



January 2018

MEDICAID DEMONSTRATIONS

Evaluations Yielded Limited Results, Underscoring Need for Changes to Federal Policies and Procedures

Accessible Version

GAO Highlights

Highlights of [GAO-18-220](#), a report to congressional requesters

Why GAO Did This Study

Demonstrations—which represented roughly a third of the more than \$300 billion in federal Medicaid spending in 2015—are a powerful tool to test new approaches to providing coverage and delivering Medicaid services that could reduce costs and improve beneficiaries' outcomes. Evaluations are essential to determining whether demonstrations are having their intended effects. States are required to evaluate their demonstrations and CMS can initiate its own federal evaluations of demonstrations.

GAO was asked to examine evaluations of demonstrations, including how the results have been used to inform Medicaid policy. This report examines (1) state-led evaluations and (2) federal evaluations. GAO reviewed evaluation documentation for eight states with high demonstration expenditures that varied in the number of years their demonstrations had been in effect and by geography. GAO also reviewed documentation for the ongoing federal evaluations and interviewed state and federal Medicaid officials. GAO assessed evaluation practices against federal standards for internal control and leading evaluation guidelines.

What GAO Recommends

GAO recommends that CMS: (1) establish written procedures for requiring final evaluation reports at the end of each demonstration cycle, (2) issue criteria for when it will allow limited evaluations of demonstrations, and (3) establish a policy for publicly releasing findings from federal evaluations of demonstrations. HHS concurred with these recommendations.

View [GAO-18-220](#). For more information, contact Katherine Iritani at (202) 512-7114 or iritanik@gao.gov.

January 2018

MEDICAID DEMONSTRATIONS

Evaluations Yielded Limited Results, Underscoring Need for Changes to Federal Policies and Procedures

What GAO Found

Under section 1115 of the Social Security Act, the Secretary of Health and Human Services (HHS) may approve Medicaid demonstrations to allow states to test new approaches to providing coverage and for delivering services that can transform large portions of states' programs. However, GAO found that selected states' evaluations of these demonstrations often had significant limitations that affected their usefulness in informing policy decisions. The limitations included gaps in reported evaluation results for important parts of the demonstrations. (See table.) These gaps resulted, in part, from HHS's Centers for Medicare & Medicaid Services (CMS) requiring final, comprehensive evaluation reports after the expiration of the demonstrations rather than at the end of each 3- to 5-year demonstration cycle. CMS has taken a number of steps since 2014 to improve the quality of state-led evaluations, and in October 2017, officials stated that the agency planned to require final reports at the end of each demonstration cycle for all demonstrations. However, the agency has not established written procedures for implementing such requirements, which could allow for gaps to continue. CMS also plans to allow states to conduct less rigorous evaluations for certain types of demonstrations but has not established criteria defining under what conditions limited evaluations would be allowed.

Examples of Gaps in States' Evaluations of Medicaid Section 1115 Demonstrations

Arizona	The state was required to evaluate whether providing long-term services and supports under a managed care delivery model improved access and quality of care. The evaluation report lacked information on important measures of access and quality.
Arkansas	The state was required to evaluate the effects of using Medicaid funds to purchase private insurance for more than 200,000 beneficiaries. The evaluation did not address a key hypothesis that using private insurance would improve continuity of coverage for these beneficiaries, who were expected to have frequent changes in income that could lead to coverage gaps.
Massachusetts	The state was required to evaluate the effectiveness of its approach of providing up to \$690 million in incentive payments to seven hospitals to improve quality of care and reduce per capita costs. Evaluation reports submitted after 5 years provided no conclusions on the impact of the payments in these areas.

Source: GAO. | GAO-18-220

Federal evaluations led by CMS have also been limited due to data challenges that have affected the progress and scope of the work. For example, delays obtaining data directly from states, among other things, led CMS to considerably reduce the scope of a large, multi-state evaluation, which was initiated in 2014 to examine the impact of state demonstrations in four policy areas deemed to be federal priorities. Though CMS has made progress in obtaining needed data, it is uncertain when results from the multi-state and other federal evaluations will be available to policymakers because CMS has no policy for making results public. By not making these results public in a timely manner, CMS is missing an opportunity to inform important federal and state policy discussions.

Contents

Letter	1
Background	5
Limitations in State-Led Evaluations Hindered Their Usefulness and May Not Be Fully Addressed by CMS Improvements	12
Ongoing Federal Evaluations Led by CMS Have Been Limited by Data Challenges and It Is Uncertain When Results Will Be Available	23
Conclusions	30
Recommendations for Executive Action	31
Agency Comments and Our Evaluation	32
<hr/>	
Appendix I: Characteristics of Selected States' 1115 Demonstrations	35
Appendix II: Comments from the Department of Health and Human Services	37
Appendix III: GAO Contact and Staff Acknowledgments	41
GAO Contact	41
Staff Acknowledgments	41
<hr/>	
Appendix IV: Accessible Data	42
Agency Comment Letter	42
<hr/>	
Tables	
Table 1: Demonstration Types Included in Federal Multi-State Evaluation of Medicaid Section 1115 Demonstrations, Initiated in 2014	11
Table 2: Data Challenges Encountered in the Four Demonstration Types Examined in CMS's Multi-State Evaluation	24
Table 3: Rapid Cycle Reports Produced as Part of the Multi-State Evaluation, as of October 2017	27
Table 4: Characteristics of Selected States' Medicaid Section 1115 Demonstrations Selected for Our Review	35

Figures

Figure 1: Federal Expenditures under Section 1115 Demonstrations as a Percentage of Total Federal Medicaid Expenditures, by State, Fiscal Year 2015	7
Figure 2: Data Illustrating Gaps in Evaluation Reporting for Section 1115 Demonstrations in Selected States, as of November 2017	17

Abbreviations

CMS	Centers for Medicare & Medicaid Services
HHS	Department of Health and Human Services
DSRIP	delivery system reform incentive payment
MLTSS	managed long-term services and supports
PPACA	Patient Protection and Affordable Care Act
STC	special terms and conditions
T-MSIS	Transformed Medicaid Statistical Information System

This is a work of the U.S. government and is not subject to copyright protection in the United States. The published product may be reproduced and distributed in its entirety without further permission from GAO. However, because this work may contain copyrighted images or other material, permission from the copyright holder may be necessary if you wish to reproduce this material separately.



January 19, 2018

Congressional Requesters

Medicaid section 1115 demonstrations, which allow states to test and evaluate new approaches for delivering Medicaid services, have become a significant feature of the Medicaid program, increasing both in number and cost over the years and affecting millions of beneficiaries.¹ In November 2016, nearly three-quarters of states operated at least part of their Medicaid program under section 1115 demonstrations, and, in fiscal year 2015, federal demonstration expenditures amounted to \$109 billion or about one-third of Medicaid program expenditures that year.²

Under section 1115 of the Social Security Act, the Secretary of Health and Human Services may waive certain federal Medicaid requirements and approve new types of expenditures that would not otherwise be eligible for federal Medicaid matching funds for experimental, pilot, or demonstration projects that, in the Secretary's judgment, are likely to promote Medicaid objectives.³ For example, the Centers for Medicare & Medicaid Services (CMS), the agency within the Department of Health and Human Services (HHS) that oversees the Medicaid program, has approved states' proposals to extend Medicaid coverage under demonstrations to populations or for services that would not otherwise be covered under Medicaid.⁴ CMS has also allowed states to use Medicaid funds to finance costs that would not otherwise be eligible for federal funds, such as incentive payments to providers to improve access to and quality of care.

¹The Medicaid program—a joint, federal-state program that finances health care coverage for low-income and medically needy populations—covered an estimated 72.2 million individuals at an estimated cost of \$575.9 billion in fiscal year 2016, including about \$363.4 billion in federal spending and \$212.5 billion in state spending, according to 2016 estimates from the Centers for Medicare & Medicaid Services' Office of the Actuary.

²For further analysis of federal expenditures for Medicaid section 1115 demonstrations, see GAO, *Medicaid Demonstrations: Federal Action Needed to Improve Oversight of Spending*, [GAO-17-312](#) (Washington, D.C.: Apr. 3, 2017).

³42 U.S.C. § 1315(a).

⁴The authority to approve section 1115 demonstrations ultimately resides with the Secretary of Health and Human Services; the Secretary has delegated the approval and administration of Medicaid section 1115 demonstrations to CMS.

Because Medicaid section 1115 demonstrations (hereafter referred to as demonstrations) are intended to test new approaches to providing coverage and delivering Medicaid services, evaluations of the demonstrations are essential to determining whether the new approaches are having their intended effect. Evaluations are also critical to ensuring that information on the effects of demonstrations, such as on beneficiary access to care, quality of care, and costs of care is available to inform federal and state policy decisions about new approaches to coverage and care. Further, because demonstrations allow states to use Medicaid funds for costs that would not otherwise be covered under the program, evaluations serve as an important check of whether such funds are achieving federal Medicaid objectives.⁵ CMS has long required states to conduct evaluations of demonstrations.⁶ In addition, CMS has initiated its own federal evaluations of selected Medicaid demonstrations.

Given continued state interest in undertaking Medicaid section 1115 demonstrations and their budgetary significance and programmatic scope, you asked us to examine evaluations of demonstrations, including how the results have been used to inform Medicaid policy. This report examines:

1. state-led evaluations of demonstrations; and
2. federal evaluations of demonstrations led by CMS.

To examine state-led evaluations of demonstrations, we reviewed documentation for demonstrations in eight states—Arizona, Arkansas, California, Indiana, Kansas, Maryland, Massachusetts, and New York. We selected these states by first identifying the 15 states with the highest average demonstration expenditures for fiscal years 2013 through 2015—the most current, complete years of data available at the time we began our work.⁷ From those, we selected eight states to achieve variation with

⁵For more information on costs not otherwise eligible for funding approved under Medicaid demonstrations, see GAO, *Medicaid Demonstrations: Approval Criteria and Documentation Need to Show How Spending Furthers Medicaid Objectives*, [GAO-15-239](#) (Washington, D.C.: Apr. 13, 2015).

⁶While CMS has historically required states to evaluate their demonstrations, the Patient Protection and Affordable Care Act (PPACA) required the Secretary of HHS to establish a process for the periodic evaluation of demonstrations. The implementing regulations became effective in April 2012. See Pub. L. No. 111-148, § 10201(i), 124 Stat. 119, 922 (2010); 42 C.F.R. pt. 431.

⁷For the purposes of this report, a reference to states includes the 50 states and the District of Columbia.

regard to (1) total spending on the demonstrations, including as a percent of the state's total Medicaid spending, (2) the number of years the state's most comprehensive demonstration had been in place, and (3) geography.⁸ Together, demonstration spending in the eight states accounted for about 47 percent of total demonstration spending for fiscal year 2015.⁹ (See appendix I for more information on the characteristics of the demonstrations in our selected states.)

For each state-led demonstration, we reviewed the following (1) evaluation requirements delineated in the contract negotiated between CMS and the state—referred to as the special terms and conditions (STC), (2) evaluation design plans submitted by the state, and (3) evaluation reports submitted by the state, including any stated limitations or gaps in evaluation findings. For seven of our eight states—those which had completed more than one demonstration cycle—we reviewed the documentation for the most recently completed and current demonstration cycles as of the time of our review.¹⁰ For Kansas, which was in its first demonstration cycle at the time of our review, we reviewed the evaluation documentation for this cycle.¹¹ We also reviewed, when available, documentation of CMS's review of design plans and reports. We supplemented the documentation review by interviewing CMS officials about the agency's policies and procedures for overseeing state-led evaluations, including recent and planned changes in the agency's policies and procedures and the agency's use of evaluation findings in decision making. We also interviewed state Medicaid officials (in five of our eight selected states) to gain an understanding of the design and implementation of their evaluations and their interactions with CMS during the evaluation process.¹² In evaluating this information, we compared CMS's policies and procedures against standards for internal control in

⁸States may be approved to operate multiple Medicaid demonstrations. If a state operated more than one demonstration, we focused our review on the most comprehensive demonstration.

⁹Calculation is based on GAO's analysis of expenditure data from CMS's Quarterly Medicaid Statement of Expenditures for the Medical Assistance Program, known as the CMS-64 form, extracted on October 3, 2016.

¹⁰Demonstrations are typically approved for an initial 5-year period, which we refer to as a cycle, and can be renewed for subsequent cycles.

¹¹Kansas's first demonstration cycle began in 2013.

¹²We interviewed state officials in Arizona, Arkansas, Indiana, Massachusetts, and New York to achieve a mix of states in terms of demonstration age and spending.

the federal government, including those related to control activities and communication, and the American Evaluation Association's recommendations for evaluations of federal programs, which include recommendations related to the scope, quality, and transparency of evaluations.¹³

To examine federal evaluations of demonstrations led by CMS, we reviewed documents in the contract files for the two contract task orders (hereafter referred to as contracts) that CMS awarded in 2014 and 2015 to conduct the agency's ongoing federal evaluations of demonstrations. The options for these contracts were exercised annually and work was ongoing as of November 2017. The documentation we reviewed included the contract scopes of work that define the purposes of the contract, the timeframes for execution, and the expected products, or "deliverables;" monthly contractor progress reports; evaluation design documents; and other contract deliverables, including any reports of findings submitted as of October 2017.¹⁴ We reviewed the documents to assess the progress of the evaluations, including identifying any challenges encountered. We also interviewed CMS officials and one of CMS's contractors about the progress and status of the federal evaluations and about the agency's policies and procedures for conducting federal evaluations, including policies for identifying demonstrations for federal evaluation and for making evaluation results public. We compared CMS's policies and procedures against the American Evaluation Association's recommendations for evaluations of federal programs.¹⁵

We conducted this performance audit from November 2016 to January 2018, in accordance with generally accepted government auditing standards. Those standards require that we plan and perform the audit to obtain sufficient, appropriate evidence to provide a reasonable basis for

¹³GAO, *Standards for Internal Control in the Federal Government*, GAO-14-704G (Washington, D.C.: September 2014). Internal control is a process effected by an entity's oversight body, management, and other personnel that provides reasonable assurance that the objectives of an entity will be achieved.

American Evaluation Association, "An Evaluation Roadmap for a More Effective Government," accessed August 14, 2017, <http://www.eval.org/evaluationroadmap>.

¹⁴The contract deliverables included a variety of documents, such as evaluation design documents, rapid cycle reports, and contractor monthly progress reports. Some documents were obtained as draft copies as they were still under CMS review at the time of our study.

¹⁵American Evaluation Association, "Evaluation Roadmap."

our findings and conclusions based on our audit objectives. We believe that the evidence obtained provides a reasonable basis for our findings and conclusions based on our audit objectives.

Background

Medicaid Section 1115 Demonstrations

Nearly three-quarters of states (37 as of November 2016) have CMS-approved Medicaid section 1115 demonstrations, which allow states to test new approaches to coverage and to improve quality and access or generate savings or efficiencies. CMS has approved demonstrations for a wide variety of purposes. For example, under demonstrations, states have extended coverage to populations or for services not otherwise eligible for Medicaid, made payments to providers to incentivize delivery system improvements, and, more recently, expanded Medicaid to certain low-income adults by using Medicaid funds to purchase private health insurance coverage.¹⁶ While state demonstrations vary in size and scope, many are comprehensive in nature, affecting multiple aspects of states' Medicaid programs simultaneously. For example, Kansas's demonstration, approved in 2012, significantly expands the use of managed care to deliver physical, behavioral, and long-term care services to almost all the state's Medicaid populations, care that for some populations was previously provided on a fee-for-service basis.¹⁷ The demonstration also established a funding pool of up to \$344 million to provide payments to hospitals to finance uncompensated care.¹⁸

¹⁶Under PPACA, states may opt to expand their Medicaid programs by covering nearly all adults with incomes at or below 133 percent of the federal poverty level beginning January 1, 2014. See Pub. L. No. 111-148, 124 Stat. 119 (2010), as amended by the Health Care and Education Reconciliation Act of 2010 (HCERA), Pub. L. No. 111-152, 124 Stat. 1029 (2010). For purposes of this report, references to PPACA include the amendments made by HCERA. PPACA also provides for a 5 percent disregard when calculating income for determining Medicaid eligibility, which effectively increases this income level to 138 percent of the federal poverty level. See 42 U.S.C. § 1396a(a)(10)(A)(i)(VIII), (e)(14)(I).

¹⁷Under a managed care delivery model, states typically contract with managed care organizations to provide a specific set of Medicaid-covered services to beneficiaries and pay them a set amount per beneficiary per month to provide those services.

¹⁸This is the total costs (federal and state) over 5 years (calendar years 2013 through 2017) for which Kansas may claim federal matching funds, according to the STCs of Kansas's demonstration, dated January 29, 2014.

Kansas's demonstration expenditures accounted for about 94 percent of the state's total Medicaid expenditures in fiscal year 2015.¹⁹

In fiscal year 2015, federal spending under demonstrations represented a third of all Medicaid spending nationwide. In 10 states, federal spending on demonstrations represented 75 percent or more of all federal spending on Medicaid. (See fig. 1.)

¹⁹Calculation is based on GAO's analysis of expenditure data from the CMS-64 form, extracted on October 3, 2016.

Demonstrations are typically approved by CMS for an initial 5-year period (referred to as a demonstration cycle), but some states have operated portions of their Medicaid programs under a demonstration for decades. This can be achieved through a series of renewals approved by CMS, generally occurring every 3 to 5 years.²⁰ What a state is testing and implementing under its demonstration can change from one cycle to the next. States often make changes to their demonstrations, either through the renewal process or by requesting an amendment during the demonstration cycle. These changes can be relatively small or can be significant and can represent testing of a new approach for the state.²¹ For example, at renewal a state could request approval to expand coverage to a new population or add requirements that beneficiaries share in the cost of care by paying a monthly premium.

CMS Oversight of State-Led Evaluations

CMS has long required states to conduct evaluations of section 1115 demonstrations.²² CMS oversees the evaluations and can influence them at several key points during the demonstration process.

- **Application review and approval:** When a state applies for a demonstration, CMS reviews the state's application, which describes the goals and objectives of the demonstration and what the demonstration will test, among other things.²³ As part of the review and approval process, CMS negotiates with the state on the STCs, including evaluation requirements. These requirements might include, for example, reporting timeframes and broad standards for the evaluation, such as standards around the independence of the evaluator and acceptable evaluation methods.
- **Evaluation design phase:** After a demonstration is approved, states are required to submit an evaluation design to CMS for review and

²⁰Renewals are also referred to as extensions by CMS.

²¹CMS may, at its discretion, treat the renewal application as an application for a new demonstration, if the renewal includes substantial changes to the existing demonstration. 42 C.F.R. § 431.412(c)(1) (2016).

²²In addition to CMS's historic practice of requiring states to evaluate demonstrations, a provision in PPACA required CMS to promulgate regulations establishing a process for the periodic evaluations of demonstrations in 2012. See 42 C.F.R. pt. 431.

²³See 42 C.F.R. § 431.412(a) (2016).

approval.²⁴ The evaluation design must discuss, among other things, the hypotheses that will be tested, the data that will be used, and how the effects of the demonstration will be isolated from other changes occurring in the state. During review of the design, CMS can seek adjustments such as requiring the state to address certain objectives or using particular performance measures.

- **Demonstration renewal:** In the event that a state wishes to renew its demonstration, it must generally submit an application to CMS at least 1 year before the demonstration is scheduled to expire. The application must include, among other things, a report presenting the evaluation's findings to date, referred to as an interim evaluation report.²⁵ CMS can use the information from the interim evaluation report to negotiate changes in the STCs for the evaluation of the next demonstration cycle. If CMS renews the demonstration, the evaluation process starts over with the state submitting a new evaluation design that reflects changes in what is being tested in the new cycle.
- **Demonstration end:** CMS requires states to submit a final evaluation report for review and approval generally after the end of the demonstration, at which time the agency can work with the state to, for example, add clarity and disclose the limitations of the evaluation before the final evaluation report is made public.²⁶

Within the framework that CMS has established for state-led evaluations, states design evaluations to the specifics of their demonstrations. As the size and scope of demonstrations varies considerably across states, so, too can evaluations vary in their breadth and complexity. State-led evaluations may assess the effects of several different policies, each with its own set of hypotheses—predictions of the effects of the policy—and methods. For example, a state could evaluate the effects of moving to a managed care delivery model for providing managed long-term services and supports (referred to as MLTSS), implementing provider payment pools aimed at delivery system reform, and expanding coverage to a new population all within the same demonstration. Each of those three

²⁴See 42 C.F.R. § 431.424(c) (2016).

²⁵State interim evaluations must be published on the state's website within 30 days of submission to CMS. 42 C.F.R. § 431.424(d)(2) (2016). In some cases, CMS has required a state to submit an interim evaluation report, regardless of whether the state planned to apply to renew its demonstration.

²⁶CMS must post to its website, or link to a state's website, all evaluation materials for the purposes of sharing findings with the public within 30 days of receipt of materials. See 42 C.F.R. § 431.424(g) (2016).

elements would have its own hypotheses and methods and may have varying timeframes for the number of years of experience needed to be able to effectively measure the effects of what is being tested.

Federal Evaluations

CMS has the authority to initiate its own federal evaluations of section 1115 demonstrations, and states must fully cooperate with any such evaluations.²⁷ Between 2014 and 2016, CMS initiated three federal evaluations that were ongoing as of November 2017. The first evaluation, initiated in 2014, is a large, multi-state evaluation examining four broad demonstration types in several states.²⁸ (See table 1.) According to CMS, it selected these demonstration types—which together account for tens of billions of dollars in federal and state Medicaid spending—because they included policies that the agency considered priority areas for evaluation. CMS awarded a contract to an evaluation organization to implement the 5-year study. According to CMS, the estimated total cost of this evaluation for the 5-year life of the contract is \$8.3 million.²⁹ The evaluation was designed to produce three sets of results: a series of reports providing contextual information about the demonstrations being evaluated, referred to as rapid cycle reports; interim evaluation reports featuring early results of more in-depth analysis; and final evaluation reports.

²⁷See 42 C.F.R. § 431.420(f) (2016).

²⁸CMS officials stated that the multi-state evaluation was the first federal evaluation of demonstrations conducted in over a decade. Officials also stated that the agency considers the multi-state evaluation, which covers four demonstration types, to be four separate evaluations. For the purposes of this report, we refer to the multi-state evaluation as a single evaluation covering the four demonstration types.

²⁹CMS awarded a contract to Mathematica Policy Research to conduct the multi-state evaluation. The award is for a base year and 4 option years, so up to 5 years in total, if CMS decides to exercise— or fund—each of the options. The contract includes tasks beyond conducting the multi-state evaluation, such as the contractor providing technical assistance in reviewing states' monitoring and evaluation reports, for a total contract cost of \$16 million.

Table 1: Demonstration Types Included in Federal Multi-State Evaluation of Medicaid Section 1115 Demonstrations, Initiated in 2014

Demonstration type	Description
Delivery system reform incentive payment (DSRIP)	<ul style="list-style-type: none"> Demonstrations provide incentive payments to providers that engage in various improvement projects that align with state delivery system reform objectives. Examples of reform objectives include improving clinical quality, improving population health, establishing value-based payment systems. As of February 2017, states with the largest DSRIP demonstrations include California, Massachusetts, New York, and Texas. In total, CMS authorized these states to spend almost \$38.7 billion for these programs nonconsecutively from 2011 through 2022.
Premium assistance to purchase exchange coverage	<ul style="list-style-type: none"> Demonstrations allow states to use Medicaid funds to provide premium assistance to beneficiaries newly eligible under the Patient Protection and Affordable Care Act to purchase private insurance offered on the state's health insurance exchange. As of August 2017, 2 states, Arkansas and New Hampshire, had such demonstrations in effect.
Beneficiary engagement/premiums	<ul style="list-style-type: none"> Demonstrations allow states to use incentives to encourage personal responsibility and/or healthy behaviors among beneficiaries. Examples include providing financial rewards or enhanced benefits to beneficiaries who seek preventive care, or requiring monthly premium payments or copayments that can be reduced by completing certain services, or disenrolling beneficiaries at certain income levels for failure to make monthly premium payments. As of August 2017, 6 states had been approved to implement beneficiary engagement/premiums for newly eligible adults: Arizona, Arkansas, Indiana, Iowa, Michigan, and Montana.
Managed long-term services and supports (MLTSS)	<ul style="list-style-type: none"> Demonstrations allow states to use managed care delivery systems to provide long-term services and supports to beneficiaries, including care in institutions and home and community-based care. As of May 2017, 22 states provided long-term services and supports under section 1115 and other Medicaid authorities.^a

Source: GAO review of Centers for Medicare & Medicaid Services (CMS) documentation. | GAO-18-220

^aSome states operate MLTSS programs under Medicaid waiver authorities other than section 1115. CMS's contractor examined MLTSS programs regardless of federal authority because, according to the contractor, the type of federal authority states use is likely to be less important to affecting outcomes than other program features.

CMS contracted with another evaluation organization to conduct two federal evaluations examining demonstrations in single states—Indiana and Montana—over 4 years. As of September 2017, the estimated cost of this contract, inclusive of all options, was \$8.2 million.³⁰ In total, spending for Indiana's and Montana's demonstrations was about \$2 billion in fiscal year 2015, including \$1.6 billion in federal spending.

- Indiana:** CMS initiated this evaluation in 2015. CMS officials told us they started this evaluation to better understand how policies in

³⁰CMS awarded a contract to Social & Scientific Systems, Inc. to conduct the federal evaluations of Indiana's and Montana's demonstrations. The award is for a base year and up to 3 options years, so 4 years in total if CMS decides to exercise each of the options.

Indiana's demonstration, many of which were unprecedented, were affecting beneficiaries. These policies included, for example, charging monthly contributions for most newly eligible adults with incomes from 0 to 138 percent of the federal poverty level; imposing a lock-out period of 6 months for nonpayment of premiums for most people with incomes above the federal poverty level; and charging co-payments above statutory levels for non-urgent use of emergency room services. The federal evaluation is aimed at estimating the effects of Indiana's demonstration on health insurance coverage and access to and use of care, and documenting beneficiary understanding of enrollment, disenrollment, and copayment policies, among other things.

- **Montana:** CMS initiated this evaluation in 2016. CMS officials told us they started this evaluation to provide a point of comparison to Indiana's demonstration, as Montana was implementing similar policies to Indiana but with some variations. For example, under Montana's demonstration, the state charges premiums to most newly eligible adults with incomes between 51 and 138 percent of the federal poverty level; and disenrolls beneficiaries with incomes above the federal poverty level for nonpayment of premiums, with reenrollment when overdue premiums are paid. Similar to the federal evaluation of Indiana's demonstration, the evaluation of Montana's demonstration is aimed at estimating the effects of the demonstration on insurance coverage, access to and use of care, and documenting beneficiary understanding of and experience with premiums, copayments, enrollment, and disenrollment, among other things.

Limitations in State-Led Evaluations Hindered Their Usefulness and May Not Be Fully Addressed by CMS Improvements

State-led evaluations of demonstrations in selected states often had significant methodological weaknesses and gaps in results that affected their usefulness for federal decision-making. Though CMS has been taking steps since 2014 to improve the quality of these evaluations, the agency has not established written procedures to help implement some of these improvements.

State-Led Evaluations in Selected States Often Had Significant Limitations That Affected Their Usefulness in Informing Federal Decision-Making

The state-led evaluations we reviewed in our selected states often had methodological limitations that affected what could be concluded about the demonstration's effects. CMS hired a contractor to review state evaluation designs and reports, and that contractor identified a number of methodological concerns with the evaluations in our selected states.³¹ For example, CMS's contractor raised concerns about the comparison groups, or lack thereof, used to isolate and measure the effects of the demonstrations in the Arkansas, California, Indiana, and Maryland evaluations.³² The contractor also raised concerns with the sufficiency of sample sizes and survey response rates for beneficiary surveys in Indiana. These surveys were key methods for assessing the effect of demonstrations on access, beneficiary understanding, and perceptions on affordability. Finally, the contractor raised concerns with the analysis of the effects of the demonstration on cost in Arkansas, California, and Maryland. Officials in several states told us that some of the methodological limitations in their evaluations were difficult to control. For example, officials in two states told us that isolating the effects of the demonstration was difficult given other changes happening in the state's health care system at the same time. Some state officials also noted that state resources, including both funding and staff capacity, present challenges in completing robust evaluations.³³

³¹This work was performed as one of the tasks in CMS's 5-year contract with Mathematica Policy Research.

³²Though not a state for which we had documentation of CMS's review, officials from Arizona also told us that their evaluation lacked a control group for a key portion of the demonstration, which made it difficult to measure impact.

³³Among the five states where we interviewed officials, officials from three states told us that evaluation costs ranged from an average cost of \$800,000 for a 5-year period in Arizona to \$3.9 million for New York's DSRIP program. Arizona's demonstration expenditures amounted to \$10.6 billion in fiscal year 2015. New York's DSRIP program has a spending limit of \$12.8 billion for its 5-year term that started in 2015.

Policy Areas Identified by CMS as Priority for Evaluation

- Delivery System Reform Incentive Payment (DSRIP) programs
- Premium assistance to purchase exchange coverage
- Beneficiary engagement policies, such as requiring monthly contributions
- Use of managed care to deliver long-term services and supports (MLTSS)

Source: Centers for Medicare & Medicaid Services (CMS). | GAO-18-220

In addition, we found that state-led evaluations in our selected states often had gaps in results for significant aspects of the demonstrations, including those aspects that CMS officials had identified as high priority policy areas for evaluation. (See sidebar.) These gaps included, for example, not answering key hypotheses and not reporting on key outcome measures. Specific examples of the gaps in the results from state-led evaluations included the following:

- **Massachusetts:** Massachusetts' demonstration includes a DSRIP program, with approved funding up to about \$690 million.³⁴ Under the demonstration STCs, the state was required to evaluate whether the seven hospitals participating in the DSRIP were able to show improvements on certain outcome measures related to improving quality of care, improving population health and access to care, and reducing the per capita costs of health care. However, the evaluation report, submitted by the state 5 years after approval of the DSRIP program, provided only descriptive or summary information about the number and types of projects implemented by the hospitals receiving payments and did not provide any data to measure or conclusions on the effects of those payments.
- **Arkansas:** Under its demonstration, the state was testing the effects of using Medicaid funds to provide premium assistance for the more than 200,000 beneficiaries newly eligible under PPACA to purchase private insurance offered through the state's health insurance exchange. The state's evaluation was designed to assess whether beneficiaries would have equal or better access to care and equal or better outcomes than they would have had in the Medicaid fee-for-service system. The evaluation was also aimed at examining continuity of coverage for beneficiaries, as the expansion population was anticipated to have frequent income fluctuations leading to changes in eligibility and gaps in coverage. However, evaluation results submitted over two and a half years into the demonstration—the only results submitted for the state's first cycle—were limited to data only from the first year of the demonstration and did not provide data on continuity of coverage. Achieving continuity of coverage was part of the state's rationale for using an alternative approach to Medicaid expansion.

³⁴Massachusetts' demonstration was first implemented in 1997 and added a DSRIP program to the demonstration in 2011. Massachusetts referred to the DSRIP component of its demonstration as the Delivery System Transformation Initiative.

- **Arizona:** Among other things, Arizona’s demonstration includes MLTSS, including for the particularly complex populations of adults who have intellectual and developmental disabilities and for children with disabilities.³⁵ As part of its evaluation, the state was assessing whether the quality of and access to care, as well as quality of life, would improve during the demonstration period for long-term care beneficiaries enrolled in MLTSS. However, evaluation results submitted in October 2016—the only results submitted for the state’s most recently completed demonstration cycle—lacked data on key measures of access, such as hospital readmission rates, and on quality of life, such as beneficiaries’ satisfaction with their health plan, provider, and case manager.³⁶

A key contributor to the gaps in the information included in the state-led evaluations we reviewed was that CMS historically had not required the states to submit final, comprehensive evaluation results at the end of each demonstration cycle. As a result, for our selected states, including those discussed above, CMS had received only interim evaluation reports that were generally based on more limited data from the early years of the demonstration cycle and did not include all of the analyses planned. Though CMS had required final evaluation reports in the demonstration STCs, the due dates for those reports were tied to the expiration of the demonstrations or, in one case, CMS did not enforce the specified due date.³⁷ Under such conditions, due dates for final evaluation reports were effectively pushed out when the demonstrations were renewed. Evaluation due dates could be pushed out for multiple cycles. CMS officials acknowledged that the lack of data in the interim evaluation reports from the more mature years of the demonstration affected the conclusions that could be drawn from them.

³⁵Arizona’s MLTSS program includes all populations eligible for LTSS, including individuals with developmental or physical disabilities and the elderly.

³⁶Arizona submitted its interim evaluation in October 2016, one month after CMS had approved a 5-year renewal of the demonstration. According to CMS requirements, the report should have been submitted with the state’s application to renew the demonstration, about 1 year earlier. According to CMS officials, though a report was not submitted prior to approval, the agency subsequently required the state to submit an interim evaluation by November 2016.

³⁷New York’s STCs originally required the state to submit a final evaluation report by a specific date; but, when the state subsequently sought and received approvals for several temporary extensions of the demonstration, CMS did not adjust the due date for the final evaluation or otherwise require a final evaluation report for this demonstration cycle.

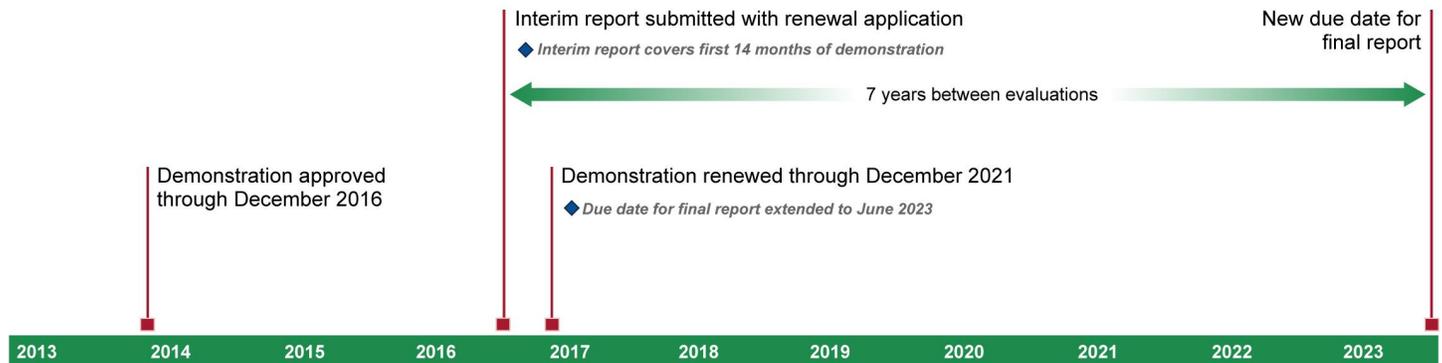
We found that due dates for final evaluation reports were pushed out upon renewal in all seven of our states that had completed a demonstration cycle, leading to a gap in evaluation reporting of up to 6 or 7 years for several states.³⁸ In Maryland, for example, CMS approved the demonstration to run from 2013 to 2016 with a final evaluation report due 120 days after the expiration of the demonstration. In 2016, CMS extended the demonstration, pushing the deadline for the final evaluation report to 18 months following the end of the new cycle, or June 2023. At that time, it will be 7 years since the interim evaluation report was submitted.³⁹ See figure 2.

³⁸In two states, Arizona and Arkansas, when demonstrations were renewed, CMS required the states, as a condition of their renewal, to submit final evaluation reports for the previous demonstration cycle relatively soon after the end of that cycle, making this gap smaller.

³⁹If the state elected to renew the demonstration again, the state would be required to submit another interim evaluation report as part of its application for renewal at that time.

Figure 2: Data Illustrating Gaps in Evaluation Reporting for Section 1115 Demonstrations in Selected States, as of November 2017

Maryland



State	Start date of most recently completed demonstration cycle	Date interim evaluation report submitted	Date final evaluation report due ^a	Number of years between submission of interim evaluation report and due date for final evaluation report
CA	November 2010	March 2015	December 2021	7
MD	November 2013	June 2016	June 2023	7
MA	October 2014	July 2016	December 2022	6.5
AR	September 2013	June 2016	July 2018	2
IN	January 2013	June 2014 and July 2016 ^b	April 2018 ^c	2 ^c
NY	August 2011	May 2014	No date specified ^d	Undetermined
AZ	October 2011	October 2016	-- ^e	Undetermined

Source: GAO analysis of documentation from the Centers for Medicare & Medicaid Services (CMS). | GAO-18-220

Notes: In the Maryland illustration, if the state were to seek another renewal, the state would be required to submit an interim evaluation report as part of its application. Similarly, if other states sought renewal, an interim evaluation report would be required a year prior to the expiration of the demonstration.

^aThe date that the final evaluation report is due represents the new due date set in the special terms and conditions (STCs) when the demonstration was renewed.

^bIndiana submitted an interim evaluation for its demonstration in June 2014. In 2015, CMS approved an expanded version of the demonstration, for which the state submitted an interim evaluation in July 2016.

^cAs of October 2017, Indiana had an application pending with CMS to renew its demonstration through January 31, 2021. If the demonstration is renewed, the due date for the final evaluation and the number of years between evaluation reports could change.

^dCMS did not specify in New York's current demonstration cycle when a final evaluation report would be due. As of October 2017, CMS officials were planning to issue a technical correction to the STCs to address this.

^eArizona was required to submit a final evaluation report for its most recently completed demonstration cycle by November 2016, but, as of October 2017, the state had not submitted this report. According to CMS officials, Arizona's final evaluation was delayed by data collection issues and is expected by the end of 2017.

The limitations in state-led evaluations—including methodological weaknesses and gaps in results—have, in part, hindered CMS’s use of them to inform its policy decisions. CMS officials told us that, historically, state-led evaluations have generally provided descriptive information but lacked evidence on outcomes and impacts. As a result, officials noted that they consider the data reported in the evaluations but, generally, state-led evaluations have not been particularly informative to their policy decisions. CMS officials told us that there have been cases where data, but not the conclusions, from state-led evaluations have informed their thinking on certain policy changes. For example, CMS officials said that data reported in early evaluations of DSRIP programs helped them in considering whether and how the agency should modify the basic policy structure of these programs. State officials had mixed perspectives on whether state-led evaluations influenced CMS decision-making around renewing their demonstrations. Officials in one state told us that while CMS reviewed their interim evaluation results, the results did not appear to influence the negotiations around the demonstration renewal. In contrast, officials from another state told us that discussion of interim evaluation results and limitations was a significant part of negotiations in 2016 regarding whether CMS would be willing to reauthorize funding for certain programs, including a new DSRIP investment and broader delivery system reforms the state was trying to implement. Officials in several states told us that there was value to state-led evaluations and in the federal-state partnership in designing the evaluations.

CMS Is Taking Steps to Improve the Quality of State-Led Evaluations, but Lacks Written Procedures to Ensure That All Evaluations Will Be Subject to New Requirements

CMS has implemented several procedures since 2014 aimed at improving the quality of state-led evaluations. CMS officials told us that these changes were part of CMS placing increased focus on monitoring and evaluation, which also resulted in CMS establishing a new office in 2015 that is responsible for these activities. One of the key changes CMS began implementing in 2014 was to set more explicit requirements for evaluations in the STCs, including requirements to improve the evaluation methodologies. According to CMS officials, the agency realized that one reason why state-led evaluations had generally lacked rigor and been of limited usefulness was that CMS had not been setting clear expectations for evaluations in the STCs. The officials said that CMS began strengthening evaluation requirements starting in 2014 with

demonstrations implementing approaches in CMS's high priority policy areas.

In our review of the STCs for current demonstration cycles in our seven selected states that had completed a demonstration cycle, all of which were approved in 2014 or later, we found evidence of CMS's efforts. Specifically, we found an increased focus on the use of independent evaluators and more explicit expectations for rigor in the design and conduct of evaluations:

- **Consistent requirements for independent evaluators.** The STCs for the most recently approved cycle of demonstrations in all seven states required the state to use an independent evaluator to conduct the evaluation.⁴⁰ In some cases, the STCs also required that the evaluation design discuss the process to acquire the independent evaluator, including describing the contractor's qualifications and how the state will assure no conflict of interest. These requirements were new in most states.
- **More explicit expectations for rigor.** In four of the seven states we reviewed, the STCs for the most recently approved cycle of states' demonstrations included new, explicit language requiring state evaluations to meet the prevailing standards of scientific and academic rigor. These included standards for the evaluation design and conduct as well as the interpretation and reporting of findings. Some states' STCs further specified the characteristics of rigor that CMS expected, including using the best available data, discussing the generalizability of results, and using controls and adjustments for and reporting the limitations of data and their effects on results. According to CMS, in the past, states have not always discussed methodological limitations in their evaluation reports.

In addition to strengthening evaluation requirements, CMS has also taken steps since 2014 to enhance its oversight during the design and early stages of state-led evaluations, and, according to officials, some of these steps are likely to improve the usefulness of evaluations. Specifically, CMS has provided technical assistance to help states design their evaluations, sometimes leveraging expertise from other parts of HHS, including the HHS Office of the Assistant Secretary for Planning and Evaluation and the Center for Medicare & Medicaid Innovation as well as

⁴⁰In the absence of this requirement, the state Medicaid agency could conduct its own evaluation rather than seeking an outside entity to do so.

outside contractors. For example, officials stated that the agency assists states in developing relevant and standardized measures and provides assistance to help address states' data limitations. Officials said this has resulted in more robust evaluation designs with increased potential to isolate outcomes and impacts.⁴¹

CMS has also used contractors to help in its review of state evaluation designs, including sampling designs, and evaluation reports. Since 2014, one contractor has provided over 30 assessments of evaluation designs and findings in at least 11 states. According to officials, this has increased CMS's capacity to identify methodological weaknesses and negotiate changes with states to improve the usefulness of evaluations. For example, CMS's contractor reviewed four draft survey instruments that Indiana planned to use in its evaluation, providing comments on the sampling frames and the structure and organization of survey questions. In response to the contractor's feedback, Indiana made changes to the surveys to gather more reliable information and improve their readability.

Finally, CMS has begun making changes to how it sets due dates for final evaluation reports. CMS officials told us that in spring 2017, CMS began requiring states to submit a comprehensive evaluation report for demonstrations in its high priority policy areas for evaluation at the end of each demonstration cycle, rather than after the expiration of the demonstration.⁴² CMS's recent demonstration renewals in Florida and Missouri—approved in August and September of 2017, respectively—required a final, summative evaluation report at the end of the demonstration cycle, consistent with the policy. In October 2017, CMS officials stated that the agency was expanding this policy and was now planning to require final reports at the end of each cycle for all demonstrations, as they are approved or renewed. However, CMS had not established written procedures for implementing this new policy.

⁴¹For example, CMS officials said they have been placing an increased emphasis on states conducting impact analyses and have leveraged evaluation expertise from the Center for Medicare & Medicaid Innovation and the Office of the Assistant Secretary for Planning and Evaluation to help states with some of the statistical techniques involved in estimating the effects of their demonstrations.

⁴²There is evidence that CMS had been taking steps in this direction prior to this stated policy change. In five of our selected states, the STCs for the states' most recently approved demonstration cycles required the states to submit final evaluation reports after the end of that cycle, which could prevent the gaps in reporting that we found in those states going forward.

It is too soon to assess the effectiveness of CMS's recent efforts to strengthen state-led evaluations. CMS has been implementing the strategies on a rolling basis as states apply for demonstration renewals and new demonstrations. If implemented and enforced consistently, CMS's efforts to improve the quality of state-led evaluations have the potential to result in more conclusive evaluations. Further, CMS's efforts to improve the quality of state-led evaluations and its plan to require final reports after each demonstration cycle are consistent with evaluation guidance from the American Evaluation Association that recommends that federal agencies conduct evaluations of public programs and policies throughout the programs' life cycles, not just at their end, and that agencies use evaluations to improve programs and assess their effectiveness.⁴³ Federal internal control standards also state that management should implement control activities through policies.⁴⁴ However, CMS does not have written procedures for implementing its planned policy, for example, for ensuring that the requirement is included in the STCs for all demonstrations, despite unique negotiations with each state, and that those requirements are consistently enforced. As a result, some state-led evaluations could continue to produce only more limited, interim findings that leave critical questions about the effects of these demonstrations on beneficiaries and costs unanswered.

CMS oversight of state-led evaluations may see further changes, as CMS officials told us that their oversight procedures are still evolving. For example, CMS officials told us that as of October 2017 the agency plans to begin to make distinctions in the level of evaluation required across demonstrations. They said that they are considering, for example, whether longstanding and largely unchanged components of a demonstration, and approaches previously tested by a number of other states without concern, require the same level of evaluation as testing a new approach to Medicaid expansion.⁴⁵ Officials said that they plan to include language in demonstration STCs, as the agency did in the recent

⁴³American Evaluation Association, "Evaluation Roadmap."

⁴⁴[GAO-14-704G](#).

⁴⁵CMS issued guidance to states in November 2017 indicating the agency will also be considering these factors in approving demonstrations and determining the reporting frequency for ongoing monitoring. The guidance indicates that CMS may approve the extension of routine, successful, non-complex demonstrations in a state for a period up to 10 years and that CMS will reduce the frequency of required monitoring reports for demonstrations that, among other things, are long-standing, noncomplex, and unchanged and that have previously been rigorously evaluated and proven successful.

renewals for Florida and Missouri, instructing the state to consider those factors as the state designs its evaluation. Specifically, in the evaluation design submitted for CMS approval, the state should include in the discussion of limitations whether the demonstration is long-standing, noncomplex, has previously been rigorously evaluated and found to be successful, or is also considered to be successful without issues or concerns. CMS officials said that the expected level of rigor for the evaluation could be balanced against such factors.

The implications of limiting evaluation requirements for certain types of demonstration approaches would depend on CMS's definitions of what is, for example, noncomplex or has previously been rigorously evaluated. As of October 2017, CMS had not established specific criteria for determining when a demonstration component would require less rigorous evaluation. Agency officials told us they were planning to develop such criteria after concluding a pilot of alternative criteria and expectations in certain demonstrations related to providing services for family planning and former foster care children. They said that when these pilots have concluded they will evaluate the results. It is unclear how these narrowly scoped demonstrations—scoped for a particular type of service or population—can be used to inform criteria for comprehensive demonstrations that can affect a state's entire Medicaid population and all services. Further, though CMS has begun indicating to states, including those with comprehensive demonstrations, that the agency may allow less rigorous evaluations for certain types of demonstration approaches, CMS has not established timeframes for issuing the criteria defining those conditions.

Federal standards for internal control stress that management should implement control activities through policy and should internally and externally communicate necessary information to achieve the agency's objectives. If CMS does not establish clear criteria for components of demonstrations that require limited evaluation, characteristics such as "long-standing" or "noncomplex" could be broadly interpreted. This could result in demonstrations that receive significant amounts of federal funds and affect many beneficiaries not being thoroughly evaluated. Written criteria could also reduce the potential for inconsistencies in the level of evaluation required across demonstrations.

Ongoing Federal Evaluations Led by CMS Have Been Limited by Data Challenges and It Is Uncertain When Results Will Be Available

Data and other challenges have significantly limited the scope and progress of CMS's large, multi-state evaluation and the agency's evaluation of Indiana's demonstration. Further, CMS has not released available evaluation results from the multi-state evaluation nor set timeframes for making these and future federal evaluation findings public.

Data Challenges Have Limited the Scope and Progress of Federal Evaluations

CMS encountered numerous data challenges in its multi-state evaluation that significantly reduced the scope of the analyses planned. These data challenges included limitations in the quality of CMS data and delays obtaining data directly from states.⁴⁶ These limitations caused CMS to narrow the evaluation's scope, often by reducing the number of state demonstrations evaluated or limiting what was being examined. All four demonstration types targeted in the multi-state evaluation—which reflect CMS's high priority policy areas—were affected by these challenges.⁴⁷ In the most extreme case, data limitations reduced the scope of the MLTSS evaluation to two states out of the more than 20 states operating such programs. As a result, the evaluation findings will not be generalizable to all MLTSS programs. (See table 2.) The data challenges were in addition to other challenges that affected the evaluation. For example, there were difficulties in trying to isolate demonstration effects in the context of rapidly changing health systems, or recent demonstrations had not been

⁴⁶CMS regulations require states to fully cooperate with federal evaluations of demonstrations, including submitting all requested data and information to CMS or its independent evaluator. 42 C.F.R. § 431.420(f) (2016). In some of our selected states, we observed that STCs effective beginning in 2017 contained more explicit requirements for full and timely state cooperation with federal evaluations. These provisions included requiring states to provide data and analytic files to CMS and to enter into a data use agreement. In some cases, the STCs further allowed CMS to defer \$5 million in federal payments to states for not complying with these requirements.

⁴⁷In some cases, CMS hoped alternative data sources would become available as the evaluation progressed, allowing the agency to include additional states and address research questions in the final evaluation report that had to be deferred due to data issues.

in operation long enough to allow CMS to appropriately assess longer-term effects.⁴⁸

Table 2: Data Challenges Encountered in the Four Demonstration Types Examined in CMS’s Multi-State Evaluation

Delivery system reform incentive payments (DSRIP)

- **Focus of evaluation:** For states making payments to providers to incentivize delivery system improvements, examine the effects of these projects on quality of care, delivery system transformation, population health, cost of care, and readiness for value-based purchasing.
- **Data challenges:** Available data were sufficient to examine DSRIP programs in only 3 of the 10 states with such programs for the interim evaluation—California, Texas, and New Jersey. In 2 of the 3 states complete and usable inpatient encounter records for adults were not available.
- **Implications for scope:** Large DSRIP programs will not be examined in New York, which has a spending limit of \$12.8 billion, and Massachusetts, which has a spending limit of \$691 million.^a Interim evaluation outcome measures will be limited to those that rely on emergency room and outpatient data.

Premium assistance to purchase exchange coverage

- **Focus of evaluation:** For states using Medicaid funds to purchase exchange coverage for newly eligible adults, examine whether enrollees with exchange coverage access care at similar or better rates than those in traditional Medicaid and whether individuals enroll in such programs at the same rates they enroll in Medicaid in other Medicaid expansion states. The evaluation will also examine how per-beneficiary spending compares with other Medicaid expansion states.
- **Data challenges:** CMS’s data were not sufficient for these analyses, due in part to the agency’s transition to a new Medicaid data system. CMS had to request data directly from the 3 states testing this approach. Obtaining data for Arkansas, the first state approved to test this approach, took over 1 year. Data for a second state, New Hampshire, were limited due to the demonstration having less than 1 year of operational experience. Finally, available data on costs were not detailed enough to examine administrative costs.
- **Implications for scope:** Only limited information on Arkansas, the state with the most experience, and New Hampshire will be included in the interim evaluation. The interim evaluation will rely mainly on analyses of Iowa’s program, which was discontinued after experiencing implementation challenges. An assessment of the administrative costs of the demonstration is being deferred until the final evaluation.

Beneficiary engagement policies

- **Focus of evaluation:** For states testing provisions requiring beneficiaries to pay monthly contributions or provisions incentivizing healthy behaviors (6 states as of August 2017), examine, among other things, whether adults enroll in Medicaid programs with these features at the same rates they enroll in Medicaid in other states, and which behavior incentives have the greatest gains.
- **Data challenges:** CMS did not obtain needed data from Indiana, the state that has implemented the strongest incentives. Sufficient data on beneficiary cost sharing and on administrative costs were not available.
- **Implications for scope:** Complete information on Indiana’s experience, such as the effect of a 6-month lock out from Medicaid for failure to pay contributions, will not be included in the evaluation. Questions about how program incentives affect access and utilization are being deferred until the final evaluation. The interim evaluation will also not address the administrative costs incurred with beneficiary engagement policies.

⁴⁸For example, DSRIP program effects may not be apparent immediately. The interim evaluation plans to focus on the more immediate effects of DSRIP programs, including delivery system transformation. Longer-term effects, such as on population health and total costs, will not be addressed until the final evaluation.

Managed long-term services and supports (MLTSS)

- **Focus of evaluation:** For states testing the delivery of long-term services and supports through managed care, examine the effects of these programs on spending, access, and quality of care.
- **Data challenges:** Limitations in the available data, including the quality of managed care encounter data, reduced the number of potential study states from 20 to 2—New York and Tennessee. Sufficient data on the costs of services and on access for New York were not available.
- **Implications for scope:** Evaluation findings for New York and Tennessee cannot be generalized to all MLTSS programs. Key questions on access will not be addressed. Per capita spending as states move from fee-for-service to MLTSS will not be assessed but descriptive information on spending trends will be included.

Source: GAO analysis of Centers for Medicare & Medicaid Services (CMS) contract documents. | GAO-18-220

^aCMS plans to examine additional DSRIP programs, including those in New York and Massachusetts, in the final evaluation report due in 2019.

Many of the data challenges CMS encountered in the multi-state evaluation reflect long-standing concerns with the lack of accurate, complete, and timely Medicaid data. Specifically, we and others have found that data states are required to submit to CMS have, at times, been incomplete or have not been reported at all, particularly managed care encounter data.⁴⁹ Complicating the availability of these data is CMS's ongoing transition to a new data system, the Transformed Medicaid Statistical Information System (T-MSIS), which is CMS's primary effort to improve Medicaid expenditure and utilization data. States' transitions to T-MSIS, however, have introduced substantial delays in state data submissions. For example, by 2015, a large number of states had stopped submitting data through the legacy information system until they established T-MSIS submissions, which meant CMS had to obtain data directly from individual states for the multi-state evaluation.⁵⁰ New data challenges have also emerged as states under demonstrations have enrolled newly eligible beneficiaries in health insurance exchange coverage. Lack of accessible data on beneficiaries enrolled in plans offered through the exchange resulted in the delays in obtaining data for

⁴⁹See GAO, *Medicaid: Program Oversight Hampered by Data Challenges, Underscoring Need for Continued Improvements*, [GAO-17-173](#) (Washington D.C.: Jan. 6, 2017) and: HHS-OIG, *Not All States Reported Medicaid Managed Care Encounter Data as Required*, OEI-07-13-00120 (Washington, D.C.: July 2015).

⁵⁰CMS's legacy data system—the Medicaid Statistical Information System—is a national data system maintained by CMS that collects data from state records on fee-for-service claims for services rendered to Medicaid beneficiaries and managed care encounter records for services delivered through managed care. CMS planned to have all states submitting T-MSIS data, which requires more extensive and frequent data reporting, by July 2014 and stopped requiring states to submit MSIS data in 2015. However, according to HHS's Office of Inspector General, as of December 2016, only 21 states were submitting T-MSIS data. CMS has since made progress collecting T-MSIS data from states. As of November 2017, 49 states have begun submitting T-MSIS data.

Arkansas for the multi-state evaluation.⁵¹ In the past, we have made recommendations to CMS to take action to improve the data available for Medicaid program oversight, including to T-MSIS.⁵²

Indiana's demonstration

Indiana's demonstration—Healthy Indiana Plan 2.0, approved in January 2015—included previously untested provisions that the Centers for Medicare & Medicaid Services stated warranted focused attention. These included:

- monthly contributions for individuals with incomes below federal poverty levels,
- a “lock-out” period for non-payment of monthly contributions for individuals with incomes above federal poverty levels, and
- co-payments above statutory levels for non-urgent use of emergency room services.

Source: GAO. | GAO-18-220

As with the multi-state evaluation, data challenges, particularly obtaining needed data from the state, also proved to be a significant hurdle in CMS's evaluation of Indiana's demonstration. CMS initiated its federal evaluation of Indiana's demonstration in 2015 to understand how the approaches being tested in Indiana's demonstration affected beneficiaries (see sidebar). However, in 2016, Indiana raised concerns about sharing enrollee data with CMS's evaluation contractors. Specifically, in a letter to CMS, the state cited concerns about the controls that CMS had in place to ensure that its contractors would protect enrollee information consistent with state and federal privacy protections.⁵³ Despite assurances by CMS, CMS's contractor and the state were not able to execute a data use agreement. This effectively halted the evaluation's progress. The data use agreement was necessary for the contractor to access state enrollment data that drove a number of planned evaluation activities, including a key beneficiary survey. In October 2017, CMS officials told us that they were continuing to work with the state and anticipated that a data use agreement would be executed and the federal evaluation of Indiana's demonstration would proceed. They did not have timeframes for when the agreement would be reached.

⁵¹Arkansas officials told us that the state was not collecting the specific data from exchange plans requested by the federal evaluator and that collecting those data would result in additional cost. CMS officials stated that, as of October 2017, the agency had resolved the data issues in Arkansas and planned to include data on the state in an updated version of the interim evaluation report that its contractor is required to submit in 2018.

⁵²See [GAO-17-173](#) and *Medicaid: Further Action Needed to Expedite Use of National Data for Program Oversight*, [GAO-18-70](#) (Washington, D.C.: Dec. 8, 2017). As of September 2017, CMS had begun targeted efforts to assess and improve T-MSIS data available for Medicaid program oversight, including initiating a pilot study to identify data anomalies and obtaining input from external experts on data quality. According to CMS officials, CMS's contractor has been reviewing the adequacy of T-MSIS data for the multi-state evaluation and plans to use it.

⁵³In December 2015, Indiana also questioned the need for CMS's single-state evaluation given its own state-led evaluation. One key difference between the evaluations is that the federal evaluation will compare the outcomes from Indiana's demonstration to those for states that expanded Medicaid without a demonstration and to some states that did not expand Medicaid. The state-led evaluation did not attempt such comparisons.

Despite the data challenges and delays, CMS’s evaluations of Medicaid demonstrations, as planned, are likely to provide new information on the effects of demonstrations in different states to inform policy decisions. The multi-state evaluation, for example, is expected to provide information on whether living in a state that collects monthly contributions from beneficiaries affects the likelihood of beneficiaries enrolling in Medicaid and how per-beneficiary spending differs between premium assistance demonstration states and states that have implemented more traditional Medicaid expansions. CMS officials emphasized that federal evaluations allow for cross-state evaluations that can be used to validate the findings of related studies and also to identify which findings are generalizable to other states and populations.

CMS Has Not Released Rapid Cycle Reports and It Is Uncertain When Final Evaluation Results Will Be Available

CMS has yet to make initial reports from the multi-state evaluation publicly available, limiting the potential use of those findings by states and other federal policymakers. As of October 2017, CMS’s contractor had produced 15 rapid cycle reports on states’ progress in implementing demonstrations in the high priority policy areas. These reports provide information on states’ implementation of their demonstrations and variations in design and provide details that can help with the interpretation of evaluation results, inform federal policymaking, and provide lessons learned to states and other stakeholders. The reports also describe policy and other challenges states encountered in implementing their programs, which could be useful to other states interested in replicating these models. (See table 3.)

Table 3: Rapid Cycle Reports Produced as Part of the Multi-State Evaluation, as of October 2017

	Date submitted to CMS	Title of report	Number of states discussed in report
1.	July 2015	Who Enrolls in State Managed Care Programs Covering Long-Term Services and Supports?	22
2.	August 2015	Wraparound Benefits in Premium Assistance Demonstrations	3
3.	August 2015	Beneficiary Engagement Strategies in Medicaid Demonstrations	4
4.	August 2015	Paying for Medicaid Coverage: An Overview of Premium Payments in Section 1115 Demonstrations	5
5.	August 2015	Achieving Coordination of Care to Improve Population Health: Provider Collaboration in Delivery System Reform Incentive Payment Programs	5

	Date submitted to CMS	Title of report	Number of states discussed in report
6.	March 2016	Premium Assistance Demonstrations: The Challenges of Interagency and Public/Private Coordination	3
7.	May 2016	Do Managed Care Programs Covering Long-Term Services and Supports Reduce Waiting Lists for Home and Community-Based Services?	8
8.	July 2016	Incentives to Change Health Behaviors: Beneficiary Engagement Strategies in Indiana, Iowa, and Michigan	3
9.	July 2016	Implementing Measurement Requirements in DSRIP Demonstration Programs	5
10.	September 2016	Educating Beneficiaries about the Cost of Care: Health Accounts in Arkansas, Indiana, and Michigan	3
11.	February 2017	Beneficiary Understanding of Incentives: Evidence from Interim Demonstration Evaluation Reports in Indiana, Iowa, and Michigan	3
12.	March 2017	Attribution in DSRIP Demonstration Programs: A Spotlight on New Jersey and New York	6
13.	August 2017	Continuing with the Same Issuer in Transitions Between Medicaid and the Marketplace: Premium Assistance Models in Arkansas, Iowa, and New Hampshire	3
14.	August 2017	How Do MLTSS Programs Interface with Federal Long-Term Services and Supports-Related Initiatives?	4
15.	September 2017	Designing Medicaid Delivery System Reform Incentive Payment Demonstrations to Reward Better Performance	6

Source: Centers for Medicare & Medicaid Services (CMS). | GAO-18-220

However, despite having received some of these reports from its contractor in 2015, CMS had not released these findings as of October 2017. CMS officials said that the reports were still under agency review and acknowledged that since some of the rapid cycle reports were almost 2 years old, CMS’s contractor was reviewing and updating the information in them. CMS officials noted that the rapid cycle reports had provided useful information and had influenced ongoing work with states designing related demonstrations. For example, according to officials, findings from the rapid cycle reports played a part in how the agency structured the latest DSRIP demonstrations. They also said that rapid cycle reports on beneficiary engagement have shed light on the effectiveness of different beneficiary education strategies, such as what approaches are more successful in capturing beneficiaries’ attention and what strategies are easiest for states to implement. In October 2017, CMS officials stated that they had recently decided to make the rapid cycle reports public, although the agency’s clearance process for the reports was still being decided and the officials did not have timeframes for the reports’ release.

It is also uncertain when CMS will make interim and final evaluation reports from the multi-state evaluation public. By September 2017, CMS’s contractor for the multi-state evaluation produced three interim evaluation

reports covering the four demonstration types.⁵⁴ CMS officials regard these as draft interim evaluation reports, and, as of October 2017, said they were under agency review and would not be publicly released. CMS expects the contractor to submit final interim evaluation reports, which are anticipated to include some additional information beyond the draft reports, by September 2018, about 1 year later than when the final interim evaluation reports were originally due.⁵⁵ CMS officials said that the agency planned to release the final interim evaluation reports, although there was no specific timetable for this. Timeframes for the completion and release of final evaluation results are even more uncertain, both because of the delays in the evaluation progress and because CMS has no standard policy for timeframes for releasing evaluation results.⁵⁶

It is also uncertain when evaluation results will be available and made public for CMS's evaluations of the Indiana and Montana demonstrations. Two years after the approval of the contract for the Indiana evaluation, CMS's contractor has produced an evaluation design but no evaluation findings. CMS had not posted the evaluation design on its website until November 2017, according to officials, about 1 year after it was originally submitted. As discussed above, the lack of findings is due to the contractor and state not having negotiated a data use agreement. To the extent that Indiana's evaluation moves forward and evaluation reports are produced, CMS officials said the agency plans to release the final evaluation report but did not indicate whether interim findings, available a year earlier, would be released.⁵⁷ With regard to the Montana evaluation, CMS expects to receive the interim evaluation report by September 2018 and the final evaluation report by September 2019. How soon these findings would be publicly available, however, is difficult to estimate, as CMS officials told us the agency must review these before making them publically available and does not have timeframes for this review.

⁵⁴Two of the demonstration types—Medicaid-supported enrollment in qualified health plans and beneficiary premiums and engagement—are covered in one report.

⁵⁵For example, CMS's contractor anticipated that the final interim evaluation reports will include information on Arkansas's experience with Medicaid-supported exchange coverage.

⁵⁶According to the contract for the multi-state evaluation, final evaluation reports are due in June 2019.

⁵⁷CMS has pushed back the contract delivery date for the interim evaluation report to September 2018.

The lack of a standard policy for the public release of findings from federal evaluations of Medicaid demonstrations is inconsistent with recommendations of the American Evaluation Association. The Association recommends that evaluation findings related to public accountability be disseminated to the public, and that evaluation results be made available in a timely manner and be easily accessible through the internet.⁵⁸ For state-led evaluations, CMS must post on its website, or provide a link to the state's website, all evaluation materials, including research and data collection, for the purposes of sharing findings with the public within 30 days of receiving the materials. CMS has not established a comparable policy for the release of findings from federal evaluations of demonstrations. CMS officials stated that federal evaluations provide a unique cross-state perspective that states typically do not have the capacity to provide in their own state-led evaluations; however, if these reports are not made public in a timely fashion, opportunities may be missed to inform federal and state policymakers and other stakeholders on the effects of Medicaid demonstrations.

Conclusions

Section 1115 demonstrations have long been an important tool for providing states with the flexibility to test new approaches to providing and financing Medicaid coverage. Given the potential effects on millions of beneficiaries and significant federal investment in these demonstrations—over \$100 billion in 2015—it is critical that they be evaluated. Evaluating Medicaid demonstrations is complex, both within a single state and across states. These programs are dynamic, and there are many factors affecting outcomes, making it challenging to isolate the effects of policy changes implemented under a demonstration. Further, persistent challenges with Medicaid data that we have highlighted over the years add to the complexity of evaluating demonstrations. Despite these challenges, targeted and well-designed evaluations offer the potential to identify policies that improve outcomes for beneficiaries and reduce costs to Medicaid. With the growing complexity of Medicaid programs and limited resources, that information could prove key in helping to sustain the program.

⁵⁸American Evaluation Association, "Evaluation Roadmap."

CMS's approach to overseeing state-led evaluations in the past has resulted in limited information about the effects of demonstrations, leaving gaps in evidence about policies that might improve state Medicaid programs. CMS's efforts since 2014 to improve the usefulness of evaluations in informing state and federal Medicaid policy decisions have promise. If CMS consistently sets and enforces clear expectations and provides support for rigorous and timely state-led evaluations for all demonstrations as planned, those evaluations could yield more useful information within the next several years. However, CMS has not established written procedures for requiring final, comprehensive evaluation reports at the end of each cycle for all demonstrations, a key step in improving the usefulness of state-led evaluations. Further, CMS is planning to allow less rigorous evaluations for some demonstrations but has not yet established specific criteria for doing so.

Federal evaluations led by CMS also show promise. The evaluations currently underway—despite challenges that caused delays and reduced scope—are likely to provide a cross-state look at the effects of policies that are of great interest to CMS, Congress, and other states. However, CMS has not yet made potentially useful rapid cycle reports public and has no established policy for making future evaluation reports public. By not making the results of the federal evaluations public in a timely manner, CMS is missing an opportunity to inform important policy discussions happening at the state and federal levels.

Recommendations for Executive Action

We are making the following three recommendations to CMS:

The Administrator of CMS should establish written procedures for implementing the agency's policy that requires all states to submit a final evaluation report after the end of each demonstration cycle, regardless of renewal status. (Recommendation 1)

The Administrator of CMS should issue written criteria for when CMS will allow limited evaluation of a demonstration or a portion of a demonstration, including defining conditions, such as what it means for a demonstration to be longstanding or noncomplex, as applicable. (Recommendation 2)

The Administrator of CMS should establish and implement a policy for publicly releasing findings from federal evaluations of demonstrations, including findings from rapid cycle, interim, and final reports; and this

policy should include standards for timely release. (Recommendation 3)

Agency Comments and Our Evaluation

We provided a draft of this report to HHS for review and comment. HHS concurred with all three recommendations. Regarding our first recommendation that CMS establish written procedures for implementing its policy requiring states to submit final evaluation reports after the end of each demonstration cycle, HHS said that it is in the process of developing such written procedures. HHS said that it is currently making this a requirement through the STCs for each demonstration as demonstrations are approved or renewed. Regarding our second recommendation that CMS issue written criteria for when the agency will allow states to limit evaluations of their demonstrations, HHS said it is in the process of testing such criteria, and that once it has experience with the criteria, it will develop written guidance. Regarding our third recommendation that CMS establish and implement a policy for publicly releasing findings from federal evaluations of demonstrations, HHS said that CMS is in the process of establishing such a policy. HHS added that CMS plans to have all finalized federal rapid cycle reports and final interim evaluation reports publicly available in the near future.

HHS also provided technical comments, which we incorporated as appropriate. HHS's comments are reproduced in appendix II.

As agreed with your offices, unless you publicly announce the contents of this report earlier, we plan no further distribution until 30 days from the report date. At that time, we will send copies of this report to the Secretary of Health and Human Services, appropriate congressional committees, and other interested parties. The report will also be available at no charge on the GAO website at <http://www.gao.gov>.

If you or your staff members have any questions about this report, please contact me at (202) 512-7114 or iritanik@gao.gov. Contact points for our Offices of Congressional Relations and Public Affairs may be found on the last page of this report. GAO staff who made key contributions to this report are listed in appendix III.

A handwritten signature in black ink that reads "Katherine Iritani". The signature is written in a cursive style with a large, prominent initial "K".

Katherine M. Iritani
Director, Health Care

List of Requesters

The Honorable Orrin G. Hatch
Chairman
Committee on Finance
United States Senate

The Honorable Lamar Alexander
Chairman
Committee on Health, Education, Labor & Pensions
United States Senate

The Honorable Greg Walden
Chairman
Committee on Energy and Commerce
House of Representatives

The Honorable Michael Burgess
Chairman
Subcommittee on Health
Committee on Energy and Commerce
House of Representatives

The Honorable Fred Upton
House of Representatives

Appendix I: Characteristics of Selected States' 1115 Demonstrations

The Medicaid section 1115 demonstrations (referred to as demonstrations) in our eight selected states varied in terms of the number of years the demonstrations had been in effect and cost, among other things. For example, three of the more mature demonstrations—those in Maryland, Massachusetts, and New York—had been in place for two decades. Demonstrations in Arkansas and Kansas represented more recent approvals, both approved in 2013. (See table 4.) With regard to cost, all of the selected states were among the top 15 states in terms of amount of spending under demonstrations. Together, spending under demonstrations in our selected states accounted for about 47 percent of all spending under demonstrations in fiscal year 2015.

Table 4: Characteristics of Selected States' Medicaid Section 1115 Demonstrations Selected for Our Review

State	Demonstration start year	Scheduled demonstration expiration	Fiscal year 2015	
			Expenditures for demonstrations (billions)	Percent of Medicaid expenditures for demonstrations
Arizona	2011 ^a	September 2021	10.6	99.7
Arkansas	2013	December 2021	1.6	29.8
California	2005	December 2020	42.6	50.1
Indiana	2008 ^b	January 2018	2.1	22.5
Kansas	2013	December 2018	2.8	94.2
Maryland	1996	December 2021	6.6	70.1
Massachusetts	1997	June 2022	9.5	62.1
New York	1997	March 2021	1.8 ^c	3.0 ^c

Source: GAO review of documentation and data from the Centers for Medicare & Medicaid Services. | GAO-18-220

Notes: Medicaid section 1115 demonstrations are typically approved for an initial 5-year cycle and may be renewed for subsequent cycles. In the table, the demonstration start year notes the date when the demonstration was first approved and the demonstration expiration notes the date the demonstration is expected to expire based on the most recently approved renewal, as of October 2017. The start dates reflect the dates of the state's most comprehensive demonstration. Demonstration expenditures represent expenditures reported by states for fiscal year 2015 as of October 3, 2016. Expenditures reflect spending for all Medicaid section 1115 demonstrations in states that had more than one demonstration. Expenditures represent federal and state spending for medical costs. They do not include administrative costs.

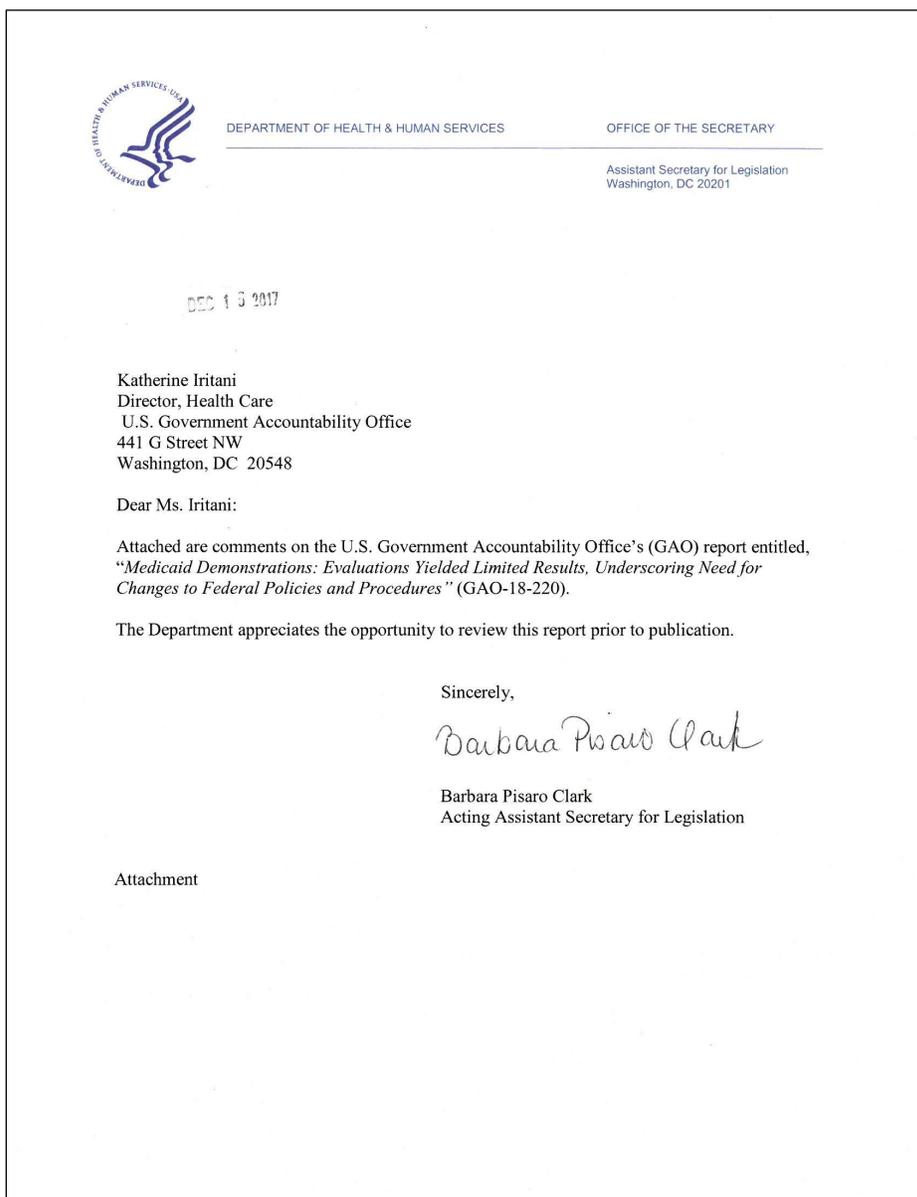
**Appendix I: Characteristics of Selected States'
1115 Demonstrations**

^aArizona has operated the Arizona Health Care Cost Containment System through a section 1115 demonstration since 1982. In 2011, Arizona received approval for a new demonstration to continue the program and also add new demonstration authorities.

^bIndiana's demonstration, referred to as Healthy Indiana Plan, became effective in 2008. In 2015, CMS approved an expanded version of the demonstration, referred to as Healthy Indiana Plan 2.0.

^cExpenditure data for New York were incomplete at the time of our analysis, therefore these data may underestimate demonstration spending as a proportion of total Medicaid spending. In fiscal year 2014, demonstration expenditures for New York totaled about \$19 billion, making the state the third highest in the amount of demonstration expenditures.

Appendix II: Comments from the Department of Health and Human Services



GENERAL COMMENTS OF THE DEPARTMENT OF HEALTH AND HUMAN SERVICES (HHS) ON THE GOVERNMENT ACCOUNTABILITY OFFICE'S DRAFT REPORT ENTITLED – MEDICAID DEMONSTRATIONS: EVALUATIONS YIELDED LIMITED RESULTS, UNDERSCORING NEED FOR CHANGES TO FEDERAL POLICIES AND PROCEDURES (GAO-18-220)

The U.S. Department of Health and Human Services (HHS) appreciates the opportunity to review and comment on the Government Accountability Office's (GAO) draft report on Medicaid Section 1115 demonstrations. HHS takes seriously its role in oversight of Medicaid demonstrations and advancing understanding of program innovation impacts to inform policy making.

Section 1115 of the Social Security Act gives the Secretary of HHS authority to approve experimental, pilot, or demonstration projects that are likely to promote the objectives of Medicaid and/or the Children's Health Insurance Program (CHIP). As noted in the draft report, the Secretary has delegated this approval authority to the Centers for Medicare & Medicaid Services (CMS). These demonstrations give states additional flexibility to design and evaluate innovative policy approaches to improve their Medicaid programs and use innovative service delivery systems that can be expected to improve care, increase efficiency, and reduce costs.

Evaluations are used to determine whether demonstrations are having their intended effects, and are conducted at both the state and the Federal level. Evaluations provide unique and important contributions in understanding the effects of 1115 demonstration policy developments. For example, they can provide insight into implementation challenges that can refine the state's demonstration policy and inform the efforts of other states that are contemplating similar innovations.

During the demonstration application process, HHS outlines key evaluation requirements with the state, and since 2014, HHS has taken important steps to improve the quality of state-led evaluations. For example, the evaluation requirements, including timelines for submission of evaluation reports to HHS, are included in the Special Terms and Conditions for an approved demonstration. Additionally, for demonstration renewals, HHS has begun to require that states use the prevailing standards of scientific and academic rigor in their evaluation reports, such as using the best available data and using controls and adjustments for reporting the limitations of data and their effect on results.

The recently released CMS informational bulletin related to demonstrations under section 1115 emphasizes flexibility in obtaining approval of demonstrations and outlines expectations of stronger accountability for monitoring and evaluation.¹ The informational bulletin also contains guidance on additional enhancements required for state-led evaluations. These enhancements include encouraging states to use existing nationally recognized metric sets and quality measures, and the continued availability of technical assistance from CMS in developing milestones and performance metrics.

In an effort to address data availability challenges that can hinder evaluations, HHS is working to transition all states to T-MSIS and has made significant progress. As of November 21, 2017, 49

¹ <https://www.medicaid.gov/federal-policy-guidance/downloads/cib110617.pdf>

GENERAL COMMENTS OF THE DEPARTMENT OF HEALTH AND HUMAN SERVICES (HHS) ON THE GOVERNMENT ACCOUNTABILITY OFFICE'S DRAFT REPORT ENTITLED – MEDICAID DEMONSTRATIONS: EVALUATIONS YIELDED LIMITED RESULTS, UNDERSCORING NEED FOR CHANGES TO FEDERAL POLICIES AND PROCEDURES (GAO-18-220)

states have begun submitting T-MSIS data. These 49 entities represent 96 percent of the Medicaid and CHIP population. With a majority of states submitting T-MSIS data, HHS has begun to develop tools for T-MSIS users, as well as work with states to improve the quality of data submitted. For example, HHS is developing a data quality scorecard for users, which aggregates data quality findings in a user-friendly tool. In addition, HHS has developed a data quality database that houses post-production data quality findings to help develop solutions for common problems seen across multiple states. These efforts will help states report complete and comparable T-MSIS data and will support Federal evaluations moving forward.

GAO's recommendations and HHS' responses are below.

Recommendation

The Administrator of CMS should develop written procedures for implementing the agency's policy that requires all states to submit a final evaluation report after the end of each demonstration cycle, regardless of renewal status.

HHS Response

HHS concurs with this recommendation. HHS is in the process of developing written procedures requiring all states to submit a final evaluation report after the end of each demonstration cycle, regardless of renewal status. HHS is currently making this a requirement through the Special Terms and Conditions for each demonstration as new demonstrations are approved or existing demonstrations are renewed.

Recommendation

The Administrator of CMS should issue written criteria for when CMS will allow limited evaluation of a demonstration or a portion of a demonstration, including defining conditions, such as what it means for a demonstration to be longstanding or noncomplex, as applicable.

HHS Response

HHS concurs with this recommendation. HHS is in the process of refining and testing criteria for when CMS will allow modified evaluation requirements for a demonstration. Once HHS has experience with these criteria, HHS will develop written guidance.

Recommendation

The Administrator of CMS should establish and implement a policy for publicly releasing findings from federal evaluations or demonstrations, including findings from rapid cycle, interim, and final reports; and this policy should include standards for timely release.

HHS Response

HHS concurs with this recommendation. CMS is in the process of establishing a policy for publicly releasing findings from federal evaluations for section 1115 demonstrations. CMS plans to have all finalized federal rapid cycle reports and final interim evaluations that have already been produced publicly available in the near future. While CMS agrees that there is important

GENERAL COMMENTS OF THE DEPARTMENT OF HEALTH AND HUMAN SERVICES (HHS) ON THE GOVERNMENT ACCOUNTABILITY OFFICE'S DRAFT REPORT ENTITLED – MEDICAID DEMONSTRATIONS: EVALUATIONS YIELDED LIMITED RESULTS, UNDERSCORING NEED FOR CHANGES TO FEDERAL POLICIES AND PROCEDURES (GAO-18-220)

purpose in and need for releasing the federal reports, CMS has been using these reports to inform policy work.

Appendix III: GAO Contact and Staff Acknowledgments

GAO Contact

Katherine M. Iritani, (202) 512-7114 or iritanik@gao.gov

Staff Acknowledgments

In addition to the contact named above, Susan Barnidge (Assistant Director), Linda McIver (Analyst-in-Charge), John Lalomio, Hannah Locke, and Corissa Kiyon-Fukumoto made key contributions to this report. Also contributing were Laurie Pachter and Emily Wilson.

Appendix IV: Accessible Data

Agency Comment Letter

Text of Appendix II: Comments from the Department of Health and Human Services

Page 1

Katherine Iritani Director, Health Care

U.S. Government Accountability Office 441 G Street NW

Washington, DC 20548 Dear Ms. Iritani:

Attached are comments on the U.S. Government Accountability Office's (GAO) report entitled, "Medicaid Demonstrations: Evaluations Yielded Limited Results, Underscoring Need for Changes to Federal Policies and Procedures" (GAO-18-220).

The Department appreciates the opportunity to review this report prior to publication.

Sincerely ,

Barbara Pisaro Clark

Acting Assistant Secretary for Legislation

Attachment

Page 2

The U.S. Department of Health and Human Services (HHS) appreciates the opportunity to review and comment on the Government Accountability Office's (GAO) draft report on Medicaid Section 1115 demonstrations. HHS takes seriously its role in oversight of Medicaid demonstrations and advancing understanding of program innovation impacts to inform policy making.

Section 1115 of the Social Security Act gives the Secretary of HHS authority to approve experimental, pilot, or demonstration projects that are likely to promote the objectives of Medicaid and/or the Children's Health Insurance Program (CHIP). As noted in the draft report, the Secretary has delegated this approval authority to the Centers for Medicare & Medicaid Services (CMS). These demonstrations give states additional flexibility to design and evaluate innovative policy approaches to improve their Medicaid programs and use innovative service delivery systems that can be expected to improve care, increase efficiency, and reduce costs.

Evaluations are used to determine whether demonstrations are having their intended effects, and are conducted at both the state and the Federal level. Evaluations provide unique and important contributions in understanding the effects of 1115 demonstration policy developments. For example, they can provide insight into implementation challenges that can refine the state's demonstration policy and inform the efforts of other states that are contemplating similar innovations.

During the demonstration application process, HHS outlines key evaluation requirements with the state, and since 2014, HHS has taken important steps to improve the quality of state-led evaluations. For example, the evaluation requirements, including timelines for submission of evaluation reports to HHS, are included in the Special Terms and Conditions for an approved demonstration. Additionally, for demonstration renewals, HHS has begun to require that states use the prevailing standards of scientific and academic rigor in their evaluation reports, such as using the best available data and using controls and adjustments for reporting the limitations of data and their effect on results.

The recently released CMS informational bulletin related to demonstrations under section 1115 emphasizes flexibility in obtaining approval of demonstrations and outlines expectations of stronger accountability for monitoring and evaluation.¹ The informational bulletin also contains guidance on additional enhancements required for state-led evaluations. These enhancements include encouraging states to use existing nationally recognized metric sets and quality measures, and the continued availability of technical assistance from CMS in developing milestones and performance metrics.

¹ <https://www.medicaid.gov/federal-policy-guidance/downloads/cib110617.pdf>

In an effort to address data availability challenges that can hinder evaluations, HHS is working to transition all states to T-MSIS and has made significant progress. As of November 21, 2017, 49

Page 3

states have begun submitting T-MSIS data. These 49 entities represent 96 percent of the Medicaid and CHIP population. With a majority of states submitting T-MSIS data, HHS has begun to develop tools for T-MSIS users, as well as work with states to improve the quality of data submitted. For example, HHS is developing a data quality scorecard for users, which aggregates data quality findings in a user-friendly tool. In addition, HHS has developed a data quality database that houses post-production data quality findings to help develop solutions for common problems seen across multiple states. These efforts will help states report complete and comparable T-MSIS data and will support Federal evaluations moving forward.

GAO's recommendations and HHS' responses are below.

Recommendation

The Administrator of CMS should develop written procedures for implementing the agency's policy that requires all states to submit a final evaluation report after the end of each demonstration cycle, regardless of renewal status.

HHS Response

HHS concurs with this recommendation. HHS is in the process of developing written procedures requiring all states to submit a final evaluation report after the end of each demonstration cycle, regardless of renewal status. HHS is currently making this a requirement through the Special Terms and Conditions for each demonstration as new demonstrations are approved or existing demonstrations are renewed.

Recommendation

The Administrator of CMS should issue written criteria for when CMS will allow limited evaluation of a demonstration or a portion of a demonstration, including defining conditions, such as what it means for a demonstration to be longstanding or noncomplex, as applicable.

HHS Response

HHS concurs with this recommendation. HHS is in the process of refining and testing criteria for when CMS will allow modified evaluation requirements for a demonstration. Once HHS has experience with these criteria, HHS will develop written guidance.

Recommendation

The Administrator of CMS should establish and implement a policy for publicly releasing findings from federal evaluations or demonstrations, including findings from rapid cycle, interim, and final reports; and this policy should include standards for timely release.

HHS Response

HHS concurs with this recommendation. CMS is in the process of establishing a policy for publicly releasing findings from federal evaluations for section 1115 demonstrations. CMS plans to have all finalized federal rapid cycle reports and final interim evaluations that have already been produced publicly available in the near future. While CMS agrees that there is important

Page 4

purpose in and need for releasing the federal reports, CMS has been using these reports to inform policy work.

GAO's Mission

The Government Accountability Office, the audit, evaluation, and investigative arm of Congress, exists to support Congress in meeting its constitutional responsibilities and to help improve the performance and accountability of the federal government for the American people. GAO examines the use of public funds; evaluates federal programs and policies; and provides analyses, recommendations, and other assistance to help Congress make informed oversight, policy, and funding decisions. GAO's commitment to good government is reflected in its core values of accountability, integrity, and reliability.

Obtaining Copies of GAO Reports and Testimony

The fastest and easiest way to obtain copies of GAO documents at no cost is through GAO's website (<http://www.gao.gov>). Each weekday afternoon, GAO posts on its website newly released reports, testimony, and correspondence. To have GAO e-mail you a list of newly posted products, go to <http://www.gao.gov> and select "E-mail Updates."

Order by Phone

The price of each GAO publication reflects GAO's actual cost of production and distribution and depends on the number of pages in the publication and whether the publication is printed in color or black and white. Pricing and ordering information is posted on GAO's website, <http://www.gao.gov/ordering.htm>.

Place orders by calling (202) 512-6000, toll free (866) 801-7077, or TDD (202) 512-2537.

Orders may be paid for using American Express, Discover Card, MasterCard, Visa, check, or money order. Call for additional information.

Connect with GAO

Connect with GAO on [Facebook](#), [Flickr](#), [LinkedIn](#), [Twitter](#), and [YouTube](#).
Subscribe to our [RSS Feeds](#) or [E-mail Updates](#). Listen to our [Podcasts](#).
Visit GAO on the web at www.gao.gov and read [The Watchblog](#).

To Report Fraud, Waste, and Abuse in Federal Programs

Contact:

Website: <http://www.gao.gov/fraudnet/fraudnet.htm>
E-mail: fraudnet@gao.gov
Automated answering system: (800) 424-5454 or (202) 512-7470

Congressional Relations

Katherine Siggerud, Managing Director, siggerudk@gao.gov, (202) 512-4400,
U.S. Government Accountability Office, 441 G Street NW, Room 7125,
Washington, DC 20548

Public Affairs

Chuck Young, Managing Director, youngc1@gao.gov, (202) 512-4800
U.S. Government Accountability Office, 441 G Street NW, Room 7149
Washington, DC 20548

Strategic Planning and External Liaison

James-Christian Blockwood, Managing Director, spel@gao.gov, (202) 512-4707
U.S. Government Accountability Office, 441 G Street NW, Room 7814,
Washington, DC 20548