

Contract Final Report

National Evaluation of the CHIPRA Quality Demonstration Grant Program: Final Project Report

Prepared for:

Agency for Healthcare Research and Quality
Rockville, MD
Contract HHS29020090002191

Prepared by:

Mathematica Policy Research
Washington, DC

Authors:

Mathematica Policy Research

Henry Ireys, PhD, Project Director
Joseph Zickafoose, MD, MS
Dana Petersen, PhD
Anna Christensen, PhD

Urban Institute

Rachel Burton MPP
Kelly Devers

AHRQ Publication No. 16-0007-EF
November 2015

This contract final report was prepared for the Agency for Healthcare Research and Quality by Mathematica Policy Research under contract HHS29020090002191. The statements and opinions presented herein are those of the authors and should not be construed as the official position of the Agency for Healthcare Research and Quality, the Centers for Medicare & Medicaid Services, or the U.S. Department of Health and Human Services.

This report is in the public domain, and it may be used and reprinted without restriction. Citation as to source will be appreciated.

Suggested citation:

Ireys H, Zickafoose J, Petersen D, et al. National Evaluation of the CHIPRA Quality Demonstration Grant Program: Final Project Report. AHRQ Publication No. 16-0007-EF. Rockville, MD: Agency for Healthcare Research and Quality; November 2015.

Contents

1. Overview

CHIPRA Quality Demonstration Grant Program	1
Evaluation of the Demonstration Grant Program	2
Final Report	3

2. Synthesis of Key Findings by Category

Category A Findings	4
Category B Findings	9
Category C Findings	12
Category D Findings	18
Category E Findings	19
Cross-Cutting Findings	22

3. Observations About the Structure of the Demonstration Grant Program

Allocation of Grant Program Resources	24
Multistate Partnerships	26
Payment and Other Approaches to Sustainability	27
Grant Administration and Planning	28
Grant Structure and Rigorous Evaluation	28

4. Observations About the Evaluation

The Challenge of Impact Analyses	30
Collaboration with Grantees	33
Technical Expert Panel	34
Development and Dissemination of Evaluation Findings	34

5. Conclusion

Conclusion and Summary	35
------------------------	----

6. Endnotes

References and Notes	36
----------------------	----

Appendixes

Appendix A. National Evaluation Team	39
Appendix B. Products Produced or Initiated by the National Evaluation Team	40
Appendix C. Obstacles to the Impact Analyses of Category C Projects	42

1. Overview

CHIPRA Quality Demonstration Grant Program

In February 2010, the Centers for Medicare and Medicaid Services (CMS) awarded 10 grants, funding 18 States, to improve the quality of health care for children enrolled in Medicaid and the Children's Health Insurance Program (CHIP). Funded by the Children's Health Insurance Program Reauthorization Act of 2009 (CHIPRA), the Quality Demonstration Grant Program aimed to identify effective, replicable strategies for enhancing the quality of health care for children.

Through this program, 18 demonstration States implemented 52 projects in five categories (Table 1):

- **Category A:** Grantees enhanced their capacity to report and use the CMS Child Core Set of quality measures and other supplemental quality measures for children.
- **Category B:** Grantees developed or enhanced health information technology (IT) to improve quality of care, reduce costs, and increase transparency. Grantees pursued a range of health IT solutions, such as encouraging uptake of electronic health records (EHRs), developing a regional health information exchange, and interfacing electronic health information with eligibility systems or social service organizations.
- **Category C:** Grantees developed or expanded provider-based care models. These models include (1) the patient-centered medical home (PCMH); (2) care management entities (CMEs), which aim to improve services for children and youth with serious emotional disorders; and (3) school-based health centers (SBHCs).
- **Category D:** Grantees implemented and evaluated the impact of a model EHR format for children, which was developed under a separate Agency for Healthcare Research and Quality (AHRQ) contract, in partnership with CMS.
- **Category E:** In addition to working in at least one of the other categories, grantees proposed additional activities. These activities were intended to enhance their work under another category or focus on an additional interest area for CMS, such as strategies for improving perinatal care.

The demonstration period began on February 22, 2010, and was originally scheduled to end on February 21, 2015. However, CMS awarded no-cost extensions to all grantees who requested them (Table 1). For 11 States, the grant period will end 1 year later than the original termination date, on February 21, 2016; for three States, it will end 6 months later, on August 21, 2015; and for one State, it ended 3 months later, on May 21, 2015. Three States did not request an extension.

Table 1. CHIPRA quality demonstration projects by grant category

	Cat. A Report and Use Core Measures	Cat. B Promote Health IT	Cat. C Evaluate a Provider- Based Model	Cat. D Use Model EHR Format	Cat. E Grantee- specified	Length of No-Cost Extension
Oregon*	✓	✓	✓			6 months
Alaska	✓	✓	✓			6 months
West Virginia	✓	✓	✓			6 months
Maryland*			✓		✓	12 months
Georgia			✓		✓	12 months
Wyoming		✓	✓		✓	12 months
Utah*		✓	✓		✓	12 months
Idaho		✓	✓		✓	12 months
Florida*	✓	✓	✓		✓	12 months
Illinois	✓	✓	✓		✓	12 months
Maine*	✓	✓	✓			12 months
Vermont		✓	✓		✓	None
Colorado*			✓		✓	None
New Mexico			✓		✓	None
Massachusetts*	✓		✓		✓	3 months
South Carolina*	✓	✓	✓			12 months
Pennsylvania*	✓	✓		✓		12 months
North Carolina*	✓		✓	✓		12 months
Total Projects in Category	10	12	17	2	11	

Source: Centers for Medicare and Medicaid Services (CMS).

*Grantees. Partner States, where they exist, are listed in the rows directly below each grantee.

Evaluation of the Demonstration Grant Program

On August 9, 2010, AHRQ, in conjunction with CMS, awarded a contract to Mathematica Policy Research and its partners, the Urban Institute and AcademyHealth (hereafter referred to as the national evaluation team, or NET), to conduct a national evaluation of the demonstration grant program (see Appendix A for list of NET staff and technical expert panel (TEP) members).¹ The evaluation's primary objective was to learn about ways to improve the quality of health care for children enrolled in Medicaid and CHIP. Working under the direction of AHRQ and CMS, the NET designed the evaluation to provide insights into best practices and replicable strategies for improving children's health care quality.

To accomplish these goals, the NET gathered a substantial amount of qualitative and quantitative data regarding the demonstration projects implemented by grantees and their partners. Qualitative data sources included program documents and semi-annual and other reports; 776 key informant interviews with grantee and program staff, participating practice staff, and other stakeholders; and 12 focus groups with parents in selected States. Sources of quantitative data included administrative and claims data, self-reported assessments of medical home characteristics in selected States, and original survey data from physicians in selected States. Using a variety of methods, we analyzed these data to address a series of research questions.² (See Section 4 for information on how the evaluation design evolved over time.)

In most cases, we synthesized information from qualitative interviews with grantee and program staff and other stakeholders across similar projects to describe the implementation of demonstration activities, challenges encountered, lessons learned, and perceptions of the influence of demonstration activities on the quality of children's health care services. We used NVivo,[®] a qualitative data management and analysis tool to support our exploration of the data. We also intended to conduct formal impact analyses integrating quantitative data to determine whether particular interventions improved child health outcomes. However, for reasons related to data limitations and States' changes to their original implementation plans, we were unable to complete these analyses.

The evaluation addressed many of the original research questions, which we grouped into the five categories noted above. We also addressed additional questions that, during the course of the project, arose in response to developments in the policy environment or from insights gained during data collection and analysis. Some of these additional questions cut across or built on the five demonstration categories.

To address the needs of stakeholders—including Congress, AHRQ, CMS, States, the provider community, and family organizations—the NET disseminated results of analyses through *Evaluation Highlights*, implementation guides, manuscripts, and presentations. These products are listed in Appendix B and can be found on the national evaluation's Web site hosted by AHRQ (www.ahrq.gov/chipra/demoeval/). For a crosswalk with the complete set of CHIPRA research questions as they relate to NET products, send an email to CHIPRADemoEval@ahrq.hhs.gov.

To further document our plans and progress in meeting the evaluation's goals, we provided AHRQ with an evaluation design report (updated three times), a plan for providing evaluation-focused technical assistance (TA) to demonstration States (updated twice), a plan for developing and using our TEP (updated twice), a plan for obtaining feedback from key stakeholders, a dissemination plan (updated twice), four interim reports, and summaries of various meetings held during the course of the evaluation. These materials are available on request from AHRQ; send an email to CHIPRADemoEval@ahrq.hhs.gov.

Final Report

We have three primary goals for this final report. First, we present a synthesis of select findings contained in the products produced by the National Evaluation.³ We present this synthesis for the five original grant categories and for a category of cross-cutting findings. To develop this synthesis, we reviewed the documents, generated an initial list of key findings and themes, and held internal discussions to identify the most critical ones. Thus, our synthesis is selective, focusing on what we believe are the most useful findings for State and Federal agencies interested in improving the quality of health care for children. Additional findings—and many additional details about the programs that the demonstration States implemented—are contained in the documents themselves. Our findings are presented in Section 2.

Our second goal for this report is to present our observations about the structure of the grant program itself. Specifically, we note the program's key structural characteristics and discuss their implications for the implementation and sustainability of grantee projects and for the evaluation. We present these observations in Section 3.

Finally, we aim to identify key lessons learned in conducting the evaluation that may help AHRQ or CMS plan future evaluations. Based on our 5-year collaboration with AHRQ, CMS, and the demonstration States, we identified factors that contributed to and hindered the development of rigorous, useful findings from the evaluation. We describe these factors in Section 4 of the report.

2. Synthesis of Key Findings by Category

A final report of modest length must be selective in reporting the key findings and activities of a 61-month-long evaluation of a complex demonstration grant program. In this chapter, we have elected to synthesize the findings and insights presented in the products developed by the National Evaluation Team using, for the most part, the original grant categories. We encourage readers to review specific products for additional findings and nuances that we have not highlighted here. These products can be found in peer-reviewed journals and on the national evaluation's Web site hosted by AHRQ (www.ahrq.gov/chipra/demoeval/).

Category A Findings

Under Category A, 10 States were funded to collect, report, and assess the use of CMS' Core Set of Children's Health Care Quality Measures for Medicaid and CHIP (Child Core Set), as well as supplemental pediatric quality measures.⁴ Their objectives were to identify barriers to the collection and reporting of these measures and to build capacity for reporting and using them to improve the quality of care for children.

The Child Core Set was originally developed by AHRQ and CMS with substantial input from key stakeholders (including the organizations that developed and maintain measures included in the set). CMS released the initial technical specifications for reporting the Child Core Set in February 2011. The measures address a range of high-priority topics in child and adolescent health, such as access to primary care, preventive care (including vaccinations and developmental screenings), maternal and perinatal health (including prenatal care and low birthweight rate), behavioral health (including followup after hospitalization for mental illness), care of acute and chronic conditions (including medication management for asthma), oral health care (including dental visits for prevention and treatment), and patient/family experience with care.

State Medicaid/CHIP agencies began voluntarily reporting some State-level measures to CMS in 2011 for the Federal fiscal year (FFY) 2010 reporting period. CMS subsequently updated the Child Core Set. Specifically, CMS changed data sources for three measures for FFY 2012, retired one measure and added three measures for FFY 2013, and retired three measures for FFY 2014 reporting. The States' performance on these measures can be found in the Secretary's Annual Report on the Quality of Care for Children in Medicaid and CHIP, usually released in October of each year.⁵

In addition to State-level reporting of the Child Core Set to CMS, there is the potential to use these measures for reporting by health care organizations, such as child-serving practices, health systems, and managed care organizations (MCOs). Beyond just reporting performance on the Child Core Set, States, MCOs, health systems, and practices can use the measures in quality improvement (QI) initiatives. Focusing on these two general activities of Category A (reporting measures and using them for QI) States produced the following findings, which we discuss below in more detail:

- States encountered a variety of barriers to reporting the Child Core Set to CMS and developed diverse methods to address the barriers.
- States applied a range of strategies for using quality measures as part of broader QI initiatives.
- Practices encountered numerous challenges to reporting quality measures (including but not limited to the Child Core Set), and some developed methods to address them.
- States developed diverse strategies for overcoming barriers providers faced in using measure reporting to improve quality of care.

1. States encountered a variety of barriers to reporting the Child Core Set to CMS and developed diverse methods to address the barriers.

Using information from Illinois, Maine, Oregon, and Pennsylvania, the NET developed a manuscript (under review) entitled “What factors influence the ability of State Medicaid agencies to report the Child Core Set of health care quality measures? A multicase study.” Analysis of the study yielded the following findings:

- Key factors affecting a State’s ability to report the Child Core Set measures to CMS included:
 - Technical factors, such as clarity and complexity of measure specifications; data availability, completeness, and linkages; and software capabilities.
 - Organizational factors, such as a history and culture of data use, support from agency and other State leadership, and availability of skilled programmers.
 - Behavioral factors, such as staff motivation and external demand for measures.
 - State health care policy environment, including the structure of Medicaid and CHIP agencies, the level of managed care, and other health care reform activities.
 - Participation in external capacity-building activities, such as through the CHIPRA quality demonstration.
- States used numerous resources and significant time to interpret and apply CMS’ specifications to available State-specific data.
- Access to fee-for-service claims data enables but does not guarantee that all administrative measures can be accurately reported.
 - Providers must consistently use the billing codes in the measure specifications, otherwise the measure will underestimate quality of care.
 - In some cases, States have one billing code to cover multiple types of services (for example, developmental screening and behavioral health screening). Such codes cannot be used to measure receipt of each specific service.
- States typically faced major technical challenges linking Medicaid/CHIP data to other data sources, such as immunization registries and vital records, to produce quality measures.
- States had a difficult time producing core measures that require EHR data because most States, health systems, and practices have not yet developed the infrastructure needed to support data

transmission from providers' EHRs. Another challenge was that most Child Core Set measures are not yet specified in the standardized Health Quality Measure Format language for EHR reporting.^{6,7}

- Diverse stakeholders in most States expressed a demand for children's health care quality measures reported regularly at the health system, health plan, or practice level rather than annual reports at the State level. The Child Core Set was not designed for practice-level reporting, but many stakeholders wanted to use the measures at the practice level.
 - Adapting the Child Core Set measures for these various levels requires modifications to the original measure specifications. These modifications and other State-to-State variations in measure production processes may influence the ability of CMS and States to compare measures across States and use them to drive QI activities.

2. States applied a range of strategies for using quality measures as part of broader QI initiatives.

Evaluation Highlight 11 identified lessons learned about measure-based strategies that additional States can use to improve the quality of care. Analysis of information from Alaska, Florida, Illinois, Maine, Massachusetts, and North Carolina yielded the following findings:

- In some of these States, State-level QI activities were supported by quality reports that were developed for specific State audiences and that compared the State's performance with neighboring or similar States, as well as with national benchmarks.
- Because improving performance typically requires a collective effort from many stakeholders, some States formed workgroups or held formal meetings to review quality measure reports with key stakeholders (including staff at child-serving agencies, large or influential practices, health plans, and health systems) with the goal of focusing on specific QI priorities.
- Improving quality of care required States to move beyond producing and disseminating quality measure reports to take one or more additional steps, such as the following:
 - Establish regular procedures for monitoring quality of care at practice or health system levels, which can help identify providers who are lagging on certain measures.
 - Implement policy and programmatic changes in clinical documentation procedures or billing processes, which can make data more accurate and timely.
 - Provide individualized and group TA to practices and health systems through practice facilitation (also called QI or practice coaching), QI specialists, Webinars, and learning collaboratives that will help providers develop their own measure-based QI initiatives.
 - Initiate statewide stakeholder engagement efforts that seek to build an enduring commitment to improving quality of care for children.
 - Consider pay-for-reporting, pay-for-performance, pay-for-improvement, or other incentive programs to spur quality reporting and improvement.

Additionally, through a survey of physicians in two demonstration States (North Carolina and Pennsylvania) and one comparison State (Ohio), we found that the majority of child-serving physicians receive quality reports and believe they are effective for QI, but only one-third of these providers

actually use quality reports in their QI activities. Physicians in the demonstration States used quality reports for QI at about the same rate as physicians in Ohio.

3. Practices encountered numerous challenges to reporting quality measures (including but not limited to the Child Core Set), and some developed solutions to address them.

Two of our *Evaluation Highlights* (1 and 5) describe lessons learned about facilitators and barriers that States and practices encounter as they work to report practice-level quality measures. Analysis of information from Maine, Massachusetts, North Carolina, Pennsylvania, and South Carolina—the States covered in these *Evaluation Highlights*—yielded the following findings:

- It was critical for States to collaborate with physician practices and providers in selecting or refining measures for QI projects because it built buy-in and ensured that measures were meaningful, feasible, and useful for practice-level improvement.
 - Providers expressed preferences for measures that were timely, under the influence of the practices' activities, and useful to the practice's QI efforts.
 - Both States and practices had to be flexible to reach agreement on measures that are high-priority, actionable, and appropriate for busy practices.
- It was unexpectedly time- and resource-intensive for States to adapt measures originally designed for reporting at the health plan or State level for use at the practice level. The administrative and technical steps needed to calculate quality measures at the practice level are quite different from the steps needed for the State level.
 - States had to adjust specifications to fit the reporting capabilities and needs of practices, including testing new data sources and modifying the measure denominator to the practice level. However, the adjustments may compromise the reliability and validity of measures if specifications for practice-level measures move too far from original specifications.
 - Accurately attributing patients to providers was especially challenging because some patients are not attached to specific providers, and some are administratively linked to one provider but actually seek care at another site.
- States used a variety of data sources to produce practice-level measures, including established State databases containing Medicaid claims and enrollment and eligibility data; statewide immunization registries; and Health Information Exchanges (HIEs) and provider-submitted data (direct EHR data or manual review of EHR or paper charts).
 - It was important for States to plan for resources to manage unexpected data access and quality issues. States were able to overcome some challenges by having experienced data analysts and alternative data extraction plans in place.
- States attempted various strategies to overcome information technology (IT) and data infrastructure challenges, such as outdated or underdeveloped claims systems, HIE, and EHRs. Strategies included involving practices in data collection (via manual extraction of data from EHRs or charts) and developing workarounds with their EHRs. However, many of these activities relied on

grant funding and staff and are not sustainable to support collecting and reporting practice-level quality measures in the long run.

4. States developed diverse strategies for overcoming barriers providers faced in using measure reporting to improve quality of care.

Our first and fifth *Evaluation Highlights* included lessons learned about facilitators and barriers that States and practices encountered as they worked to use practice-level quality measures to inform their own QI efforts. Analysis of information received from Maine, Massachusetts, North Carolina, Pennsylvania, and South Carolina and covered in these *Evaluation Highlights* yielded the following findings:

- When practice staff began to apply quality measures in their own practice, they often discovered clinical documentation limitations (such as incomplete or inconsistent documentation in EHRs and paper charts) and therefore had to make improvements in documentation so they could have accurate information for their QI efforts.
- When practice staff first generated quality reports based on accurate data, they frequently discovered that their performance was worse than they expected.
- For QI activities to be effective, they required the involvement of all staff (including physicians, nurses, and administrative staff). To engage staff, practices made them aware of quality measures, why they matter, and each person's role in QI.
- Practices found measure reports more useful for identifying QI priorities than for guiding and assessing QI projects, mainly because data receipt often lagged; therefore it was difficult for them to use reports to assess and make adjustments to redesigned workflows in real-time.
- States used a variety of other strategies or combinations of strategies to support QI efforts at the practice level, including payments or stipends to participating practices, training, and TA.
 - For example, to encourage QI, Pennsylvania offered pay-for-reporting and pay-for-performance incentives to participating health systems. Incentives included \$10,000 per measure reported from an EHR for the base year (up to 18 measures, or \$180,000) and \$5,000 for each percentage point improvement per measure, up to five points, or \$25,000 per measure, capped at a total payment of \$100,000. The State offered relatively little TA.
 - In contrast, South Carolina provided extensive TA rather than payments or stipends, and used the Child Core Set as the foundation for assisting primary care practices via a multiyear learning collaborative focused on quality improvement (plan-do-study-act) cycles. The State also provided practice staff customized support from practice facilitators.
- States had to invest substantially in both the human and automated components of data extraction to support use of EHRs for practice-level reporting. EHR-based reporting will never be fully automated. For example, each time an EHR was updated or modified, programmers and analysts had to reconsider data coding and modify procedures to report the measures.

Category B Findings

The overall goal of the Category B projects was to identify effective strategies for using health IT to improve the quality of children’s health care, reduce Medicaid and CHIP expenditures, and promote transparency and consumer choice. Based on their final operational plans (developed in the first year of the demonstration), the 12 States that originally intended to implement Category B projects proposed to use several types of health IT and implementation strategies to pursue the goals for their projects (Table 2). These strategies included using various combinations of EHRs, personal health records (PHRs), and HIE pathways for multiple purposes. Purposes included automated reporting of the Child Core Set of quality measures; reporting of Early and Periodic Screening, Diagnosis, and Treatment (EPSDT) measures; supporting clinical decisionmaking; promoting QI in clinical settings; supporting the informational needs of public health agencies; fostering consumer engagement; and coordination across different types of providers (especially in connection with medical homes).

Table 2. Health IT strategies to be used by demonstration States as of June 2011

Health IT Strategies	OR ^c	AK ^c	WV ^c	WY ^b	UT ^b	ID ^b	FL ^a	IL ^a	ME ^a	VT ^b	SC ^a	PA ^a	Total
Creating or enhancing a regional child health database/warehouse					✓	✓	✓	✓	✓	✓		✓	7
Linking databases across agencies					✓	✓	✓	✓	✓	✓		✓	7
Increasing access to data for targeted users				✓	✓	✓				✓		✓	5
Encouraging practices to use EHRs and quality measures	✓	✓	✓	✓				✓			✓	✓	7
E-reporting from practice to HIE/child health database	✓	✓	✓		✓	✓	✓		✓	✓	✓		9
E-reporting from HIE/child health database to practices and/or health agencies	✓	✓	✓					✓	✓	✓	✓		7
Devising/refining/implementing incentive payments based on reporting data											✓		1

Source: State final operational plans.

✓ State planned to employ strategy in Category B demonstration.

^a State planned to link some elements of its Category B project to its Category A project.

^b State planned to linked some elements of its Category B project to its Category C project.

^c State planned to linked some elements of its Category B project to both its Category A and C projects.

Most Category B States planned to implement or improve electronic reporting from practices to an HIE or children’s health database, including developing standard reporting tools, forms, and formats. South Carolina had explicit plans to offer incentives for reporting through payment reform. Most States also intended to pursue some form of electronic reporting from an HIE or children’s health database to practices or health agencies (for example, patient-level quality measure reports).

Based on information collected for the evaluation, we identified three findings about the Category B projects, which we discuss in further detail below:

- Most demonstration States faced major challenges that hindered implementation of their Category B projects.
- Projects involving the development of electronic screening methods were able to achieve their objectives.

- Projects that aimed to develop focused health IT applications were successfully implemented.

1. Most demonstration States faced major challenges that hindered implementation of their Category B projects.

A review of information collected during site visits and other discussions with project staff underscores the following obstacles States encountered while executing Category B projects:

- The diversity and turnover of EHR products used by practices and insufficient functionality in EHRs to collect and analyze data posed barriers to EHR use.
 - As an example, South Carolina achieved limited success in producing practice-level quality measure reports by combining Medicaid claims data with EHR data. The limitation was largely because of the difficulties in developing the infrastructure and functionality needed to record and transfer pediatric data from practices' EHRs to the States. The diversity of EHRs used by practices and the amount of modifications needed to those EHRs further complicated and delayed data extraction.
- Challenges related to interoperability between the practices' EHRs and State databases, including HIEs, were common among many States. In many cases, these challenges went largely unresolved. Furthermore, most States had not yet developed the infrastructure, such as HIEs, to exchange EHR data with providers. As a result of these barriers, program staff in many States focused on other demonstration projects.
 - As an example, in West Virginia, State program staff dropped their plan to create and implement a PHR—the primary goal of their Category B project—for two reasons. First, the platform would have duplicated the function in the EHRs that practices were already using, and second the State decided not to implement an HIE, which was necessary for the PHR to be implemented as planned.
- Challenges related to data ownership and security issues also stalled projects in some States.
 - As an example, in Illinois, development of a statewide prenatal minimum electronic data set that would extract data from EHRs and link to the State HIE eventually foundered. The State was unable to finalize development because neither the State nor the vendor wanted to own the repository that was tested in the early stages of the grant.
- Practice staff often needed training and TA to effectively use their EHRs.
 - As an example, in Alaska, participating practices needed substantial assistance to improve use of their EHRs to support practice functions and QI; as a result, there were few remaining grant resources in that State available for additional work in this grant category.

2. Projects involving the development of electronic screening methods were able to achieve their objectives.

Colorado and New Mexico implemented an electronic screening questionnaire.⁸ This computer tablet-based risk screening instrument, the electronic Student Health Questionnaire (eSHQ), was used by

SBHCs to improve early identification of health risk behaviors and initiation of discussions about protective factors for adolescents.

Pennsylvania was also able to implement its electronic screening project as planned. This project involved introducing a fully electronic developmental screening questionnaire in 12 pediatric primary care sites associated with the Children's Hospital of Philadelphia between 2011 and 2013.

Additional details regarding each of these projects are also available in special innovation features⁹ posted on the national evaluation Web site. These three States' projects provide the following key findings:

- Technology can be used to streamline the administration of screening questionnaires to identify children with health risks, such as developmental delay or autism.
- The use of electronic screening tools in practices and SBHCs can enhance documentation that services were provided and can support data quality, tracking, and monitoring and a higher quality of care.
- Adolescents, families, and providers find electronic screening easy to use. Additionally, adolescents valued tablet-based screening as a way of communicating directly and privately with their doctors.
- Although electronic screeners afford many benefits, there are also costs to providers related to ongoing training and technical support.

3. Projects that aimed to develop focused health IT applications were successfully implemented.

Although many States halted their health IT efforts in response to challenges noted elsewhere in this report, two States were each able to implement stand-alone and specific health IT products that are likely to be sustained beyond the grant period.

- With support from the grant, Utah developed an online health platform that practices can use to share information about QI work including cumulative performance on quality measures and graphic depictions of data in a time sequence. This Web-based platform has been used for learning collaboratives and will form the basis of future QI activities in Utah. In addition, other States are using the platform, and their payments to Utah are now supporting maintenance costs.
- Wyoming developed a data dashboard to track CME performance on quality and output measures. The State will continue to use an expanded version of the dashboard to track CME quality under a new contract to expand CME services statewide.

Category C Findings

The goal of the Category C projects was to develop, implement, and determine the impact of selected provider-based models on the delivery of children's health care, including access, quality, and cost. All of the demonstration States except Pennsylvania implemented a Category C project. To achieve the Category C goals, grantees and partner States used one of three strategies: (1) transforming child-serving practices into PCMHs, (2) strengthening SBHCs; or (3) developing CMEs for children with serious emotional or behavioral disorders. These strategies sometimes overlapped and expanded, with SBHCs

working to develop PCMH features and many PCMH projects strengthening practices' general QI skills. We briefly describe each strategy here.

Transforming child-serving practices into PCMHs. Seven grantees, inclusive of 12 States, implemented efforts to enhance PCMH features of child-serving practices.¹⁰ These efforts involved varying combinations of strategies to promote practice transformation, including learning collaboratives, one-on-one QI facilitation, TA related to collecting and reporting quality measure data, TA related to building family engagement in practice activities and QI strategies, and practice stipends. About 140 child-serving practices participated in these efforts to some extent (excluding practices that served as comparison practices).¹¹ Through interviews with project staff in the 12 States and staff in many of the participating practices, as well as focus groups with families whose children were patients of these practices, we gathered substantial qualitative data about these PCMH transformation efforts. We also reviewed medical home survey data submitted by States. We analyzed that information to address questions about implementation processes and perceived outcomes of these models. Because practice transformation was such a predominant activity within the demonstration, we devoted considerable effort to documenting our findings in four *Evaluation Highlights* (nos. 3, 7, 9, and 13) and three manuscripts.

Strengthening SBHCs. Colorado and New Mexico collaborated on efforts to enhance PCMH features of 22 SBHCs. These projects involved practice facilitators, engagement with youth and their families, and collaboration between SBHCs and other providers. Two *Evaluation Highlights* (nos. 3 and 8) described these efforts.

Developing or enhancing CMEs. Maryland, Georgia, and Wyoming aimed to enhance or develop ways for providing services to youth with serious emotional disorders. Specifically, these States examined means for locating oversight and coordination of services for children with serious emotional disorders outside of the traditional provider setting through the use of separate CMEs. We developed an implementation guide that described and built on their efforts.

Looking across the diverse PCMH, SBHC, and CME projects implemented by the 17 States that participated in Category C, we identified seven findings that we believe are especially relevant to AHRQ, CMS, and the States:

- Learning collaboratives were useful for supporting practice transformation when implemented with appropriate clinical expertise and collaboration among State and practice staff.
- The addition of new staff members was viewed as an important factor in practices' ability to improve QI and PCMH capacity.
- Measuring progress in practice transformation was important for driving QI improvement.
- States recognized the importance of consumer engagement but noted major challenges in accomplishing this goal.
- Demonstration States identified barriers unique to providing high quality care for adolescents, as compared to children generally, and developed strategies to address them.
- Using peers to support caregivers of children with special health care needs provided valuable assistance to families.

- Successful development of CMEs to serve youth with serious behavioral and emotional disorders required a multi-pronged approach.

1. Learning collaboratives were a useful means for supporting practice transformation when implemented with appropriate clinical expertise and collaboration among State and practice staff.

Learning collaboratives were used in the 12 States that had projects focused on helping practices or SBHCs enhance or adopt features of the PCMH model.¹² Analysis of data provided by key informants in these States yielded the following findings:

- States discovered that learning collaborative topics need to be relevant to providers. Generating the topic list with substantial provider input generally resulted in engaging meaningful provider participation. Many States solicited frequent feedback from the practices and made midcourse adjustments to collaboratives' structure and content.
- Maintaining provider engagement and participation in collaboratives is challenging given competing demands for time. States found the following strategies to be useful in recruiting and ensuring the ongoing engagement of practice staff:
 - Providing practice stipends to offset some of the costs of missed revenue resulting from taking time off from care delivery to attend learning collaborative sessions.
 - Aligning demonstration efforts with professional development requirements such as offering providers Maintenance of Certification (MOC) credits in exchange for participation in the learning collaboratives.
 - Aligning demonstration efforts with external financial incentive programs, such as focusing learning collaboratives on clinical topics covered by Medicaid pay-for-performance measures.
 - Offering a combination of traditional didactic instruction and interactive learning activities such as competitions, live demonstrations, and peer networking.
 - Offering Web-based learning sessions as alternatives or complements to in-person meetings. Web-based meetings were favored by some providers because they saved on travel time, but it was harder for some States to keep attendees focused and engaged in the Web-based discussions.
 - Supplementing learning collaboratives with individualized practice facilitation allowed practices to obtain customized one-on-one assistance and kept practices on task by holding them accountable for learning collaborative "homework."
- Finding the right mix of participants in a learning collaborative can foster the exchange of information among practices. Sharing experiences was easier when participating practices had similar pre-existing QI and PCMH capacity and patient populations and were working on similar topic areas and measures.
- States felt that tracking practices' performance on quality measures over time was helpful in identifying areas for improvement and progress achieved, but reporting on these quality measures was sometimes time consuming and challenging for practices.

- To support practices' QI efforts, States learned that it was important to use a judicious number of quality measures tightly linked to the topics focused on in learning collaboratives and to not require too-frequent reporting of measure data.
- To build providers' QI abilities related to the collection, analysis, interpretation, and use of quality measure data, States learned that it was important to provide adequate supports such as learning collaborative sessions, QI materials and tools, and individualized assistance via practice facilitators.
- Although States were often able to effectively engage participating providers in learning collaborative activities, these providers frequently experienced challenges in spreading and sharing information among other practice staff who did not attend meetings or actively participate in activities. This finding was especially true if the learning collaborative participant was not the lead physician in a practice.

2. The addition of new staff members was viewed as an important factor in practices' ability to improve QI and PCMH capacity.

States used CHIPRA funds to provide participating practices with various kinds of additional staff, such as care coordinators, practice facilitators, and parent partners. These additional staff provided new or enhanced services and support specifically related to enhancing QI and PCMH capacity. Analysis of project reports and data from key informant interviews yielded the following findings:

- Adding new staff members is particularly effective when they have the required technical skills and are integrated into the existing organizational culture.
- Practices that played a substantial role in hiring new staff found it easier to integrate a care coordinator than if the State assigned new staff to a practice because practices could select individuals with the credentials, demeanor, and communication style that best fit their needs and culture.
- New staff appeared to be most effective under two conditions: (1) when existing staff, such as clinicians and administrators, valued their contributions and (2) when existing staff understood the role that the newcomers could play in achieving practice transformation and improved quality of care.
- States and practices found that practice facilitators need to limit the number of practices they work with to allow them to provide meaningful individualized support.
- In many cases, States and practices that used demonstration funds to help pay for additional staff were not able to sustain these staff after the grant period.
 - Practices that highly valued the contributions of new staff, such as care coordinators, were more likely to seek alternative funding mechanisms to support these positions after the grant period.

3. Measuring progress in practice transformation was important for driving QI improvement.

States recognized the need to assess the extent to which their projects were accomplishing the goals of practice transformation and to use these assessments to shape ongoing efforts.

- States working to enhancing PCMH features of participating practices understood the need to assess the extent to which the practices were adopting these features.
 - States tended to select assessment tools based on a variety of factors, including other medical home activities in the State, the target population for the medical home intervention, and familiarity with particular approaches. CMS did not require States to use the same assessment tool.
 - Illinois used the National Committee for Quality Assurance (NCQA) PCMH self-assessment tool; Florida, Maine, Massachusetts, Idaho, North Carolina, South Carolina, and Utah used some version of the Medical Home Index (MHI); and Oregon, Alaska, and West Virginia used components from both tools.¹³
- The States working to enhance the medical home features of SBHCs worked with practice facilitators to monitor quality measure change over time using the Medical Home Index – Revised Short Form (MHI-RSF).¹⁴
- The three CME demonstration States used grant funding to hire a contractor to design an evaluation plan that included measuring the key outcomes or results of CME adoption or expansion, as well as measuring care processes to support QI.

4. States recognized the importance of consumer engagement but noted major challenges in accomplishing this goal.

States experimented with methods to engage families and adolescent patients in QI activities, including using youth engagement specialists, family partners, family advisory councils, and community service boards. These activities yielded several key findings:

- Enlisting family caregivers to provide practices with feedback was valuable for identifying consumer perspectives, but challenging.
 - Parents had limited time available to contribute feedback due to their multiple and competing priorities.
 - Some parents were not accustomed to “advisory” roles and felt uncomfortable providing feedback. The opportunities for parents to provide feedback may not have been optimal given their preferences and abilities (e.g. long surveys, large group meetings, meetings at inconvenient times).
 - Some State staff noted that some practices resisted seeking parent feedback because they feared that parents would ask for changes that the practices deemed not feasible (such as offering evening appointments).
 - Many practices worked to change features that they believed are important to providing high quality care but that are not noticeable to parents, such as the use of team huddles, improvements to EHRs, and use of patient registries. The low profile of these improvements made it challenging for parents to detect them and provide feedback.

- Enlisting youth participation in project activities carried benefits.
 - In SBHCs, youth engagement specialists and youth advisory boards helped to increase students' and families' use of the centers.
 - Georgia noted that engaging youth and caregivers in designing peer support trainings for youth with social and emotional disorders helped develop a curriculum that was comprehensive, accessible, and relevant.
- 5. Demonstration States identified barriers unique to providing high quality care for adolescents, as compared to children generally, and developed strategies to address them.**

Colorado, New Mexico, North Carolina, and Utah implemented projects that aimed to improve health care for adolescents. These projects identified the following key challenges to providing high quality care to teenagers:

- Many primary care providers do not use adolescent risk screening tools effectively or efficiently.
- Perceived shortages of mental health professionals in some areas have made some primary care providers hesitant to screen for mental health conditions.
- Some primary care providers were uncomfortable discussing sensitive health issues or conditions with teenagers and had difficulty ensuring the confidentiality of information that teens communicate.

In the context of their CHIPRA demonstration projects, the States identified multiple strategies to overcome barriers to providing high quality adolescent health care. These strategies aimed to increase providers' willingness, frequency, and skill in administering adolescent health risk assessment questionnaires and engaging in private consultations with adolescents regarding responses. Strategies include:

- Training in tips and techniques for engaging adolescents and using screening tools effectively and efficiently.
- Implementing electronic screening methods that assess adolescents' risks and strengths, collect sensitive information confidentially, and help providers prioritize topics to discuss during office visits.
- Training in State and Federal privacy rules.
- Providing information about local referral resources by developing resource lists or collaborating with local mental health professionals.
- Working to identify reimbursement for health risk screening and anticipatory guidance for adolescents.
- Offering MOC credits for participating in educational training opportunities specifically related to providing high quality care for adolescents.

6. Using peers to support caregivers of children with special health care needs provided valuable assistance to families.

By providing emotional solace, practical tips, and general encouragement, peer support can be helpful to parents who care for children with special needs. Some States tried a provider-based approach, through which providers link parents who volunteer to provide peer support with parents who ask for such support. Some States worked to develop a peer support workforce whose services are reimbursable through Medicaid. These activities provided the following findings:

- Individuals who provided peer support needed comprehensive training on their roles and responsibilities, a clear understanding of the time commitment required, and access to a support system.
- Caregivers who were best suited to provide peer support were those who had experience navigating the health system and caring for their own child with special health care needs. However, they themselves needed support when they were faced with crises involving their own children.
- Educating health care providers about caregiver peer support helped to increase their understanding of and interest in supporting this service.
- In Maryland and Georgia—States that developed a formal mechanism for certifying and funding caregivers to provide peer support—the services were more likely to be sustained than in other States where peer support was funded only by the demonstration grant.

6. Successful development of CMEs to serve youth with serious behavioral and emotional disorders required a multi-pronged approach.

As the lead State, Maryland worked with its two partners (Georgia and Wyoming) to help them develop or improve CMEs. These States worked to identify funding streams, establish organizational infrastructures, and develop training programs. Challenges included competing priorities at the State level, resistance to a new model on the part of established service providers, and a steep learning curve for most stakeholders. The projects in these three States provided the following findings:

- CMEs can use different management structures, depending on existing service infrastructure. In Maryland (which has two CME models), CMEs are managed by an interagency State-level organization and counties; State Medicaid offices run CMEs in Georgia and Wyoming.
- Gaining financial support from multiple child-serving agencies (Medicaid, welfare, juvenile justice, health, and others) was difficult. Agencies were more willing to provide a funding stream for CMEs if they were involved in the design (for example, determining the eligibility criteria).
- When a State decided to use an out-of-State organization for CME services, State staff had to work diligently to build local trust to overcome provider reluctance to refer youth for services.

Category D Findings

The goal of the Category D projects was to assess the Children's EHR Format (Format). The Format was commissioned by CMS and AHRQ to bridge the gap between the functionality present in most EHRs currently available and the functionality that would more optimally support the care of children. The Format, officially released by AHRQ in February 2013, is a set of 695 recommended requirements for EHR data elements, data standards, usability, functionality, and interoperability that need to be present

in an EHR system to address health care needs specific to the care of children. (The current version of the Format is available at <https://ushik.ahrq.gov/mdr/portals/cehrf?system=cehrf>.)

Two demonstration grantees (Pennsylvania and North Carolina) conducted projects in this category but approached the task somewhat differently. Pennsylvania collaborated with EHR vendors and five of the State's health systems (three children's hospitals and affiliated ambulatory practice sites, one federally qualified health center, and one small hospital) to implement and test the Format and determine the extent to which EHRs could yield data for calculating the Child Core Set of quality measures. Consequently, their Category D efforts were closely linked to Category A quality measure reporting activities. In contrast, North Carolina used EHR practice facilitators to work with 30 individual practices to identify the degree to which their EHRs already were consistent with the Format and to gather feedback on Format specifications. Facilitators also focused on training staff in these practices on how to use EHR functionalities that already met Format requirements but were not being used.

Evaluation Highlight 10 presents findings related to these States' experiences assessing the Format. We summarize these findings here:

- Comparing the Children's EHR Format with existing EHRs was challenging but valuable.
- EHR vendors were reluctant to engage in the demonstration projects, especially because the U.S. Department of Health and Human Services (HHS) has not mandated that vendors adhere to the Format.
- The Format's complexity overwhelmed providers' resources to fully understand it.

1. Comparing the Children's EHR Format with existing EHRs was challenging but valuable.

One of the first steps that States and practices took was to compare their own EHRs functionality with the 695 requirements contained in the model EHR Format. This process produced the following conclusions:

- States and providers generally found the Format to be a major advance in the specification of child-oriented EHR functions. Appreciation for the Format's thoroughness, however, was diminished by the time-consuming process of comparing the Format with existing EHRs.
- Vendors and practices/health systems often were at odds about whether existing EHRs met Format requirements. It took time to resolve discrepancies—often because practice staff were not aware of their own EHRs functionalities and in some cases because of ambiguity in the Format's requirement descriptions.
- The comparison process meant that many practices learned more about the capabilities of their EHRs and worked to determine how to make Format requirements applicable to practice workflow.

2. EHR vendors were reluctant to engage in the demonstration projects, especially because HHS has not mandated that vendors adhere to the Format.

EHR vendors' reluctance stemmed in part from their need to pay attention to other priorities (such as ICD-10 transition, and achieving certification under the CMS' EHR Incentive Program). They also saw little reason to voluntarily make their products Format-compliant or to meet the needs for children's

health IT more generally. Overall, lack of vendor participation impeded progress in Category D activities in both States.

- North Carolina found that vendors needed clinical and informatics guidance to incorporate the Format requirements in a way that supports the State's desired improvement in children's health care.
- When EHR facilitators and health systems got the attention of vendors, their assessment of the Format helped them to identify and discuss providers' expectations for a child-oriented EHR.

2. The Format's complexity overwhelmed providers' resources to fully understand it.

Many stakeholders suggested that it would be more fruitful to have a Format that includes a narrower subset of EHR requirements that align closely with current QI priorities or are limited to a subset of critical/core requirements. To that end, AHRQ has convened two workgroups to further evaluate the Format and its potential uses; an abridged version including only the critical and core requirements is now available.¹⁵

Category E Findings

CMS guidelines for Category E offered States the opportunity to implement additional strategies aimed at improving health care delivery, quality, or access. The activities could relate to one of the CMS key program focus areas listed in the grant solicitation or to another area of the grantee's choice, provided it complemented the activities performed under another grant category. Because the guidelines for this category were less specific than for Categories A through D, States addressed a range of topics; 11 States fielded Category E projects:

Colorado and New Mexico worked with selected SBHCs in their States to increase youth engagement in their health care. As part of this project, the States developed a Youth Engagement in Health Services (YEHS) survey for high school and middle school students. In both States, participating SBHCs used tablet computers to administer the survey to youth. In Colorado, SBHCs will not be using the survey after the demonstration period. New Mexico integrated about half of the YEHS questions into its existing Student Satisfaction Survey, which all SBHCs that receive State funding are required to administer.

Florida and Illinois established stakeholder workgroups to focus on improving the quality of perinatal and early childhood care for children enrolled in Medicaid and CHIP. Florida provided CHIPRA dollars to the University of South Florida to promote the Florida Perinatal Quality Collaborative (FPQC). During the later years of the project, the collaborative met every 6 months, bringing together hospitals and other perinatal stakeholders to improve the quality of care for mothers and newborns. In its first QI project, the FPQC focused on reducing elective pre-term births through delivery room interventions. The project was viewed a success; rates of elective scheduled early-term deliveries decreased among the 26 participating hospitals.¹⁶ The FPQC's partners (March of Dimes, the Hospital Engagement Network, and the Blue Cross Foundation) may sustain its work after the grant period.

The Illinois Perinatal Quality Collaborative (IPQC) began with seed funds from the CHIPRA grant and now has a membership of more than 100 hospitals. State demonstration staff also were on the leadership team of the IPQC. Activities have included several statewide conferences, an early elective delivery (EED)

initiative involving 49 hospitals (41 have achieved the goal of reducing their EED rates to less than 5 percent), a neonatal nutrition initiative involving 18 neonatal intensive care units (NICUs), an initiative involving 106 hospitals to improve accuracy of 17 key birth certificate variables, and an initiative involving 28 NICUs to improve the quality of care in the first hour after a child's birth. Although CHIPRA funding supported the creation of the collaborative, the group has also received funds from other sources, including March of Dimes, Illinois Department of Healthcare and Family Services, the Illinois Hospital Association, and a Centers for Disease Control and Prevention (CDC) grant, and will continue operations after the CHIPRA demonstration ends.

Maryland, Georgia, and Wyoming used Category E funding to support their Category C work to develop or expand CMEs for youth with serious emotional and behavioral health needs. We note each State's specific activities conducted under their Category E projects and their sustainment status:

- Maryland surveyed and held focus groups with behavioral health providers, families, and youth on crisis response and family support services to understand families' experiences related to these services and identify gaps in service availability. Based on these discussions, the State developed a report outlining best practices for crisis response and disseminated it to local organizations providing these services. The State also determined an appropriate reimbursement rate for crisis and family support services and included these services in a new State plan amendment.
- Georgia established a network of certified family peer support specialists to develop related training programs and to obtain Medicaid reimbursement for the services provided by these specialists. The State was able to institute a training and certification program for family and youth peer support specialists that will continue after the grant period through separate funding mechanisms.
- Wyoming used CHIPRA funds to support the *Too Young, Too Much, Too Many* Program, which tracks patterns of psychotropic medication prescribing in Medicaid, addresses misuse by physicians, and determines whether youth need additional intervention. The State renewed and expanded its contract with their pharmacy benefit manager to continue this program after the grant period.

Massachusetts formed the Children's Health Quality Coalition, a 60-member multi-stakeholder group representing clinicians, payers, State and local government agencies, family advocacy groups, and individual parents and families. During the demonstration, the coalition reviewed child health quality measure reports to analyze gaps in care and identify priority areas, convened task forces and workgroups that advanced its agenda in priority areas, and developed a Web site with resources to help practices and families improve the quality of care. Going forward, the coalition will be incorporated into the Massachusetts Health Quality Partners' coalition agenda and initiatives. The Massachusetts Children's Health Quality Coalition's Web site¹⁷ remains live, and content has been updated to reflect its new organizational home.

Utah and Idaho, with support from the National Improvement Partnership Network (NIPN), established or strengthened State-based pediatric QI networks to support continued development of QI initiatives for children:¹⁸

- Idaho established the Idaho Health and Wellness Collaborative for Children, which will be housed at the St. Luke's Children's Hospital.
- In Utah, the CHIPRA project team was closely linked to an existing improvement partnership network (Utah Pediatric Partnership to Improve Healthcare Quality, or UPIC) that provided

intellectual leadership for the State’s demonstration grant. After the grant period, UPIC will continue to seek internal and external support for QI initiatives for children in Utah—efforts that will be informed by experiences and relationships developed through the grant.

Vermont used Category E funding to contract with NIPN to provide TA to improvement partnerships (IPs) in more than 20 States, develop core measure sets, and hold both annual operations trainings attended by representatives from IPs nationwide and monthly “all-site” conference calls. NIPN is run through the Vermont Child Health Improvement Project based at the University of Vermont’s College of Medicine.

Cross-Cutting Findings

In addition to the category-specific findings, *Evaluation Highlights 4 and 6* and a manuscript on sustainability include findings that cut across the five demonstration categories.^{19,20} Key findings include:

- To ensure that child health care remains an important topic on State health policy agendas, demonstration States leveraged the CHIPRA grant to develop or strengthen connections to key policymakers.
- Of the project elements that were in place at the end of the fifth year of the demonstration, more than half were, or were highly likely to be, sustained after the grant period was over.
- Demonstration grants allowed States to gain substantial experience, knowledge, and partnerships related to QI for children in Medicaid and CHIP—a resource we refer to as “intellectual capital.”

1. To ensure that child health care remains an important topic on State health policy agendas, demonstration States leveraged the CHIPRA grant to develop or strengthen connections to key policymakers.

Demonstration States reported that the presence of a CHIPRA grant sent State policymakers a signal about the importance of improving the quality of care for children and adolescents. The prestige of winning the grant lent legitimacy to staff efforts to improve the quality of care for children. In many States, it also allowed key staff to participate in policy discussions and supported them in including children in the broader health reform activities occurring in the State. Project staff in several States also learned how to leverage data and analysis generated through the CHIPRA quality demonstration to engage policymakers, raise awareness about pediatric health issues, and suggest potential solutions. For example, demonstration staff in Maryland used behavioral health claims data to identify gaps in the availability of crisis response tools throughout the State and made recommendations for a redesign of the State’s crisis response system.

The strategies that States used to elevate children on health policy agendas reflected the political and administrative context in each State. Common to all of these efforts, however, were the new connections formed among State officials, policymakers, providers, provider associations, private-sector payers and insurance plans, patient representatives, staff of various State and Federal reform initiatives and demonstrations, and other key stakeholders.

In addition, States aligned their efforts with—and used their CHIPRA quality demonstration project experiences to directly inform—broader Federal and State health reform initiatives. For example, States

most commonly linked their efforts to existing statewide reform initiatives, particularly those related to PCMH implementation.

2. Of the project elements that were in place at the end of the fifth year of the demonstration, more than half were, or were highly likely to be, sustained after the grant period was over.

During the demonstration, States implemented projects that included multiple elements. For example, some State projects aimed to support PCMH transformation, and these projects typically included separate elements such as learning collaboratives, practice facilitation, financial and labor resources provided to participating practices, and health care training or certification programs. We defined each of these activities as a separate element, because some were sustained and others were not. Using this definition, States implemented 115 elements by the end of the grant program's fifth year. Our analysis of the sustainment of project elements yielded the following findings:

- Across all States, 57 percent of elements were or were highly likely to be sustained. The percentage of sustained elements varied by topic, with elements related to patient engagement being least likely to be sustained and elements related to practice facilitation and quality reporting being most likely to be sustained.
- Seventeen demonstration States implemented 40 elements used singly or in combination for service delivery transformation and sustained just over half of these elements. Some types of elements within this topic area were more likely to be sustained than others.
 - States sustained 77 percent of their facilitation programs, compared with 60 percent of their training and certification elements; 42 percent of their learning collaboratives; and 20 percent of their programs to provide payments to practices for participating in QI activities.
- Eight States developed strategies for reporting quality measures to CMS, and all of the States sustained or hoped to sustain those elements after the grant period. Consistent with our findings related to challenges in developing quality reports, States were somewhat less successful in sustaining program elements related to quality measure reports for stakeholders within the State or to payments and technical assistance to providers to produce or use reports on quality measures
- Twelve States implemented a diverse range of elements related to health IT that involved providing TA to improve data from EHRs, achieving data system interoperability, and establishing Web sites with information for providers or families; about half of these elements were sustained. Although demonstration States encountered challenges in health IT-related projects, the sustainment of nearly half of them implies States are committed to using health IT as a platform for improving quality of care generally.
- States planned to spread more than half of sustained elements following the demonstration. For elements related to service transformation, spreading the program elements typically involved increasing the number of practices that States were reaching through learning collaboratives or practice facilitation. States also spread concepts and approaches from the demonstration to QI programs in the adult health realm.
- States implemented about one-quarter of all sustained elements statewide as part of the demonstration and therefore had already maximized the spread of these elements. For example, one State developed and is highly likely to sustain a new administrative infrastructure to analyze

data from multiple child-serving agencies—an element that was designed to be spread statewide from its inception.

- Even though many States had contracted with evaluation teams to conduct various types of monitoring and evaluation studies, States reported few opportunities to make sustainment decisions based on empirical data.

3. Demonstration grants allowed States to gain substantial experience, knowledge, and partnerships related to QI for children in Medicaid and CHIP—a resource we refer to as “intellectual capital.”

Demonstration staff in all 18 States garnered a great deal of experience through partnerships with officials, providers, and quality specialists in their own and other States. The intellectual capital acquired during the demonstration will be sustained in varying forms in 13 States. For example:

- Six States will build on demonstration activities through new scope of work provisions in pre-existing contracts with State universities.
- In five States, key State staff either stayed in their positions or moved to other positions in the Medicaid agency and remained closely involved in QI activities. In contrast, key staff that provided leadership for the demonstration grant in five other States will not be supported after the grant period.
- New entities were developed in two States; one developed a new statewide partnership to continue QI activities for children; the other State will establish a new administrative unit within the Medicaid agency to support QI learning collaboratives and related initiatives begun under the demonstration grant.

3. Observations About the Structure of the Demonstration Grant Program

The national evaluation team has worked closely with AHRQ, CMS, and the demonstration States during the 5-year evaluation period. As a result, we have had many opportunities to observe and reflect on the design of the grant program itself. In this section, we discuss our observations about four of the program’s key characteristics:

- The grant program’s resources were spread across many discrete projects.
- Multistate partnerships heightened cross-State learning but posed administrative challenges for demonstration staff.
- The quality demonstration grant program did not explicitly encourage the development of payment models or other approaches to promote sustainability of successful projects after the grant period.
- Implementation of grantees projects was supported by several administrative structures, including full-time project directors, an initial planning period, and no-cost extensions.
- Several aspects the demonstration structure affected the likelihood of obtaining rigorous evaluation results from the beginning.

Allocation of Grant Program Resources

After Congress passed CHIPRA in February 2009, CMS developed the details of the CHIPRA quality demonstration grant program, with input from AHRQ, and issued the grant solicitation on September 30, 2009. Although constrained by four categories stipulated in the CHIPRA legislation, CMS was able to make several decisions that affected the scope of the grants. The first was to restrict applicants for the grants to State Medicaid agencies and not award grants directly to providers. The second was to create Category E, a broad addition to the Congressionally-mandated categories. The third decision was to allow applicants to apply for funding in more than one of the categories. The final decision was to encourage States to collaborate and submit multistate applications, which were permitted by the authorizing legislation. One result of these decisions was a large number of separate projects—52 overall.

Congress appropriated \$100 million for the CHIPRA quality demonstration grant program, a substantial Federal investment designed to learn about ways to improve quality of care for children. The value of the 10 grants ranged from approximately \$9.8 million to \$11.3 million over 5 years and supported from three to nine discrete projects (Table 3). Although projects were not the same size, it is instructive to calculate average per-project funding amounts. As shown in Table 3, the average amount available per project per year varied substantially across the grantees, depending on the number of partner states and the number of categories covered by each partner.

The figures in the last two columns in Table 3 should be interpreted as general indices of the average level of funds available, rather than precise amounts spent on any given project in a particular year. Moreover, grantees and their partner States established grant operations in very different ways, with varying degrees of subcontracting and in-kind contributions. Few, if any, of the States would be able to report dollars per project, because many individuals paid by grant dollars were working on multiple projects at any one time. In addition, most grantees and States requested a no-cost extension, meaning that their award was stretched beyond a 5-year period.

Table 3. CMS grant amounts received, number of projects, and average amount per project per year, 10 CHIPRA quality demonstration grantees

Grantee (Total # States)	Total Amount Received	Total Number of Projects ¹	Average Amount Per Project ²	Average Amount Per Project Per Year ³
Oregon (3)	11,277,361	9	1,253,040	250,608
Florida (2)	11,277,361	8	1,409,670	281,934
Maryland (3)	10,979,602	7	1,568,515	313,703
Utah (2)	10,277,361	6	1,712,894	342,579
Maine (2)	11,277,362	6	1,879,560	375,912
Colorado (2)	7,784,030	4	1,946,008	389,202
Massachusetts (1)	8,777,361	3	2,925,787	585,157
North Carolina (1)	9,277,361	3	3,092,454	618,491
South Carolina (1)	9,277,361	3	3,092,454	618,491
Pennsylvania (1)	9,777,361	3	3,259,120	651,824
Total	99,982,521	52	1,922,741	384,548

Source: Centers for Medicare & Medicaid Services (CMS).

Note: The figures in this table do not include in-kind contributions from the States or other Federal agencies, which in many cases were substantial.

¹ Number of discrete projects implemented by grantee and partners (see Table I).

² Total amount received divided by number of projects.

³ Average amount divided by 5. (We did not account for the no-cost extension period.) Amount reflects average level of funds available, rather than precise amount spent on any given project in a particular year.

Overall, the grant program's large number of projects had benefits and drawbacks. On one hand, the number and breadth of projects provided many opportunities to identify QI strategies across diverse topic areas. Involving a considerable number of States in a large number of projects may have attracted greater contributions by States, health plans, practices, and other funders, leveraging the Federal investment. On the other hand, the States were limited in the scope of certain projects because grant funding was spread thin across so many efforts. For their Category C projects, for example, most States engaged a relatively small number of sites in grant activities. Alaska engaged the fewest practices (three), and Illinois the most (about 25 practices signed up for learning collaboratives). Most others had between 10 and 18 practices. Not only did this limit the demonstration's potential to have a direct impact on a large number of children's lives, but the small number of sites also interfered with the ability to conduct rigorous evaluation, as noted in Section 4.

The purpose of the grant program was to "evaluate promising ideas for improving the quality of children's health care provided under [Medicaid and CHIP]." ²¹ States varied in how they pursued promising ideas, which had implications for how grant funds were spent. For some States, this meant demonstrating proof of concept. Alaska, for example, used grant dollars to explore and operationalize the concept of a medical home in a frontier environment. For other States, it meant implementing a pilot study focused in a few locations, with the potential to spread the intervention if the pilot study were successful. For example, Utah used grant funds to support care coordinators in 12 practices; after the grant funding ended, it used another source of funds to spread the use of coordinators to other practices.

Still other States pursued promising ideas by building on previous efforts. The Maryland team, for example, used grant funds to strategically explore avenues for supporting CMEs. Its eventual pursuit of a Medicaid waiver opened a new funding stream that could serve many more children. Another example is Vermont, whose CHIPRA team used demonstration funds to accelerate the timeline for implementing an ongoing statewide initiative (*Blueprint for Health*) with pediatric practices.

The abundance of projects allowed many efforts to move forward simultaneously in the demonstration States. But demonstration projects may have suffered from being under-funded, making them poorer tests of the promising ideas they explored. Furthermore, the diversity and multitude of projects made it more difficult to summarize the demonstration's lessons for policy and program administrators. A more focused grant program could have produced more definitive results on fewer topics, rather than drawing more limited conclusions across more topics.

Multistate Partnerships

As noted above, six of the quality demonstration awards involved multistate partnerships (see Table 1). States in these partnerships were committed to learning from and sharing ideas with each other. In all cases, the States allocated time and resources to support these partnerships, although the methods and amount of resources varied. Two of the six grantees (Illinois/Florida and Maryland/Georgia/Wyoming) hired independent organizations to convene the partners and foster cross-State learning.

As described in detail in our sixth *Evaluation Highlight*, these partnerships had significant benefits and challenges. Several States collaborated closely with their partners by developing joint projects, integrating activities, and setting up complementary implementation schedules. States shared information through activities such as visiting each other's sites, trading key materials and reports, and

scheduling regular teleconferences or in-person meetings. Generally speaking, States found that they offered each other complementary, rather than redundant, skills and expertise.

Interviews with staff and presentations made during several monthly grantee calls hosted by CMS noted the following benefits of these partnerships:

- Fairly rapid and easy dissemination of information about tools, training resources, and other QI initiatives across partner States, thereby filling gaps in expertise and capacity.
- An opportunity to learn the operational details needed to implement a particular strategy from more experienced State staff or consultants, thus potentially avoiding some mistakes.
- Opportunities to expand the spread and potential impact of a project across States.

Staff in most States felt the benefits of partnering outweighed the costs, but also noted the following challenges:

- Working together is both time- and labor-intensive. States reported that project activities took longer to implement than they might have if a State were “going it alone,” especially with regard to financing project work across States, reporting, and decisionmaking.
- Establishing and maintaining contracts and agreements between State governments can result in implementation delays.

Payment and Other Approaches to Sustainability

States tested models for improving child health care delivery, but most did not establish associated payment mechanisms to sustain these models after the grant ended. For example, some States used grant funds to offer payments to practices for participating in QI collaboratives or to provide stipends or salaries for care coordination, but they did not establish ongoing financing approaches, such as care coordination as a Medicaid billable service. As a result, most States did not have the administrative infrastructure or alternative source of revenue in place at the end of the grant to institutionalize incentives for practices to continue QI activities. Notable exceptions to this general observation include Pennsylvania’s continuation of its pay-for-reporting and pay-for-improvement program, South Carolina’s creation of a new children’s health care quality office in its Medicaid agency, and Maryland’s Medicaid State Plan Amendment that provides a funding stream to support CMEs.

Efforts to transform the delivery system are unlikely to be successful unless new payment models emerge to support them. To help promote sustainability of successful interventions, CMS and other funders could consider requiring efforts at payment reform or other sustainability planning to be explicit parts of projects through the application, operational planning, and implementation stages.

Grant Administration and Planning

CMS required grantees to ensure that project directors were available full time for the grant activities. As a result, the 10 project directors were well informed about the operational activities that the States and their partners were implementing through the grant. As was frequently evident on CMS’ all-grantee calls, this allowed CMS to build a community of individuals consistently engaged around and

knowledgeable about the goals of the demonstration. One potential drawback to full-time project directors became apparent toward the end of the grant period when project directors sought other positions in anticipation of the grant's termination. In some cases, the project directors moved to other positions in the State or partnering organizations, and it was difficult to maintain contact with these individuals. Not unexpectedly, some individuals who stepped in to serve as project directors during the grant's last phase often lacked historical knowledge of grant activities. This could be addressed in future grant programs by providing education to grantees on planning for leadership succession and management approaches to maintain institutional knowledge.

CMS required each State to submit an operational plan, which was due approximately 10 months after the grant award. Key stakeholders in some States noted that this planning period substantially helped subsequent program implementation by better aligning grant activities with what was considered feasible. The period between grant award and submission of the plan allowed the States to refine their proposed plans in response to many factors that were likely to have evolved significantly from the original grant application period. As a result, in certain key respects, some States' operational plans differed significantly from their applications. In some cases, States realized during this planning period that certain projects proposed in their applications (including some of the health IT-related efforts) could not be practically implemented and therefore shifted funds to other grant efforts.

CMS granted no-cost extensions (NCEs) ranging from 3 to 12 months to nine grantees, and as a result, 15 States continued to operate some aspects of their projects beyond the original end date of February 22, 2015. States used their remaining funds to continue selected program elements, such as quality measure reporting or statewide partnerships, or to complete their own evaluation reports. This extension period also allowed us to gather information about program sustainment that otherwise might have been difficult to collect because key staff would have been hard to contact. Because most States' NCEs extended beyond the end of this evaluation contract, we were unable to fully assess the influence of NCEs on demonstration activities and sustainability. Future demonstration programs could provide clear and early guidance to participants on whether NCEs might be available and how and when decisions about NCEs might be made.

Grant Structure and Rigorous Evaluation

Because of the emphasis on learning from the demonstration grants, CMS made two important decisions to address the program's goals for evaluation. First, CMS elected to fund and have AHRQ lead an evaluation of the entire grant program. CMS required grantees to work collaboratively with the national evaluation team contracted by AHRQ and provide access to program data and staff. Second, CMS let States conduct their own independent evaluations using grant funds as long as they were not duplicative of the national evaluation. CMS did not, however, require that States develop or support rigorous approaches to evaluating the impact of their programs, such as having comparison groups to control for non-demonstration influences on demonstration sites. (See section 4 for further discussion of the implications of this limitation.) Furthermore, many projects were underway and intervention sites had been selected and enrolled before the evaluation contract was awarded, limiting the ability to make changes to support rigorous evaluation.

Many applicants to the program requested grant funds for independent evaluations. When grant awards were less than the amount grantees had applied for, some grantees cut back evaluation budgets. This may have contributed to the need for evaluation-focused TA that the national evaluation team provided

to the States and their independent evaluators. (See section 4 for more on evaluation-focused TA.) It also may have contributed to the level of cooperation States could give to the national evaluation. For example, when the NET attempted to get claims data from States, there were often delays because State programming resources were scarce. Once the data were obtained, often after multiple attempts to get the requested format, the data required extensive cleaning. (See Appendix C for further details.) To minimize such problems, future grant programs could earmark grantee funds for conducting or cooperating with evaluation activities.

We believe that future grant programs with similar goals should include review criteria in the grant solicitation on how well applicants demonstrate that their proposed projects could be rigorously evaluated. Proposals that do not meet minimum standards would not be eligible for funding. Additionally, CMS and other funders could consider including standards for evaluability as part of the process for approving operational plans.

Future grant programs could also provide evaluation-focused TA from the beginning of the grant to increase the opportunities for rigorous evaluation and help build relationships between the grantees and evaluation team. This strategy could be further supported by a requirement for at least quarterly communication between State-based and national evaluation teams to encourage them to develop evaluation plans that complement and build on each other.

4. Observations About the Evaluation

Like many complex evaluations of grant programs, the 5-year national evaluation of the CHIPRA quality demonstration grant program faced key challenges as the NET worked on a variety of tasks and produced numerous and diversified products. Our five interim reports for AHRQ (submitted in final form in August 2011, August 2012, February 2013, October 2014, and May 2015) detail the tasks we undertook during the evaluation, the challenges we faced, and the solutions we devised to address them. Here we discuss four overall conclusions regarding the evaluation itself:

- The national evaluation accomplished many of its goals, but it did not include impact analyses for demonstration projects because of challenges related to program design, program implementation, and data availability.
- We developed diverse methods for collaborating with grantees, such as providing evaluation-focused TA.
- Our technical expert panel was helpful in the early stages of the evaluation, but need for their input diminished once the evaluation design solidified.
- We developed and disseminated evaluation findings throughout the evaluation period, emphasizing emerging lessons learned about program implementation at first and synthesizing findings about program outcomes and effects in the last months of the project.

The Challenge of Impact Analyses

In addition to listing more than 100 potential questions that the evaluation could address, AHRQ's request for task order (RFTO) noted that the evaluation's purpose was to provide CMS and States with: (1) "insight into how best to implement quality improvement programs" for children and (2) "information on how successful programs can be replicated." As noted above, the NET has generated a large number of products that provide insights into strategies for improving quality of care, suggesting that the national evaluation accomplished the first goal. These products also address a large proportion of AHRQ's original research questions (www.ahrq.gov/policymakers/chipra/demoeval/index/html).

We did, however, face significant challenges in reaching the second goal—determining the success of the demonstration projects based on quantitative measures of care derived from claims or other types of quantifiable data. Typically, evaluations use rigorous research designs to estimate the impact of programs on designated outcomes. Strong research designs include randomized controlled trials or comparison group designs that draw on data collected before and after program initiation, from both the group receiving the intervention and a comparison group that is similar in characteristics but is not involved in the intervention. These designs are considered strong because they provide evidence about what would have happened in the absence of the intervention. Comparing outcomes for the group affected by the intervention with outcomes for the comparison group allow one to estimate the impact of the intervention beyond what would have happened anyway.

The fact that States were not required to have comparison groups as a condition of their grant impeded the use of these rigorous methods. Furthermore, many projects were underway, and intervention sites often selected and enrolled before the evaluation contract was awarded. Therefore, we worked actively throughout the evaluation period to determine opportunities to work with States and implement comparison group designs for at least one project in each demonstration State. These efforts included:

- Asking States to use their grant funds to identify and obtain data from comparison group practices as part of their Category C projects.
- Asking Pennsylvania to use a lagged implementation approach for its Category B work, so practices that implemented electronic screeners in later years of the grant could be used as comparison practices in earlier years.
- Requesting that States with projects designed to enhance medical home features use a standard measure of medical homeness so that we could combine data or compare outcomes across States.
- Working with States to ensure that we had the quantitative information necessary to develop claims-based measures of service use and to attribute children to specific intervention and comparison group practices.
- Providing evaluation-focused TA to States to ensure that they gathered the data needed for quantitative analyses.

Although we examined each project to determine whether impact analyses would be feasible, we focused on Category C projects because they appeared particularly conducive to a rigorous impact evaluation. In particular, 12 States planned to implement a PCMH model to improve quality of care for children in selected practices. As described in our first evaluation design report, we planned to collect Medicaid administrative data and practice-reported PCMH surveys from the CHIPRA intervention practices and a set of comparison practices to assess whether outcomes (such as receipt of well child care and avoidable emergency department visits) improved more among children in the intervention

versus comparison practices. Moreover, by combining data across States with similar interventions, we expected to have enough statistical power to detect project impacts on children's health care.

Unfortunately, for numerous reasons, we could not conduct these analyses as planned. (Appendix C provides a detailed account of the problems we encountered.) In fact, as we worked with each State during the evaluation period, we encountered obstacles beyond our control that made it impossible to implement our plans for quantitative analyses:

- The number of intervention practices in some States' Category C projects was so small that the chance of detecting differences in service use for children in these practices and children in comparison practices did not warrant the substantial investment of resources required to conduct impact estimates. For example, Alaska and Idaho each worked with only three practice sites in their Category C projects.
- The quality and comprehensiveness of the Medicaid administrative data were compromised by lack of encounter data from managed care organizations. Many of the Category C States have high use of Medicaid managed care among child beneficiaries, with 45 to 90 percent of children in managed care in Florida, Massachusetts, Oregon, South Carolina, Utah, and West Virginia. Most states could not provide managed care encounter data. Without data from managed care organizations, our evaluations would have represented a small proportion of intervention children in those States.
- States altered original plans for their interventions so substantially that the project's actual implementation was far less likely to achieve the effects originally intended. For example, several States focused on a narrow range of PCMH transformation activities, rather than implementation of the full model as they had originally planned.
- States initially agreed to identify and collect data from comparison groups but then did not do so because they did not want to impose data collection burdens on practices without providing some benefits, the costs of which had not been included in the grant's budget.
- Because of the selection process for identifying intervention sites, developing an equivalent group of comparison sites was not feasible, especially in the less populous States (for example, the intervention sites were the largest and most sophisticated in the State).
- Post-intervention data were unavailable from CMS' data files because of major lags in the data submitted by States and because of major delays as CMS transformed its data file structure from one system (MSIS) to another (T-MSIS).

In addition to these obstacles, many of the demonstration projects were designed to enhance the State's infrastructure for QI activities—as requested in the original solicitation. Infrastructure programs are typically designed to affect children statewide; as a result, there are no "intervention" or "comparison" groups. For example, Massachusetts sought to assemble a group of stakeholders, The Massachusetts Child Health Care Quality Coalition, to develop a shared understanding of child health care quality priorities, create a platform for formulating systemwide goals and objectives, and implement activities to support those goals.

Moreover, evaluating infrastructure programs requires a substantial period of time. Following implementation (which may require several years of planning and activity), the effects of such programs on beneficiaries' service use are likely to be measurable only after a substantial amount of time has passed. For example, Wyoming spent more than 3 years of the grant designing and developing the administrative infrastructure for the State's first care management entity to improve care for children

with severe behavioral health care needs; then, the State piloted the program with around 150 youth in the final years of the demonstration.

Our inability to complete impact analyses for at least one demonstration project in at least one State was a major disappointment. Nonetheless, our efforts to do so led to three positive developments.

First, because of our early work with Massachusetts to support enrollment of comparison practices into their Category C program activities, the State was in a position to conduct its own impact analyses using its Medicaid claims and managed care encounter data (whereas we were able to use only its fee-for-services claims).²² At the State's request, we provided TA for these analyses. Consequently, we anticipate that the Massachusetts team will complete a manuscript describing the impact of their Category C program and will likely submit the manuscript to a peer-reviewed journal in late 2015. Their unpublished findings show that children with chronic conditions attributed to CHIPRA practices for the full 3-year demonstration had a significant reduction in potentially avoidable emergency department use, whereas comparison children with chronic conditions had no such reduction over the same time period.

Second, we reassigned resources originally allocated for impact analyses to a quantitative survey of physicians in three States to ensure that we could address questions related to provider perceptions of QI efforts based on practice-level quality reports—an issue that is directly related to many of the demonstration States' Category A and Category C projects. One of our journal manuscripts (submitted for publication) presents quantitative findings based on analysis of data from this survey.

Third, we used the baseline claims data received from three States to conduct an analysis of the association between a practice's degree of medical homeness and health care utilization of child Medicaid beneficiaries in those practices. Although this work is not technically an evaluation of the CHIPRA demonstration activities, the publication of our analysis²³ in a peer-reviewed journal contributed to the field's limited knowledge of the effect of PCMH for children, using data that were already collected and cleaned in preparation for the planned impact evaluation.

Although many of the challenges we encountered could not have been foreseen, we believe that future grant programs could avoid some of these problems by adhering to recommendations made in Section 3 of this report.

Collaboration with Grantees

From the beginning of the evaluation, the NET worked carefully to develop productive working relationships with the demonstration States and engage them in our work. We remained mindful of the need to avoid imposing unnecessary burdens on the States and the importance of acknowledging the value of their experiences and perspectives. We also provided States with an opportunity to comment on our products, and make factual corrections as needed. We held Webinars and conference calls to discuss overarching issues and specific content. As noted above, the presence of full-time project directors seemed to be more conducive to external evaluation activities compared with our experiences on other similar large demonstration projects.

We also offered States evaluation-focused TA. The need for this kind of TA emerged in the first 6 months of the evaluation in response to our recognition that the States had not proposed any mechanism for gathering counterfactual information to support rigorous evaluation for the majority of the 52 projects.

In the first 12 months of the project, we strongly urged States to identify comparison practices and to administer measures of “medical homeness” to both comparison and intervention groups. (See the “challenge” section above for further discussion of this issue).

Our TA took different forms at different stages of the evaluation. In the first year of the evaluation, we helped selected States consider comparison groups for their Category C interventions, to make their projects more conducive to a rigorous evaluation. In the second year of the evaluation, we established periodic calls with all demonstration States to address issues in measuring “medical homeness.” During these calls, we provided overviews of different measurement frameworks, discussed the strengths and weaknesses of each for application in the CHIPRA demonstration projects, and answered questions from the States. In addition, we participated in several calls with the Children’s Hospital of Philadelphia (CHOP) Policy Lab to collaborate on designing an evaluation they planned to conduct of the effect of the CHIPRA developmental screening intervention on children’s receipt of early intervention services. In years three and four of the evaluation, we held a series of calls related to measuring outcomes using claims data. In the last year of the evaluation, we offered to provide assistance to States in developing technical reports related to their own State-based evaluations. For example, as described above, we worked closely with staff in Massachusetts to help them develop an impact analysis of their PCMH intervention using Medicaid claims and managed care encounter data and to present their findings in a manuscript for submission to a peer-reviewed journal.

Technical Expert Panel

We met with our 14-member Technical Expert Panel (TEP) in person in the fourth month of the evaluation and presented our overall plan for conducting the evaluation. The TEP concurred with our approach and also offered helpful suggestions for refining our methodology.

We used subsequent meetings (held by telephone beginning at the midpoint of our second year) to help prioritize the long list of research questions that AHRQ had originally posed for the evaluation. Through these deliberations, we recognized that we would not be able to address all the questions in a comprehensive manner. With the TEP’s assistance, we were able to prioritize the most important questions, which allowed us to focus our resources productively. Subgroups of the TEP also provided input on specific topics, such as the content of the physician survey.

As we moved past the design and prioritization phases of the evaluation, we realized that TEP meetings would be less useful over time, because we would be asking the TEP members to read and comment on only the products we had committed ourselves to developing. In conjunction with our AHRQ project officers, we decided to use the remaining national evaluation funds that had been allocated to run the TEP to support the writing of *Evaluation Highlights* and other evaluation products.

Development and Dissemination of Evaluation Findings

AHRQ provided consistent encouragement to the NET to develop—as soon as possible and throughout the evaluation period—products with findings that would be of use to States, in particular, and also to CMS and the field of child health care in general. In line with this emphasis, we focused on several methods for disseminating our products.²⁴ Specifically, we took the following steps:

- In August 2012, we launched the national evaluation Web page, hosted on AHRQ's Web site. Initially, we used the Web page as a venue for educating stakeholders about the demonstration and our evaluation. As the evaluation progressed, we posted our products on this page and posted links to State-generated reports as they became available. By the end of the national evaluation in September 2015, more than 8,500 individuals had become subscribers to the CHIPRA national evaluation updates. AHRQ used the GovDelivery platform, along with its Electronic Newsletter, Child and Adolescent Health Periodic Digest, and Twitter feed to inform subscribers and others when new information was posted to the national evaluation Web site.
- We developed dissemination partners to help broaden the reach of our findings. In addition to the States themselves, we worked closely with the Maternal and Child Health Bureau (MCHB), the Association of Maternal and Child Health Programs (AMCHP), the Catalyst Center, the National Association of Medicaid Directors (NAMD), the National Academy for State Health Policy (NASHP), the Children's Hospital Association (CHA), the American Academy of Pediatrics (AAP), the American Academy of Family Physicians (AAFP), Voices for America's Children, and the National Initiative for Children's Healthcare Quality (NICHQ). Some of these organizations (AMCHP, the Catalyst Center, and NASHP, for example) helped us in our dissemination efforts by including announcements of our products in their newsletters and through other means. Other organizations (the AAP and the AAFP, for example) were less interested in helping with dissemination.
- We presented findings to the demonstration States during several CMS-hosted conference calls, at CMS-sponsored quality conferences, and at various professional conferences (including several of AcademyHealth's annual conferences and at its first National Child Health Policy Conference).
- During the last 3 months of the projects, we helped organize several Webinars in conjunction with key dissemination partners. At the time of writing this report, we had Webinars scheduled with (1) NASHP to present to their CHIP directors and Children in the Vanguard learning networks, (2) the State-University Partnership Learning Network hosted by AcademyHealth, (3) the National Improvement Partnership Network led by the University of Vermont, and (4) the Association of Medicaid Medical Directors.

Overall, the Web page on AHRQ's Web site provided a sturdy platform for making available to interested individuals both the products developed by the NET and links to State reports. The number of subscribers increased steadily during the evaluation period. The visits to and downloads of our products typically peaked in the month of their publication and then waned. Introduction of a new product often produced some traffic to earlier publications.

The major challenge we faced with our dissemination work was the short period of time between completing our final analyses (July 2015) and the end of the contract (September 8, 2015). We were unable to develop Webinars with key dissemination partners until we had a reasonably clear idea of our results. As our findings emerged from our analyses during the spring of 2015, we began reaching out to our partners. In most cases, they indicated that they would be willing to collaborate on Webinars during the fall, rather than during the summer. Hence, we worked to plan the Webinars and develop the necessary slides and materials during the contract period.

5. Conclusion and Summary

The CHIPRA quality demonstration grant program was an ambitious Federal effort to evaluate promising strategies for improving quality of care for children enrolled in Medicaid and CHIP. States implemented a wide array of projects that provided examples of such strategies, many of which will be sustained and spread after the demonstration has ended. These projects underscore the importance of marshaling resources over several years to enhance the capacity of States to report and use quality measures, address the thorny problems of implementing new health IT applications, and develop the stakeholder relationships that underpin successful efforts to transform service delivery systems. Challenges in assessing the impact of these projects emphasize the need to both embed evaluation considerations in designing grant programs and enhance access to the administrative and claims data needed to assess quality of care for populations enrolled in Medicaid and CHIP.

Overall, findings from the evaluation of the grant program provide policymakers at the Federal and State levels with a strong foundation for considering next steps to improve quality of care for children enrolled in Medicaid and CHIP. For example, CMS could build on demonstration States' successful experiences in developing their capacity for reporting and using the Child Core Set of quality measures, possibly by supporting other States in replicating these capacity-building strategies or incorporating lessons learned in future TA efforts. As Federal and State policymakers develop new efforts to stimulate innovation in service delivery systems, they could look to the outcomes of this demonstration for ideas about pathways to further explore (and to avoid). In sum, results from the demonstration grant program and its national evaluation suggest numerous strategies that can inform future policy development and new grant-making programs to improve care for children.

6. References and Notes

¹ We use the term “national evaluation” to distinguish our work from the activities undertaken by evaluators who are under contract with many of the demonstration grantees to assess the implementation and outcomes of State-level projects. The word “national” should not be interpreted to mean that our findings are representative of the United States as a whole.

² A detailed description of our evaluation goals and methods can be found in the design plan submitted to AHRQ in April 2014 available at <http://www.ahrq.gov/sites/default/files/wysiwyg/policymakers/chipra/demoeval/what-we-learned/finaldesignplan.pdf>.

³ For a complete listing of these products and their relevance to the research questions, please send an email request to CHIPRADemoEval@ahrq.hhs.gov.

⁴ The 10 grantee States are Alaska, Florida, Illinois, Maine, Massachusetts, North Carolina, Oregon, Pennsylvania, South Carolina, and West Virginia.

⁵ Burwell SM. 2014 annual report on the quality of care for children in Medicaid and CHIP. November 2014. Available at <http://www.medicaid.gov/medicaid-chip-program-information/by-topics/quality-of-care/downloads/2014-child-sec-rept.pdf>. Accessed August 7, 2015.

⁶ CMS. Electronic specifications for clinical quality measures. Available at: http://www.cms.gov/Regulations-and-Guidance/Legislation/EHRIncentivePrograms/Electronic_Reporting_Spec.html. Accessed on June 3, 2015.

⁷ CMS. Guide for reading eligible professional (EP) and eligible hospital (EH) eMeasures, version 4. May 2013. Available at: https://www.cms.gov/Regulations-and-Guidance/Legislation/EHRIncentivePrograms/Downloads/Guide_Reading_EP_Hospital_eCOMs.pdf. Accessed on September 2, 2015.

⁸ Colorado and New Mexico implemented the electronic screening questionnaire as a Category E project. We include them in this section because their screening program is conceptually similar to the Pennsylvania Category B project; both projects are health IT applications.

⁹ See Electronic Student Health Questionnaire (eSHQ) enhances risk assessment for adolescents (available at <http://www.ahrq.gov/policymakers/chipra/demoeval/what-we-learned/co-nm-specialinnovation.html>) and Introducing electronic screening tools for developmental delay and autism into pediatric primary care (available at <http://www.ahrq.gov/policymakers/chipra/demoeval/what-we-learned/pa-specialinnovation.html>).

¹⁰ The 12 States are Oregon, Alaska, West Virginia, Utah, Idaho, Florida, Illinois, Maine, Vermont, Massachusetts, South Carolina, and North Carolina.

¹¹ Illinois, Massachusetts, Maine, North Carolina, South Carolina, and West Virginia collected medical home survey data from more than 80 child-serving “comparison” practices that did not participate in CHIPRA practice transformation activities.

¹² The 12 States were Alaska, Florida, Idaho, Illinois, Maine, Massachusetts, North Carolina, Oregon, South Carolina, Utah, Vermont, and West Virginia.

¹³ Information on the medical home assessment method used by Vermont is not available.

¹⁴ The MHI-RSF instrument can be found at: <http://www.ahrq.gov/policymakers/chipra/demoeval/resources/mhirsf.html>.

¹⁵ The Abridged Children’s EHR Format can be found at https://ushik.ahrq.gov/mdr/lists/administeredItems/Requirements?filterColumn_8=yes&system=cehrf&enableAsyncronousLoading=true.

¹⁶ Oshiro BT, Kowalewski L, Sappenfield W. A multistate quality improvement program to decrease elective deliveries before 39 weeks of gestation. *Obstet Gynecol* 2013 May;121(5):1025-31.

¹⁷ See CHQC Massachusetts at <http://www.masschildhealthquality.org/>.

¹⁸ The goals of State-level improvement partnerships and NIPN are to facilitate collaboration and the translation of knowledge across programs, so that States can learn from each other about strategies that work (and do not work) to improve quality of care for children under Medicaid and CHIP.

¹⁹ *Evaluation Highlight 4* focuses on how the demonstration helped to elevate children on State health policy agendas.

²⁰ *Evaluation Highlight 6* addresses the issue of partnerships among the States in multistate grants. We review findings from this *Evaluation Highlight* in Section 3 because this topic pertains to the structure of the grant program.

²¹ Children’s Health Insurance Program Reauthorization Act of 2009, P.L. 111-3.

²² During our baseline analysis, Massachusetts was unable to link managed care provider identification numbers to participating CHIPRA practices, a connection that is necessary for us to be able to attribute children to intervention and comparison practices. So while they sent us managed care encounter data, we could not identify which managed care patients were cared for in CHIPRA practices. By the time we determined Massachusetts had solved the linkage problem and could use managed care data, they were already in the process of conducting their own evaluation.

²³ Christensen AL, Zickafoose JS, Natzke B, et al. Associations between practice-reported medical homeness and health care utilization among publicly insured children. *Acad Pediatr* 2015 May-June;15(3):267-74. Available at [http://www.academicpedsjnl.net/article/S1876-2859\(14\)00429-X/abstract](http://www.academicpedsjnl.net/article/S1876-2859(14)00429-X/abstract). Accessed October 27, 2015.

²⁴ Details regarding our dissemination methods are available in the original dissemination plan we submitted to AHRQ in August 2011 and the updates submitted in May 2012, August 2013, and May 2014. Available on request; send an email to CHIPRADemoEval@ahrq.hhs.gov.

Appendix A. National Evaluation Team and Technical Expert Panel Members

National Evaluation Team Members

Mathematica Policy Research	Urban Institute	AcademyHealth	CMS	AHRQ
Grace Anglin, M.P.H.	Fiona Adams	Ellen Albritton	Barbara Dailey, M.S.	Linda Bergofsky, M.B.A., M.S.W.
Kirsten Barrett, Ph.D.	Rachel Burton, M.P.P.	Margo Edmunds, Ph.D.	David Greenberg, M.B.A.	Cindy Brach, M.P.P.
Veronica Barrios-Monroy	Nicole Cafarella Lallemand, M.P.P.	Gerry Fairbrother, Ph.D.	Elizabeth Hill, M.P.H.	Tennille Brown, M.A.
Joseph Brown	Kelly Devers, Ph.D.	Lisa Simpson, M.B., B.Ch., M.P.H.	Karen LLanos, M.B.A.	Stacy Farr, Ph.D.
Steve Bruns	Vanessa Forsberg, M.P.P.	Veronica Thomas		
Anna Christensen, Ph.D.	Elaine Grimm, M.A.			
Sharon Clark	Emily Hayes			
Jared Coopersmith, Ph.D.	Ian Hill, M.S.W., M.P.A.			
Cheryl DeSaw	Embry Howell, Ph.D.			
Wesley Dunlap	Michael Huntress, M.S.			
Leslie Foster, M.P.A.	Margarita Hurtado, Ph.D.*			
Alicia Haelen Demers, M.A.	Genevieve Kenney, Ph.D.			
John Hall, M.S.	Ellen Kurtzman, M.P.H., R.N.			
Tricia Higgins, Ph.D.	Emily Lawton			
Mynti Hossain, M.P.P.	Juliana Macri			
Henry Ireys, Ph.D.	Stacey McMorro, Ph.D.			
Cathy Lu, M.S.	Amanda Napoles			
Daryl Martin	Ashley Palmer, M.P.P.			
Catherine McLaughlin, Ph.D.	Rebecca Peters, M.P.H.			
Amanda Mims, M.S.	Amy Phillips			
Brenda Natske, M.P.P.	Christal Ramos, M.P.H.			
Dana Petersen, Ph.D.	Elizabeth Richardson, M.S.			
Frank Potter, Ph.D.	Arnav Shah			
Amanda Reiter, Ph.D.	Adam Weiss			
Betsy Santos, Ph.D.	Margaret Wilkinson			
Lisa Schottenfeld, M.P.H., M.S.W.	Elena Zarabozo			
John Schurrer, M.P.P.				
Lauren Smith, M.P.P.				
Adam Swinburn, M.P.P.				
Frank Yoon, Ph.D.				
Hanzhi Zhou, Ph.D.				
Joseph Zickafoose, M.D., M.S.				

*External consultant

Technical Expert Panel Members

Bruce Bagley, M.D.	Janet Marchibroda, M.B.A.
Stephen Blumberg, Ph.D.	Cynthia Minkovitz, M.D., M.P.P.
Carl Cooley, M.D.	Jeff Schiff, M.D., M.B.A.
Sharron L. Docherty, Ph.D., C.P.N.P.	Sarah Scholle, M.P.H., Dr.P.H.
Elena Fuentes-Afflick, M.D., M.P.H.	Xavier Sevilla, M.D., M.B.A.
Bruce Lesley	Carol Tobias, Ph.D.
Rita Mangione-Smith, M.D., M.P.H.	Mark Weissman, M.D.

Appendix B. Products Produced By the National Evaluation Team

Products produced or initiated, featured States, data sources, and month of posting or publication, as of June 2015

Title	Featured States	Grant Category Examined	Primary Data Source	Month Posted or Published
Evaluation Highlights				
1. How are CHIPRA demonstration States approaching practice-level quality measurement and what are they learning?	Maine, Massachusetts, North Carolina, Pennsylvania	Category A	Interviews	January 2013
2. How are States and evaluators measuring medical homeness in the CHIPRA Quality Demonstration Grant Program?	Massachusetts, North Carolina, South Carolina, West Virginia, Alaska, Oregon	Category C	Interviews, Practices' responses on the Medical Home Index-Revised Short Form (MHI-RSF)	May 2013
3. How are CHIPRA quality demonstration States working to improve adolescent health care?	Colorado, New Mexico, North Carolina, Utah	Category C	Interviews	August 2013
4. How the CHIPRA quality demonstration elevated children on State health policy agendas	Maine, Maryland, Massachusetts, Vermont, Oregon	Cross-cutting	Interviews	October 2013
5. How are CHIPRA demonstration States encouraging health care providers to put quality measures to work?	Pennsylvania, South Carolina	Category A	Interviews	October 2013
6. How are CHIPRA quality demonstration States working together to improve the quality of health care for children?	Alaska, Colorado, Florida, Georgia, Idaho, Illinois, Maine, Maryland, New Mexico, Oregon, Utah, Vermont, West Virginia, Wyoming	Cross-cutting	Interviews	January 2014
7. How are CHIPRA quality demonstration States designing and implementing caregiver support programs?	Georgia, Idaho, Maryland, Utah	Category C	Interviews	February 2014
8. CHIPRA quality demonstration States help school-based health centers strengthen their medical home features	Colorado, New Mexico	Category C	Interviews	May 2014
9. How are CHIPRA quality demonstration States supporting the	Alaska, Idaho, Massachusetts, Oregon,	Category C	Interviews	July 2014

Title	Featured States	Grant Category Examined	Primary Data Source	Month Posted or Published
use of care coordinators?	Utah, West Virginia			
10. How are CHIPRA demonstration States testing the children's electronic health record format?	North Carolina, Pennsylvania	Category D	Interviews	August 2014
11. How are CHIPRA quality demonstration States using quality reports to drive health care improvements for children?	Alaska, Florida, Illinois, Maine, Massachusetts, North Carolina	Category A	Interviews	April 2015
12. How are CHIPRA quality demonstration States improving perinatal care?	Florida, Illinois	Category E	Interviews	May 2015
13. How did CHIPRA quality demonstration States employ learning collaboratives to improve children's health care quality?	Alaska, Florida, Idaho, Maine, Massachusetts, North Carolina, Oregon, Utah, West Virginia	Category C	Interviews	June 2015
Implementation guides				
1. Engaging stakeholders to improve the quality of children's health care	Georgia, Idaho, Massachusetts	Category E	Interviews	July 2014
2. Designing care management entities for youth with complex behavioral health needs	Georgia, Maryland, Wyoming	Category C	Interviews	September 2014
Manuscripts				
1. Nine States' use of collaboratives to improve children's health care quality in Medicaid and CHIP	Florida, Illinois, Maine, Massachusetts, North Carolina, Oregon, South Carolina, Utah, West Virginia	Category C	Interviews	November 2013
2. Associations between medical homes and health care utilization among publicly insured children	Illinois, North Carolina, South Carolina	Category C	Medicaid fee-for-service claims data	May 2015
3. What factors influence the ability of State Medicaid agencies to report the Child Core Set of health care quality measures? A multicase study	Illinois, Maine, Oregon, Pennsylvania	Category A	Interviews, DHHS report	Under review
4. Primary care physicians' experiences with and attitudes toward pediatric quality reporting	North Carolina, Pennsylvania, Ohio ¹	Category A	Interviews, physicians' responses to survey	Under review
5. After the demonstration: what States sustain after the end of Federal grants to improve children's health care	All 18 States	Cross-cutting	Interviews	Under review

Title	Featured States	Grant Category Examined	Primary Data Source	Month Posted or Published
quality				
6. Parent experiences in child-serving patient-centered medical homes in the CHIPRA quality demonstration	Florida, Oregon, South Carolina, Utah	Category C	Interviews	Under review
Special features				
Introducing electronic screening tools for developmental delay and autism into pediatric primary care	Pennsylvania	Category B	State-provided data	August 2015
The electronic Student Health Questionnaire (eSHQ) enhances risk assessment for adolescents	Colorado, New Mexico	Category B	State-provided data	August 2015

Notes: For all products, we consulted relevant State reports and contacted State officials as needed for clarification and fact-checking.

¹ Ohio was included as a comparison State and did not participate in any demonstration activities.

Appendix C. Obstacles to Impact Analysis of Category C Projects

As noted in Section 4, we intended to conduct impact analyses of Category C projects. Of the 17 States participating in Category C, 12 focused on PCMH models, three were focused on finding new strategies for funding CMEs or developing new ones (Maryland, Georgia, Wyoming), and two were focused on SBHCs (Colorado, New Mexico). We wanted to ensure that we had the potential to combine data across the States to maximize the chances of having sufficient analytic power to estimate impact. Hence, we excluded the five States not working to implement PCMH models because those models and the expected effects were sufficiently different from the PCMH approach that combining them would not have been conceptually plausible.

CMS did not require or encourage States to develop comparison groups for any of their CHIPRA projects, although some were planning to do so. By the time the national evaluation had developed a foundation of knowledge about the State projects (in December 2010, about 10 months after the start of the States' demonstrations), the States were completing their planning process and, in some cases, were well underway with implementation—too late for major design changes that might have supported a comparison group and a more rigorous evaluation. However, the NET was able to work with some of the PCMH States to develop a comparison group when they otherwise would not have done so.

Of the 12 PCMH States, we excluded from our planned analysis the States that did not agree to recruit comparison practices (Alaska, Florida, Idaho) or were unable to identify practices that were not participating in other PCMH initiatives in the State (Oregon, Vermont). Of the seven States that agreed to recruit comparison practices, Utah and West Virginia were unable to provide the Medicaid administrative data necessary for the analysis despite several months of negotiation.

We received data from the five remaining States (Illinois, Maine, Massachusetts, North Carolina, and South Carolina), including Medicaid enrollment and claims data and PCMH survey data,^a from more than 140 intervention and comparison practices. There were severe data quality issues in the baseline data for three of the States (Maine, Massachusetts, and North Carolina) that required considerable time and resources to resolve via multiple rounds of data submission and discussions with the States. Unable to overcome data limitations and faced with budget constraints, we ultimately excluded Maine and Massachusetts from the analysis.

In 2013, we analyzed baseline data (2009–2010) from the three remaining States (Illinois, North Carolina, South Carolina), presenting findings at the 2013 and 2015 AcademyHealth annual research conferences and publishing a manuscript in the journal *Academic Pediatrics*.^b

^a Maine, Massachusetts, North Carolina, and South Carolina used the Medical Home Index or the Medical Home Index-Revised Short Form developed for the evaluation. Illinois used the National Committee for Quality Improvement's Patient-Centered Medical Home practice self-assessment.

^b Christensen AL, Zickafoose JS, Natzke B, et al. Associations between practice-reported medical homeness and health care utilization among publicly insured children. *Acad Pediatr* 2015 May-June;15(3):267-74. Available at [http://www.academicpedsjnl.net/article/S1876-2859\(14\)00429-X/abstract](http://www.academicpedsjnl.net/article/S1876-2859(14)00429-X/abstract). Accessed October 27, 2015.

Site visits conducted by the NET in 2014 revealed that North Carolina and South Carolina shifted the focus of their projects from general PCMH improvement efforts as originally conceived to more specific QI projects that each practice conducted on targeted topics (for example, increasing dental visit rates, lowering obesity rates, or improving rates of documented developmental screenings). We decided not to conduct an impact evaluation on these projects because (1) no uniform set of outcomes applied to all practices; (2) the number of practices focusing on a given outcome was too small to expect an impact analysis to detect changes; and (3) we would be unable to assess some of the key outcomes (for example, BMI screening) through analysis of claims data, and alternative data collection methods, such as chart reviews, were not feasible.

Illinois was the only remaining State whose Category C project was designed to help practices enhance their PCMH features. Illinois provided PCMH survey data and Medicaid administrative data for calendar years 2009–2013 to support our impact evaluation. However, the State cannot yet provide data for the full post-intervention period (which includes calendar year 2014) because of the recent transition in CMS' Medicaid data systems. For 2009–2013 data, Illinois sent us Medicaid Statistical Information System (MSIS) data approved by CMS. The State is transitioning to the new data system (the Transformed Medicaid Statistical Information System, or T-MSIS) for all claims occurring January 2014 and later. Illinois halted MSIS production beyond that date and will not provide 2014 claims until CMS has transitioned to and validated the T-MSIS claims, the timeline of which has been delayed and remains unknown. Hence, we have no way of accessing the post-intervention data we need to conduct an impact analysis.