systems in which global budgets are an expected form of payment.

Of particular concern is care related to CSHCN, children with common chronic health conditions (eg, asthma, attention-deficit/hyperactivity disorder [ADHD]), acute life-threatening illnesses that generally necessitate time-limited intensive treatment (eg, congenital heart disease, cancer), and conditions that affect the life course (eg, Down syndrome, sickle cell disease, cystic fibrosis, HIV). CSHCN make up 13% to 26% of the pediatric population. Poor-quality health care for CSHCN is thought to contribute to excess morbidity (eg, ~2 million annual unnecessary pediatric asthma exacerbations) and poor academic achievement (eg, inadequate ADHD management). Moreover, parents of CSHCN who receive fragmented or poor-quality health care more often report having to reduce their work hours or leave the workforce altogether. Health care spending for CSHCN can also be 2.5 to 20 times greater than for other children and accounts for 45% to 75% of total public and private pediatric health insurance expenditures.

Despite these concerns, there is a paucity of information about how payment incentives can promote care quality or spending for CSHCN. To our knowledge, there is no published information on the intended and unintended effects of a global budget arrangement in a general pediatric population or in CSHCN. Thus, this study aims to examine whether BCBSMA’s AQC was effective in its first 2 years at improving care quality for CSHCN and non-CSHCN across measures that are and are not tied to P4P and at slowing pediatric spending growth for CSHCN and non-CSHCN.

**METHODS**

**Setting, Study Design, and Participants**

We studied a commercially insured pediatric population enrolled in BCBSMAs health maintenance organization plan. BCBSMA is the largest private insurer in Massachusetts, with a commercial market share of ~50%.

We used a difference-in-differences study design to examine the AQC’s effect on pediatric quality and spending at the enrollee level for the AQC first and second year. This study design allows us to account for unmeasured differences between case and comparison groups in a manner that enables estimation of treatment effects in a nonrandomized setting.

Our cases were 0- to 21-year-old enrollees who designated a primary care physician (PCP) within a provider group that joined the AQC; our comparisons were 0- to 21-year-old enrollees who designated PCPs within a provider group that had not joined the AQC (ie, continued fee-for-service contracts that bore no financial risk and were only tied to up to 5 pediatric P4P measures for which related year-end bonuses were approximately one-fifth that available in the AQC). Seven provider groups joined the AQC in 2008 and were included in the 2-year analysis. For each year, our comparisons were propensity score–matched enrollees using age, gender, and health risk score (an International Classification of Diseases, Ninth Revision-based method for estimating the medical complexity of patients). We compared quality and spending from the 3 years preceding the AQC (2006–2008) with each of the 2 years after the AQC began (2009 and 2010).

For each year, we limited our study to enrollees who were continuously enrolled for at least 1 full calendar year. The only exception was for infants in their birth year (eg, an infant born August 1 and continuously enrolled thereafter...
contributed 5 months of data for that year.

The 7 groups engaged in the AQC were large; the total number of physicians per group ranged from 636 to 6786. On average, pediatric physicians (pediatricians and family practitioners) constituted 6% (SD 4%) and 4% (SD 1%) of the total PCP workforce within these groups, respectively. Children accounted for 26% (SD 12%) of the total enrollees in these groups. AQC and non-AQC groups usually included dozens to scores of small, medium, and large single-specialty practices, many of which had affiliations with a local independent practice association or other physician contracting organization. In both arms, large multispecialty practices were only a small proportion of the total number of practices. Groups varied in degree of integration with the total number of physicians per group ranging from 2 to 200, although pediatricians and family practitioners constituted a similar proportion of the total PCP workforce in these non-AQC groups.

Harvard Medical School’s and Boston Children’s Hospital’s institutional review boards approved this study.

Variables of Interest

Quality

We assessed pediatric care quality by using 11 claim-based measures endorsed by the Healthcare Effectiveness Data and Information Set or the Agency for Healthcare Research and Quality (Table 1). We examined the 6 dichotomous measures that the AQC tied to P4P rewards as a composite with each measure weighted equally; we individually analyzed 5 additional measures that were not tied to P4P bonuses in the AQC contract. We followed Healthcare Effectiveness Data and Information Set specifications for all these quality measures. P4P measures covered preventive and acute care; measures not tied to P4P pertained to CSHCN-related care, specifically care for the 2 most common chronic conditions affecting the physical and behavioral health of children, asthma and ADHD, respectively. For CSHCN-related quality measures, we limited the analysis sample to the 80% of the study population with full prescription drug benefits through BCBSMA because these measures rely on prescription drug information to ascertain care quality.

Spending

We assessed spending by calculating each enrollee’s total annual medical spending. We included all medical claims regardless of whether services

<table>
<thead>
<tr>
<th>Measure Name</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Well visits: infants</td>
<td>Whether enrollees turning 15 mo of age during the measurement year had at least 6 well visits in the first 15 mo of life&lt;sup&gt;a,b&lt;/sup&gt;</td>
</tr>
<tr>
<td>Well visits: children</td>
<td>Whether enrollees turning 3–6 y of age during the measurement year had at least 1 well visit in the measurement year&lt;sup&gt;a,b&lt;/sup&gt;</td>
</tr>
<tr>
<td>Well visits: adolescents</td>
<td>Whether enrollees turning 12–21 y of age during the measurement year had at least 1 well visit in the measurement year&lt;sup&gt;a,b&lt;/sup&gt;</td>
</tr>
<tr>
<td>Chlamydia screening</td>
<td>Whether sexually active female enrollees aged 16–21 y had at least 1 test for chlamydia during the measurement year&lt;sup&gt;a,b&lt;/sup&gt;</td>
</tr>
<tr>
<td>Acute care</td>
<td>Whether enrollees aged 2–18 y who were diagnosed with pharyngitis and were dispensed an antibiotic received a group A streptococcal testing&lt;sup&gt;a,b&lt;/sup&gt;</td>
</tr>
<tr>
<td>Upper respiratory infection treatment</td>
<td>Whether enrollees aged 5 mo–18 y who were given a diagnosis of a upper respiratory infection were not dispensed an antibiotic prescription&lt;sup&gt;a,b&lt;/sup&gt;</td>
</tr>
<tr>
<td>Quality Measures Not Tied to P4P Persistent asthma</td>
<td>Percentage of enrollees aged 5–21 y with persistent asthma who had ≥1 emergency department visit for which asthma was the primary diagnosis during the measurement year&lt;sup&gt;a&lt;/sup&gt;</td>
</tr>
<tr>
<td>Appropriate medications</td>
<td>Percentage of enrollees aged 5–21 y with persistent asthma who were dispensed controller medications (eg, inhaled corticosteroids, leukotriene modifiers) during the measurement year&lt;sup&gt;a&lt;/sup&gt;</td>
</tr>
<tr>
<td>Medication management</td>
<td>Number of days out of the year that enrollees aged 5–21 y with persistent asthma were dispensed controller medications (eg, inhaled corticosteroids, leukotriene modifiers)&lt;sup&gt;b&lt;/sup&gt;</td>
</tr>
<tr>
<td>ADHD Follow-up: initiation</td>
<td>Whether enrollees aged 6–12 y who were dispensed a stimulant medication had at least 1 follow-up visit to a practitioner with prescribing authority during the 30 d initiation phase&lt;sup&gt;a,b&lt;/sup&gt;</td>
</tr>
<tr>
<td>Follow-up: maintenance</td>
<td>Whether children aged 6–12 y who were dispensed a stimulant medication and remained on the medication for at least 210 d had ≥2 follow-up visits to a practitioner with prescribing authority within 270 d after the initiation phase&lt;sup&gt;a,b&lt;/sup&gt;</td>
</tr>
</tbody>
</table>

<sup>a</sup> Agency for Healthcare Research and Quality<sup>42</sup>

<sup>b</sup> Healthcare Effectiveness Data and Information Set 2011<sup>43</sup>
were delivered by the enrollees’ own PCP, another physician, or a physician in a different organization entirely. We incorporated claims paid by BCBSMA at its negotiated fee-for-service prices plus expenses incurred by enrollees or families due to cost sharing. For spending analyses, we excluded prescription drug spending because not all enrollees had prescription drug coverage through BCBSMA.

### Statistical Analysis

For all analyses, after using propensity score methods to match AQC enrollees to non-AQC enrollees (based on age, gender, and health risk score), we used a multivariate linear model to examine spending and care quality in AQC versus non-AQC groups. We controlled for secular trends using indicators for time (year for the quality model, quarters for the spending model) and interactions between AQC and time, treated plan type (year for the quality model, quarters for the spending model) and interactions, and adjusted for clustering of multiple observations for each person. The interaction between indicators for the AQC and the first and second years of the contract (2009 and 2010) provided our estimates of interest. For the spending analysis, we did not log-transform our models because studies show that linear models perform better than more complex functional forms and because the health risk score is designed to predict dollar spending. We used Stata version 11 (Stata Corp, College Station, TX) for all analyses and report all results by using 2-tailed P values.

### Results

#### Statistical Analysis

For all analyses, after using propensity score methods to match AQC enrollees to non-AQC enrollees (based on age, gender, and health risk score), we used a multivariate linear model to examine spending and care quality in AQC versus non-AQC groups. We controlled for secular trends using indicators for time (year for the quality model, quarters for the spending model) and interactions between AQC and time, treated plan type as fixed effects, and adjusted for clustering at the practice level. We also included enrollee age, gender, interaction between age and gender, and health risk score in all models to adjust for residual differences between enrollees not accounted for by the propensity score matching methods. The health risk score was assigned according to the risk adjustment system BCBSMA uses for its AQC program, DxCG. We used DxCG’s risk adjustment method because we wanted to simulate the AQC program in our analysis, and its risk adjustment approach is standard for the field. It has also been developed with pediatrician input and is used nationally and internationally. DxCG uses patient demographics, current-year diagnostic codes, current-year spending, and a large national database to generate a health risk score between 0 and 66 for patients (higher scores indicate greater medical complexity and predicted costs). This particular model was designed for persons <65 years old. We used Huber–White corrections to adjust SEs for clustering of multiple observations for each person. The interaction between indicators for the AQC and the first and second years of the contract (2009 and 2010) provided our estimates of interest. For the spending analysis, we did not log-transform our models because studies show that linear models perform better than more complex functional forms and because the health risk score is designed to predict dollar spending. We used Stata version 11 (Stata Corp, College Station, TX) for all analyses and report all results by using 2-tailed P values.

### Quality

Baseline (2006–2008) performance on quality measures was comparable for AQC and non-AQC groups except for the pediatric acute care quality measures (appropriate testing for pharyngitis, upper respiratory infection treatment); these began at ∼93% in AQC groups, which was 10 percentage points higher than in non-AQC groups (Table 3).

### Table 2: Demographic Characteristics of BCBSMA Enrollees Aged 0–21 Years Pre and Post Start of the AQC

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>AQC</th>
<th>Non-AQC</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pre, N (%)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Post, N (%)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Unique enrollees</td>
<td>126 975</td>
<td>415 331</td>
</tr>
<tr>
<td>Age category, y</td>
<td></td>
<td></td>
</tr>
<tr>
<td>0–&lt;2</td>
<td>65 878</td>
<td>67 430</td>
</tr>
<tr>
<td>2–&lt;7</td>
<td>6152  (9)</td>
<td>6134  (9)</td>
</tr>
<tr>
<td>7–&lt;13</td>
<td>14 630 (22)</td>
<td>14 309 (21)</td>
</tr>
<tr>
<td>13–&lt;21</td>
<td>18 074 (27)</td>
<td>17 969 (27)</td>
</tr>
<tr>
<td>Female</td>
<td>32 022 (41)</td>
<td>29 018 (43)</td>
</tr>
<tr>
<td>Health risk score</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean</td>
<td>0.523</td>
<td>0.588</td>
</tr>
<tr>
<td>Interquartile range</td>
<td>0.067–0.595</td>
<td>0.070–0.662</td>
</tr>
</tbody>
</table>

Pre = 2006; Post = 2008.

* Per proprietary risk adjustment software used by BCBSMA, DxCG.
Preintervention performance on the remaining P4P measures was comparable for both CSHCN and non-CSHCN and bore greater room for improvement. For instance, 74% of adolescent patients received the recommended number of well visits, and 55% received chlamydia testing. Across the composite of 6 P4P measures, pediatric care quality increased significantly for AQC versus non-AQC groups when we compared 2009 (Year 1) and 2010 (Year 2) with 2006–2008 (the baseline period). Individually, changes for Year 1 and Year 2 were both significant, but the magnitude of those changes was almost uniformly greater for Year 2 than Year 1. For acute care measures, care quality improved more in non-AQC than AQC groups (for whom performance was already very high).

Care quality increased more for CSHCN than for non-CSHCN on measures focused on preventive care (well visit rates for all age categories, chlamydia screening). Analogously, care quality for acute care measures tended to increase in the control group more for CSHCN than for non-CSHCN.

For quality measures pertaining to asthma and ADHD care that were not tied to P4P, baseline performance was comparable across AQC and non-AQC groups, with the exception of emergency department visit rates, for which beginning rates differed by 12.5 percentage points between the groups (Table 3). We did not find a statistically significant difference in performance on quality measures not tied to P4P between AQC and non-AQC groups.

**Spending**

Average spending levels were highest for CSHCN (Table 4). At baseline, average per capita spending levels for CSHCN in AQC ($4585) and non-AQC ($4676) groups were nearly 5 times as great as those for non-CSHCN ($996 in AQC, $987 in non-AQC). Health care spending for all 0- to 21-year-old enrollees cared for by AQC groups, including both CSHCN and non-CSHCN, rose by $12 ($P = .33) more per child in Year 1 and $131 ($P = .11) more per child in Year 2 compared with non-AQC groups. These spending patterns were the same across all the subpopulations of children examined.

### Table 3 2009 Cohort: Adjusted Difference-in-Differences in Pediatric Care Quality Pre and Post Start of the AQC

<table>
<thead>
<tr>
<th>Measure Name</th>
<th>Measure of Pediatric Quality</th>
<th>AQC</th>
<th>Non-AQC</th>
<th>Difference-in-Differencea</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Quality Measures Tied to P4P</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Aggregate</td>
<td>CSHCN</td>
<td>80.2</td>
<td>83.4</td>
<td>0.6</td>
</tr>
<tr>
<td></td>
<td>Non-CSHCN</td>
<td>79.3</td>
<td>82.6</td>
<td>1.3</td>
</tr>
<tr>
<td>Prevention and screening</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Well visits: infants</td>
<td>CSHCN</td>
<td>92.0</td>
<td>95.1</td>
<td>2.2</td>
</tr>
<tr>
<td></td>
<td>Non-CSHCN</td>
<td>92.6</td>
<td>94.6</td>
<td>1.4</td>
</tr>
<tr>
<td>Well visits: children</td>
<td>CSHCN</td>
<td>93.6</td>
<td>95.7</td>
<td>2.1</td>
</tr>
<tr>
<td></td>
<td>Non-CSHCN</td>
<td>92.3</td>
<td>94.7</td>
<td>1.7</td>
</tr>
<tr>
<td>Well visits: adolescents</td>
<td>CSHCN</td>
<td>78.5</td>
<td>82.3</td>
<td>3.8</td>
</tr>
<tr>
<td></td>
<td>Non-CSHCN</td>
<td>73.8</td>
<td>78.0</td>
<td>4.0</td>
</tr>
<tr>
<td>Chlamydia screening</td>
<td>CSHCN</td>
<td>53.5</td>
<td>65.7</td>
<td>4.6</td>
</tr>
<tr>
<td></td>
<td>Non-CSHCN</td>
<td>55.4</td>
<td>66.2</td>
<td>6.9</td>
</tr>
<tr>
<td>Acute care</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pharyngitis testing</td>
<td>CSHCN</td>
<td>93.9</td>
<td>95.0</td>
<td>1.1</td>
</tr>
<tr>
<td></td>
<td>Non-CSHCN</td>
<td>93.7</td>
<td>96.1</td>
<td>2.4</td>
</tr>
<tr>
<td>Upper respiratory infection treatment</td>
<td>CSHCN</td>
<td>94.8</td>
<td>94.5</td>
<td>0.3</td>
</tr>
<tr>
<td></td>
<td>Non-CSHCN</td>
<td>94.7</td>
<td>95.7</td>
<td>1.0</td>
</tr>
<tr>
<td><strong>Quality Measures Not Tied to P4P</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Persistent asthma</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Emergency department visits</td>
<td></td>
<td>18.5</td>
<td>17.4</td>
<td>0.8</td>
</tr>
<tr>
<td>Appropriate medications</td>
<td></td>
<td>7.9</td>
<td>6.2</td>
<td>1.7</td>
</tr>
<tr>
<td>Medication management</td>
<td></td>
<td>174</td>
<td>159</td>
<td>1.5</td>
</tr>
<tr>
<td>ADHD</td>
<td></td>
<td>39.3</td>
<td>46.8</td>
<td>7.5</td>
</tr>
<tr>
<td>Follow-up: initiation</td>
<td></td>
<td>38.7</td>
<td>51.0</td>
<td>10.3</td>
</tr>
<tr>
<td>Follow-up: maintenance</td>
<td></td>
<td>41.3</td>
<td>48.7</td>
<td>7.4</td>
</tr>
</tbody>
</table>

Pre, adjusted mean 2006–2008; Post, adjusted mean 2009–2010; Year 1, 2009; Year 2, 2010. Difference-in-differences figures are adjusted for patient age, gender, age × gender interaction, health risk score, and time trend.

a Difference-in-differences = (AQCPost - AQCPre) - (Non-AQCPost - Non-AQCPre).
By CSHCN status

Two results stand in contrast to adult-effect on spending for children.

This preventive care measures tied to P4P.

AQC did not affect measures that were not those preventive measures. However, the AQC did not affect measures that were not tied to P4P and did not have a significant effect on spending for children.

Two results stand in contrast to adult-oriented findings. First, AQC groups may have reached a ceiling with respect to pediatric acute care measures; baseline pediatric performance was ≥94% in AQC groups, whereas non-AQC groups’ baseline performance was 2.5% to 11.8% lower. Difficulty raising quality further when it is already high has been observed in other studies of quality improvement and P4P. Second, pediatric spending trends did not change significantly, whereas aggregate spending for adults slowed modestly but significantly over the study period. Although improving care quality for CSHCN was not a specific focus of the AQC, this clinically important group benefited the most with respect to preventive care measures tied to P4P. This finding is consistent with a pediatric

P4P study that also found that children with chronic conditions had significantly greater odds of being immunized, and it supports the notion that it may be easier to achieve targeted performance rates in a population that already has higher health care utilization rates.

The AQC was not associated (positively or negatively) with changes in pediatric care quality not tied to P4P, at least not for 2 dominant groups of CSHCN (those with asthma and ADHD). The AQC could have had negative unintended consequences if it caused providers to overemphasize care tied to P4P over important care not tied to quality incentives or caused providers to stint on needed care viewed as costly. For example, following recommended care guidelines for ADHD could cause care costs to rise (through more medication management visits or referrals) without providing short-term savings as a counterbalance (eg, fewer emergency department visits for asthma). The lack of a negative spillover effect is reassuring; however, it also suggests that there was no positive spillover when there are potential savings to be gained, as is likely in the case of asthma.

The lack of significant change in pediatric spending could result from pediatric efforts being either modest or deferred compared with those for adults (eg, if groups viewed the opportunities for pediatric savings to be small relative to that available for adults). Although pediatric spending on CSHCN is 4 to 5 times greater than that on non-CSHCN, CSHCN are approximately one-fifth of all children, and average per capita pediatric spending is one-third lower than adult spending. Consequently, efforts to achieve savings by addressing pediatric care may have seemed more difficult or less pressing in the initial years of the contract. Alternatively, groups may have lacked experience with the methods that underlie pediatric performance measurement and improvement, because standard adult conventions may need to be modified to meaningfully assess and improve quality in pediatrics. Additionally, opportunities to reduce pediatric spending levels probably require a deliberate focus on care delivery for CSHCN, which means groups must build infrastructure for tracking care and spending for a clinically heterogeneous population to realize savings.

Our study has limitations. A number of factors may make the case and comparison groups different (eg, availability of quality improvement infrastructures, previous P4P exposure, or bonus amounts) and influence our difference-in-differences estimates. However, the fact that preintervention trends were similar between AQC and non-AQC groups reassures us that such factors did not play a large role in determining our results. Although we are using standard pediatric quality measures, future studies are likely to benefit from the new pediatric quality measures currently being developed and from greater investigation into the scale or scope of infrastructures (eg, quality improvement or health information technology) needed for groups to respond to quality

**DISCUSSION**

In the first 2 years of its implementation, we found that the AQC had on average a significant, positive, and small effect on pediatric preventive care quality measures tied to P4P and that CSHCN experienced significantly greater increases in performance than non-CSHCN on those preventive measures. However, the AQC did not affect measures that were not tied to P4P and did not have a significant effect on spending for children.

Two results stand in contrast to adult-oriented findings. First, AQC groups may have reached a ceiling with respect to pediatric acute care measures; baseline pediatric performance was ≥94% in AQC groups, whereas non-AQC groups’ baseline performance was 2.5% to 11.8% lower. Difficulty raising quality further when it is already high has been observed in other studies of quality improvement and P4P. Second, pediatric spending trends did not change significantly, whereas aggregate spending for adults slowed modestly but significantly over the study period. Although improving care quality for CSHCN was not a specific focus of the AQC, this clinically important group benefited the most with respect to preventive care measures tied to P4P. This finding is consistent with a pediatric

P4P study that also found that children with chronic conditions had significantly greater odds of being immunized, and it supports the notion that it may be easier to achieve targeted performance rates in a population that already has higher health care utilization rates.

The AQC was not associated (positively or negatively) with changes in pediatric care quality not tied to P4P, at least not for 2 dominant groups of CSHCN (those with asthma and ADHD). The AQC could have had negative unintended consequences if it caused providers to overemphasize care tied to P4P over important care not tied to quality incentives or caused providers to stint on needed care viewed as costly. For example, following recommended care guidelines for ADHD could cause care costs to rise (through more medication management visits or referrals) without providing short-term savings as a counterbalance (eg, fewer emergency department visits for asthma). The lack of a negative spillover effect is reassuring; however, it also suggests that there was no positive spillover when there are potential savings to be gained, as is likely in the case of asthma.

The lack of significant change in pediatric spending could result from pediatric efforts being either modest or deferred compared with those for adults (eg, if groups viewed the opportunities for pediatric savings to be small relative to that available for adults). Although pediatric spending on CSHCN is 4 to 5 times greater than that on non-CSHCN, CSHCN are approximately one-fifth of all children, and average per capita pediatric spending is one-third lower than adult spending. Consequently, efforts to achieve savings by addressing pediatric care may have seemed more difficult or less pressing in the initial years of the contract. Alternatively, groups may have lacked experience with the methods that underlie pediatric performance measurement and improvement, because standard adult conventions may need to be modified to meaningfully assess and improve quality in pediatrics. Additionally, opportunities to reduce pediatric spending levels probably require a deliberate focus on care delivery for CSHCN, which means groups must build infrastructure for tracking care and spending for a clinically heterogeneous population to realize savings.

Our study has limitations. A number of factors may make the case and comparison groups different (eg, availability of quality improvement infrastructures, previous P4P exposure, or bonus amounts) and influence our difference-in-differences estimates. However, the fact that preintervention trends were similar between AQC and non-AQC groups reassures us that such factors did not play a large role in determining our results. Although we are using standard pediatric quality measures, future studies are likely to benefit from the new pediatric quality measures currently being developed and from greater investigation into the scale or scope of infrastructures (eg, quality improvement or health information technology) needed for groups to respond to quality.
and spending incentives within a 2-year time frame. Last, it may appear that the P4P component of the AQC may have had a stronger effect than the risk-bearing capitation element. However, the possibility that spending reduction incentives could affect quality efforts remains (especially for quality measures aimed at limiting overuse of inappropriate services [eg, reducing antibiotic prescriptions for upper respiratory infections or decreasing emergency department visits for children with asthma]), and our study was not designed to distinguish the relative effects of the P4P and spending reduction components of the AQC.

This study adds to 2 distinct areas in the pediatric payment literature. First, it is the first to evaluate how large provider groups, which tend to care for adults more than children, perform with respect to pediatric care when responding to a global budget intervention. Available studies, derived mainly from the pediatric P4P literature, have been conducted in environments in which children are the main focus of a practice (eg, academic pediatric practices) or constitute a majority population (eg, a Medicaid health plan). Second, to our knowledge, this is the first empirical examination of the effectiveness of risk-sharing global budget incentives for a pediatric population. Previous literature on capitation in pediatrics has been largely descriptive. Stakeholders should know how the effects of general payment policies may differ depending on whether children are a minority or the main focus of providers.

### REFERENCES


9. Chien AT, Colman MW, Ross LF. Qualitative insights into how pediatric pay-for-performance programs are being designed. Acad Pediatr. 2009;9(3):185–191


18. Bethell CD, Read D, Blumberg SJ, Newacheck PW. What is the prevalence of children with special health care needs? Toward an understanding of variations in findings and


38. McCarthy D. Case study: the Mount Auburn Ambulatory Care Finance and Policy; 2010


Two-Year Impact of the Alternative Quality Contract on Pediatric Health Care Quality and Spending
Alyna T. Chien, Zirui Song, Michael E. Chernew, Bruce E. Landon, Barbara J. McNeil, Dana G. Safran and Mark A. Schuster

*Pediatrics* 2014;133;96; originally published online December 23, 2013; DOI: 10.1542/peds.2012-3440

Updated Information & Services
including high resolution figures, can be found at:
/content/133/1/96.full.html

References
This article cites 53 articles, 11 of which can be accessed free at:
/content/133/1/96.full.html#ref-list-1

Citations
This article has been cited by 7 HighWire-hosted articles:
/content/133/1/96.full.html#related-urls

Subspecialty Collections
This article, along with others on similar topics, appears in the following collection(s):
Administration/Practice Management
/cgi/collection/administration/practice_management_sub
Practice-Based Learning & Development
/cgi/collection/practice-based_learning_-_development_sub
Child Care
/cgi/collection/child_care_sub

Permissions & Licensing
Information about reproducing this article in parts (figures, tables) or in its entirety can be found online at:
/site/misc/Permissions.xhtml

Reprints
Information about ordering reprints can be found online:
/site/misc/reprints.xhtml
Two-Year Impact of the Alternative Quality Contract on Pediatric Health Care Quality and Spending
Alyna T. Chien, Zirui Song, Michael E. Chernew, Bruce E. Landon, Barbara J. McNeil, Dana G. Safran and Mark A. Schuster
Pediatrics 2014;133;96; originally published online December 23, 2013; DOI: 10.1542/peds.2012-3440

The online version of this article, along with updated information and services, is located on the World Wide Web at:
/content/133/1/96.full.html