February 2020

Federal Meta-Analysis Support: Section 1115 Substance Use Disorder Demonstrations

Evaluation Design Report

Prepared for

Michael Trieger
Centers for Medicare & Medicaid Services
7500 Security Boulevard, Mail Stop B3-03-30
Baltimore, MD 21244-1850

Prepared by

Anupa Bir, ScD MPH
Susan Haber, ScD
Jesse Hinde, PhD
Brian Bruen, PhD
Tami Mark, PhD
Jim Derzon, PhD
Kevin Smith, MA
Kyle Emery, MS
RTI International
3040 E. Cornwallis Road
Research Triangle Park, NC 27709

RTI Project Number 0214448.001.012.001.002

RTI INTERNATIONAL

Federal Meta-Analysis Support: Section 1115 Substance Use Disorder Demonstrations

Evaluation Design Report

February 2020

Prepared for

Michael Trieger
Centers for Medicare & Medicaid Services
7500 Security Boulevard, Mail Stop B3-03-30
Baltimore, MD 21244-1850

Prepared by

Anupa Bir, ScD MPH
Susan Haber, SCD
Jesse Hinde, PhD
Brian Bruen, PhD
Tami Mark, PhD
Jim Derzon, PhD
Kevin Smith, MA
Kyle Emery, MS
RTI International
3040 E. Cornwallis Road
Research Triangle Park, NC 27709

CONTENTS

APPENDICES

LIST OF FIGURES

Number **Page**

LIST OF TABLES

EXECUTIVE SUMMARY

Section 1115 Medicaid demonstrations offer vehicles for states to design, implement, and test new approaches that have the potential to improve the Medicaid program's effectiveness, promote the health and well-being of the low-income individuals they serve, and shape new policy directions at the federal level. To learn from each Medicaid section 1115 demonstration and the groups of such demonstrations with similar features, the Centers for Medicare & Medicaid Services (CMS) has commissioned the Federal Meta-Analysis Support contract. Under this contract, RTI International will work with CMS to conduct meta-evaluations of selected groups of Medicaid section 1115 demonstrations. These meta-evaluations will compare experiences among similar section 1115 demonstrations across states to understand the overall effectiveness of the demonstrations and how variations in state demonstration features and the context in which they are implemented contribute to differences in effectiveness. States also have the option of pursuing similar objectives through state plan amendments. As states take up state plan amendments for substance use disorder (SUD) purposes, such as amendments under Section 5052 of the Substance Use-Disorder Prevention that Promotes Opioid Recovery and Treatment (SUPPORT) for Patients and Communities Act, RTI will work with CMS and states toward understanding their activities.

Each state is required to evaluate its section 1115 demonstration and submit monitoring reports to CMS. States pursuing state plan amendments do not have an evaluation requirement, but efforts will be made to collaborate with states choosing this option to gather information and understand their experiences. The meta-analyses will primarily use data from state demonstration monitoring and evaluation reports, augmented with limited stakeholder interviews and analyses of national surveys, Transformed Medicaid Statistical Information System (T-MSIS) data, and other national data sets. Through this work, RTI will collaborate with CMS and its other contractors to provide information, including best practices and recommendations on improving demonstration policy and implementation strategies, to inform national policy making and to support scaling up and diffusion of successful policies. An additional goal of this project is to inform CMS on the rigor and limitations of state evaluations to support further improvements in CMS evaluation guidance for section 1115 demonstrations. By combining an in-depth look at the context, implementation, features, and outcomes across section 1115 demonstrations, the meta-analysis will complement state evaluations and provide added value to CMS.

Section 1115 Demonstrations

Identifying innovative approaches to address high-priority challenges is of critical interest to state Medicaid programs and their federal partner, CMS. States have long been allowed to waive certain Medicaid requirements to explore new approaches to service delivery and payment. Since the inception of Medicaid, states have been able to use research and demonstration waivers authorized under section 1115(a) of the Social Security Act to test new or existing approaches to financing and delivering services under Medicaid, while maintaining the overall goals of the Medicaid program and budget neutrality within the program. CMS has announced its commitment to supporting state innovations in their Medicaid programs and allowing flexibility for states to adapt demonstration design to reflect the uniqueness of their covered populations, resources, and policy goals.

This Evaluation Design Report for the Federal Meta-Analysis Support contract focuses on section 1115 SUD demonstrations and state plan amendments in the case of states choosing to use section 5052 of the SUPPORT Act. These approaches allow states to develop and expand comprehensive treatment strategies to address opioid use disorder (OUD) and other SUDs to increase rates of identification, initiation, and engagement in SUD treatment; increase adherence to and retention in SUD treatment; reduce overdose deaths; reduce the number of SUD-related emergency department (ED) visits and inpatient hospital admissions; reduce the number of readmissions to the same or higher level of SUD care; and improve access to care for physical health and comorbid conditions.

Overview of Meta-Analytic Approach

The meta-analysis of SUD demonstrations will document and explore variation in state baseline conditions and demonstration design, approach, and implementation to explain differences in outcomes observed across demonstrations. Each state with a section 1115 demonstration is required to conduct its own state evaluation per 42 CFR § 431.424. While state evaluations provide valuable information on demonstration experience and outcomes in individual states, the meta-analysis will synthesize data from state evaluations and other sources to understand lessons learned across states from all the related section 1115 demonstrations. The meta-evaluation approach planned will provide CMS and states with a deeper understanding of what levers affect successful outcomes—both implementation and impacts—as well as whether, under what conditions, and how these initiatives would best be replicated in other states.

Meta-analysis incorporates synthesis of qualitative and quantitative data. To support the meta-analysis, we will compile a cross-state database that includes all application documents,

implementation and evaluation plans, monitoring documents, evaluation documents, and complementary analyses of monitoring data and publicly available data conducted by RTI. The meta-analysis will incorporate several analytic methods. We will conduct qualitative analysis of secondary data and limited primary data collected from demonstration states to document demonstration implementation and contextual features that will be used in quantitative analyses. Qualitative data will also be used for cross-state syntheses that take a deep dive into targeted demonstration design and implementation topics. We will conduct both descriptive and multivariate analyses of demonstration outcomes. We will analyze descriptive data on demonstration implementation and impacts from state monitoring reports to compare outcomes across groups of states whose demonstrations share common features. We will use forest plots and meta-regression of effect sizes reported by state evaluations to estimate the overall impact of demonstrations across states and identify factors that likely explain differences in state impacts. We may also use qualitative case comparative methods to explore factors related to successful implementation and impacts. We will conduct supplemental descriptive and regression analyses using national data sets as needed to complement implementation and impact analyses from state monitoring and evaluation data. We will develop an interactive dashboard that integrates data for multiple time periods from monitoring reports, state evaluations, and supplemental national data sources to facilitate comparisons across states with common features. Our results will also be shared through a series of reports on Medicaid.gov.

SECTION 1.
INTRODUCTION

1.1 Overview of Meta-Analysis Support Contract Goals

The Centers for Medicare & Medicaid Services (CMS) has five objectives for the meta-analyses of section 1115 demonstrations:

1. Using and exploring available state and federal data, including the Transformed Medicaid Statistical Information System (T-MSIS) and other sources, to study the effectiveness of Medicaid section 1115 SUD demonstrations and compare the effectiveness of section 1115 demonstration policies and features across states

2. Provide information, including best practices and recommendations on improving demonstration policy and implementation strategies, to inform national policy making and to support scaling up and diffusion of successful demonstration policy experiments

3. Provide materials for and participate in CMS Learning Collaboratives

4. Cooperate with CMS, its evaluation and other contractors, and other federal agencies to share data and provide information to improve overall understanding of any related studies

5. Inform CMS on the rigor and limitations of state evaluation designs and reports, as well as monitoring protocols and reports, to support further improvements and capacity building in CMS monitoring and evaluation of section 1115 demonstrations.

This evaluation design focuses on section 1115 SUD demonstrations and state plan amendments in the case of states choosing to use section 5052 of the SUPPORT Act. Separate evaluation designs will be prepared for other types of section 1115 demonstrations as requested by CMS.

1.2 Overview of Section 1115 Substance Use Disorder Demonstrations

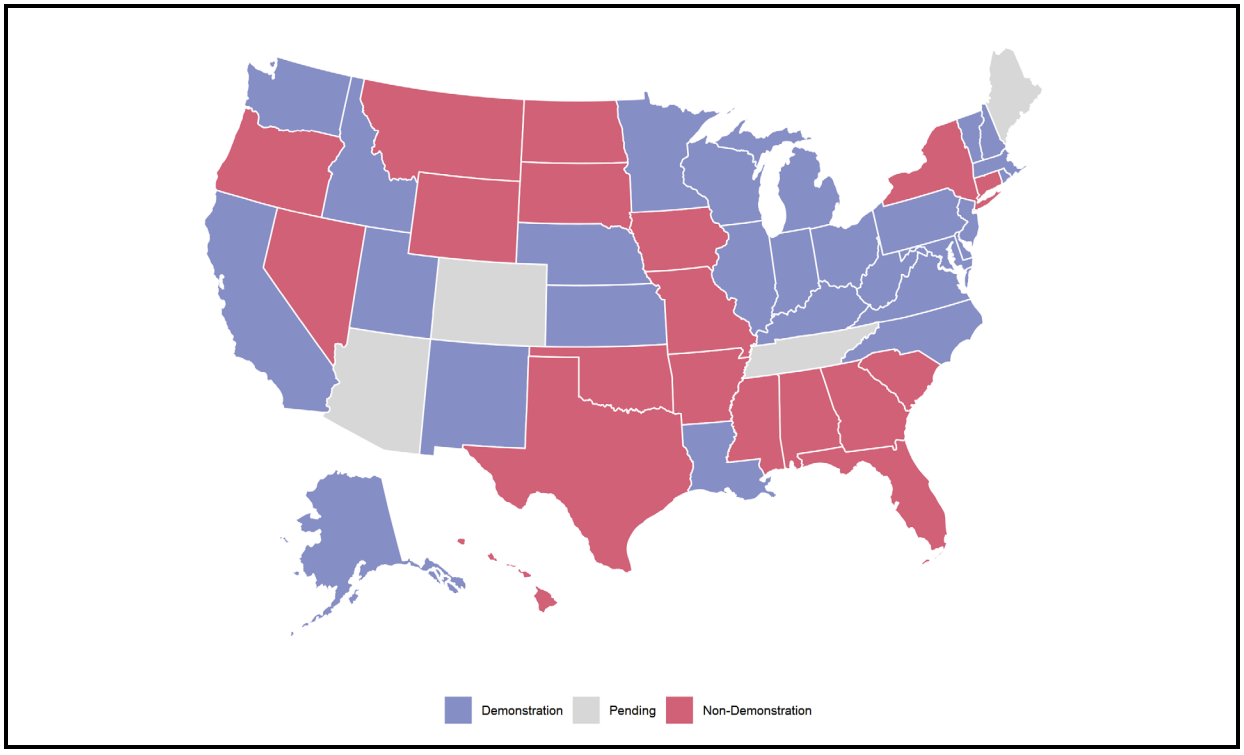
In response to the opioid epidemic, CMS and other federal and state agencies face a significant challenge in increasing the availability of high-quality substance use disorder (SUD) treatment services. Over the last two decades, opioid-related drug overdose death rates have increased by 200 percent[1,2] and recent data suggest that nearly 1 percent of the U.S. population has a SUD related to prescription pain relievers, often opioids.[3] Most individuals with opioid use disorder (OUD) do not receive any treatment, and even more receive inadequate treatment.[4] Medicaid has become an important source of insurance coverage for individuals with OUD. Approximately 28 percent of individuals who report opioid misuse in the past year were covered by Medicaid.[5]

However, states historically have had limited Medicaid SUD benefits and covered limited treatment options. To address this, section 1115 SUD demonstrations offer states flexibility to improve access to and quality of treatment for Medicaid beneficiaries with OUD and other SUDs. Starting in 2015, CMS offered states the opportunity to test Medicaid coverage of a full SUD treatment service array in the context of overall SUD service delivery transformation, provided participating states met specific requirements.[6] These requirements included the following: comprehensive evidence-based design; adoption of widely-accepted appropriate standards of care (e.g., American Society of Addiction Medicine [ASAM] criteria); strong network development; care coordination; integration of physical and SUD treatment; program integrity safeguards; benefit management; community integration; strategies to address prescription drug misuse and OUD; services to youth and adolescents with SUD; quality metrics reporting; and collaboration with the single state agency for substance abuse services.

States that implemented these requirements were allowed to receive federal financial participation (FFP) for the continuum of services to treat OUD and other SUDs including services delivered to beneficiaries in institutions for mental disease (IMDs), which normally are ineligible for FFP if the facility has more than 16 beds. The demonstration requires that the average length of stay for residential and inpatient stays be 30 days.

CMS modified the requirements for SUD section 1115 demonstrations in November 2017 to improve access to clinically appropriate treatment for OUD and other SUDs, to better support states in the development and expansion of comprehensive treatment strategies, and to incorporate improved progress and outcome monitoring.[7] The modifications added flexibility to the SUD demonstration requirements, including allowing states to implement an alternative or modified set of patient placement criteria that are widely recognized as representing an expert consensus on appropriate treatment. The revised policy incorporated provisions related to health information technology (IT) and emphasized that states ensure the availability of medication-assisted treatment (MAT) in residential treatment settings. In addition, it gave states 2 years after demonstration approval to meet the requirement for comprehensive services and an implementation plan approved by CMS for meeting this requirement. A state cannot receive federal financial participation for IMD services until its implementation plan is approved. As of May 1st, 2020, 27 states and the District of Columbia had received approval for section 1115 SUD demonstrations and several other states had pending proposals (***Figure 1-1***).

Figure 1-1. Approved and Pending Section 1115 SUD Demonstrations as of May 1, 2020



Sources: RTI review of CMS documents, Medicaid.gov, and state documents.

1.3 Other Approaches to Expanding SUD Services in Medicaid

Section 5052 of the Substance Use-Disorder Prevention that Promotes Opioid Recovery and Treatment (SUPPORT) for Patients and Communities Act of 2018 offers states an additional option to address SUDs through a state plan amendment (SPA), rather than a section 1115 demonstration. Like the section 1115 demonstration, states can receive FFP for SUD residential treatment in IMDs. However, the features and conditions for the SPAs differ from the section 1115 demonstrations in several important ways. First, residential stays under the SPA authority are limited to a maximum of 30 days. Second, there is a much more specific and demanding maintenance of effort provision for the SPA. Third, while both section 1115 demonstrations and the SPA require states to offer a full continuum of SUD services, the SPA requires the continuum of services to be in place at approval whereas the section 1115 demonstrations allow for a two-year implementation ramp-up. The SPA continuum also requires provision of early intervention and both intensive outpatient and partial hospitalization services, whereas early intervention services are encouraged, and states need only one of intensive outpatient and partial hospitalization services under the section 1115 demonstrations. Section 5052 also requires that participating IMDs be able to provide care at a lower level of intensity or have an established relationship with another facility that is able to provide care at a lower level of intensity. It is

theoretically possible that the same state will have overlapping section 5052 and section 1115 demonstrations, though the current demonstrations do not yet include this situation. Finally, states with a SPA are not required to conduct an evaluation of their program or submit monitoring data to CMS.

We will include the 5052 SPA states in our meta-analyses to the extent that we can access their data. To date, one state (Idaho) has submitted a 5052 SPA for approval, and CMS discussions continue with other interested states.

1.4 Overview of Meta-Analysis Activities

States are required to monitor and evaluate their section 1115 demonstrations per 42 CFR §431.424. CMS has enhanced the expectations for scientific rigor in state evaluations and has increased supports to states in recent years. Monitoring provides early and ongoing information about demonstration implementation and progress toward goals that CMS and states can use to identify potential problems and to make mid-course adjustments if needed. Evaluations measure implementation outcomes and demonstration impacts and provide an evidentiary base for CMS and states to inform modification of program design and implementation choices. Other states considering a demonstration can use evaluation findings to shape the design of their own demonstrations. Finally, evaluations provide evidence to support decisions about whether demonstrations should be authorized to continue, whether new demonstrations should be approved, if demonstration requirements should be modified, and, ultimately, whether federal Medicaid policy should be changed.

Individual state demonstration monitoring and evaluation results can provide valuable information about program implementation experience and impacts. However, because of the single-state focus, they are not well positioned to provide overarching lessons about what works best to improve care for OUD and other SUDs that can be used to shape broader state and federal Medicaid policies. In contrast with individual state evaluations, meta-evaluation of the cross-state data is concerned with variations in demonstration design and context that can impact an individual demonstration's outcomes and explain variation in outcomes observed **across** demonstrations. The meta-evaluation of the SUD demonstrations will draw on experience across multiple states implementing differing activities, but with the same policy goal. By doing so, meta-evaluation can provide CMS and states with a deeper understanding of what levers affect successful implementation and impacts, as well as whether, under what conditions, and how specific policy initiatives should be replicated in other states for maximum impact.

The plan for federal meta-analysis support includes the following major components.

1. We will use information abstracted from state demonstration special terms and conditions, implementation plans, evaluation designs, state quarterly and annual monitoring reports, and state evaluation reports to develop a detailed understanding of the activities and goals of each state demonstration. This information will be used to understand the features of each demonstration (see ***Section 2.2.1***), as well as contextual and implementation changes over time. This information will be coded systematically and used in subsequent analyses.

2. To validate and flesh out our understanding of state activities based on information abstracted from state documents, we will engage in discussions with state leaders and may also include payers and providers. These discussions will aim to assess our understanding of the features of each state's demonstration, contextual characteristics, and implementation experiences (***see Table 2-3***). In addition to understanding the demonstration itself, an accurate understanding of policies and activities that predated the demonstration will be critical. This information will also be incorporated in targeted case studies of selected demonstration design and implementation topics.

3. We will analyze multiple national data sets that address outcomes specified in SUD demonstration milestones and goals to understand the baseline situation in each state, as well as trends prior to the start of the demonstration as data permits. Baseline analyses provide important information for interpreting subsequent analyses of demonstration implementation and impacts. The data sources and their use is detailed in ***Table 2-5***. Metrics that can be calculated from each data source are shown in ***Tables 2-6*** and ***2-7***.

4. As data become available for the demonstration period, we will use data from state monitoring reports and state evaluations to analyze the relationship between outcomes specified by demonstration milestones and goals and demonstration features coded from report abstraction and discussions with state leaders (see ***Table 2-4***). Descriptive analyses of monitoring report data on demonstration implementation and impacts will provide early information on how outcomes compare across groups of states whose demonstrations share common features. When results from state evaluations are available, we will use forest plots and meta-regression of effect sizes reported by state evaluations to estimate the overall impact of demonstrations across states and identify factors that likely explain differences in state impacts.

5. We will supplement analyses of state monitoring and evaluation data with analyses of data from T-MSIS and other national data sets to assess the relationship between demonstration features and outcomes. The supplemental analyses will complement analyses of monitoring and evaluation data by providing metrics for outcomes that may not be reported for all demonstrations or are not measured consistently. T-MSIS may also support analyses that include comparison groups if these are not included in state evaluations. Depending on the data source and metric, we will use descriptive or multivariate regression analyses.

6. Qualitative case comparisons may be used to explore potential causal pathways between demonstration features and demonstration implementation effectiveness and outcomes. These will use the narrative parts of the state monitoring reports, the conversations with states, and a set of outcomes including implementation outcomes and demonstration impacts.

7. The data and analyses will support reporting of a summary of descriptive studies, cross-state synthesis, qualitative case comparisons, and meta-analysis results through a series of Rapid Cycle Reports and a Summative Evaluation Report, which will be shared on Medicaid.gov.

Table 1-1 summarizes our approach and methodology for the meta-analysis.

Table 1-1. Overview of Meta-Analysis Methodology

Data Abstraction and Synthesis	Data Analysis
• Extract, code, and synthesize qualitative data about state context, demonstration design, and implementation from state waiver applications and implementation plans. • Extract, code, and synthesize quantitative data about demonstration implementation and impact outcomes from state monitoring reports and state evaluations. • Conduct structured coding of data from discussions with state leaders and possibly payers and providers. • Conduct structured coding of data on state context, demonstration design, and implementation to define sets of demonstrations that share features of interest to feed into cross-state implementation and impact analyses.	• Conduct qualitative analyses to support cross-state comparisons of selected demonstration design and implementation topics. • Use descriptive analyses to compare implementation and impacts across demonstrations with common features. • Use forest plots and meta-regression of state evaluation findings to understand the relationship between demonstration features and demonstration impacts across sets of states. • Use T-MSIS and other national data sources to examine the relationship between demonstration features and demonstration outcomes. • May use qualitative comparative analysis and comparative case studies to assess relationships between demonstration features and demonstration impacts. Can also use these methods to assess relationship between demonstration features and monitoring metrics.

1.5 Organization of this Report

The next section of this report provides an overview of the outcomes that will be addressed and the data sources for the SUD demonstration meta-analysis. Section 3 describes the meta-analysis methodologies including meta-regression methods and qualitative comparative analysis. Section 4 describes the report deliverables and timelines.

SECTION 2.
META-EVALUATION DESIGN AND DATA SOURCES

2.1 SUD Demonstration Meta-Evaluation Conceptual Model

The design for the meta-evaluation of SUD demonstrations is guided by the conceptual model presented in ***Figure 2-1***. The conceptual model, which flows from left to right, is built around three key points of measurement. We need to first measure the pre-demonstration characteristics of Medicaid SUD treatment coverage and delivery to understand the potential for change in each state toward meeting the required SUD demonstration milestones. We then need to measure the changes made to meet each milestone and then the impacts of the changes made on the SUD demonstration goals. The goals and milestones are outlined in the State Medicaid Director Letter (SMDL) on SUD section 1115 demonstrations. After describing the conceptual model, we will discuss the data sources that will be used to capture these three dimensions. We will then describe measures that will be used to assess the extent to which state implementation outcomes meet demonstration milestones and measures that will be used to assess demonstration impacts. The implementation outcome measures are based on CMS's technical specifications for the SUD demonstration monitoring metrics and the impact measures are based on outcomes defined in the evaluation design technical assistance guide.

The SUD demonstrations must meet the following milestones,[7] which define implementation outcomes for the demonstrations, to support the demonstration goals:

(M1) access to critical levels of care for OUD and other SUDs (i.e., Medicaid coverage of these services);

(M2) widespread use of evidence-based, SUD-specific patient placement criteria;

(M3) use of nationally recognized, evidence-based SUD program standards to set residential treatment provider qualifications;

(M4) sufficient provider capacity at each level of care;

(M5) implementation of comprehensive treatment and prevention strategies to address opioid abuse and OUD; and

(M6) improved care coordination, care integration, and transitions between levels of care.

Achieving these milestones is expected to lead to successful performance on the impact goals specified for the demonstrations, as shown in the last two columns in ***Figure 2-1***. The figure describes a sequencing among the impacts, such that changes in appropriate SUD treatment (short-term impacts) must occur before there can be changes in overdoses, service utilization, and expenditures (long-term impacts).

CMS has identified six goals, or desired impacts, for the SUD demonstrations.[7] We list these here and describe how addressing the demonstration milestones may impact these goals:

(G1) **Increase rates of identification, initiation, and engagement in treatment**. Adding coverage of various SUD treatment services and increasing provider capacity and care coordination will increase the likelihood that Medicaid beneficiaries with an underlying SUD are identified and placed in SUD treatment.

(G2) **Increase adherence to and retention in treatment.** Increasing coverage, capacity, and care coordination; using placement criteria so patients receive the appropriate level of treatment; and ensuring that providers are appropriately certified to support a level of care will increase the likelihood that beneficiaries to stay in treatment.

(G3) **Reduce overdose deaths, particularly those due to opioids**. Increasing access to SUD treatment and emphasizing comprehensive prevention strategies will decrease the occurrence of fatal and non-fatal opioid- and non-opioid-related overdoses.

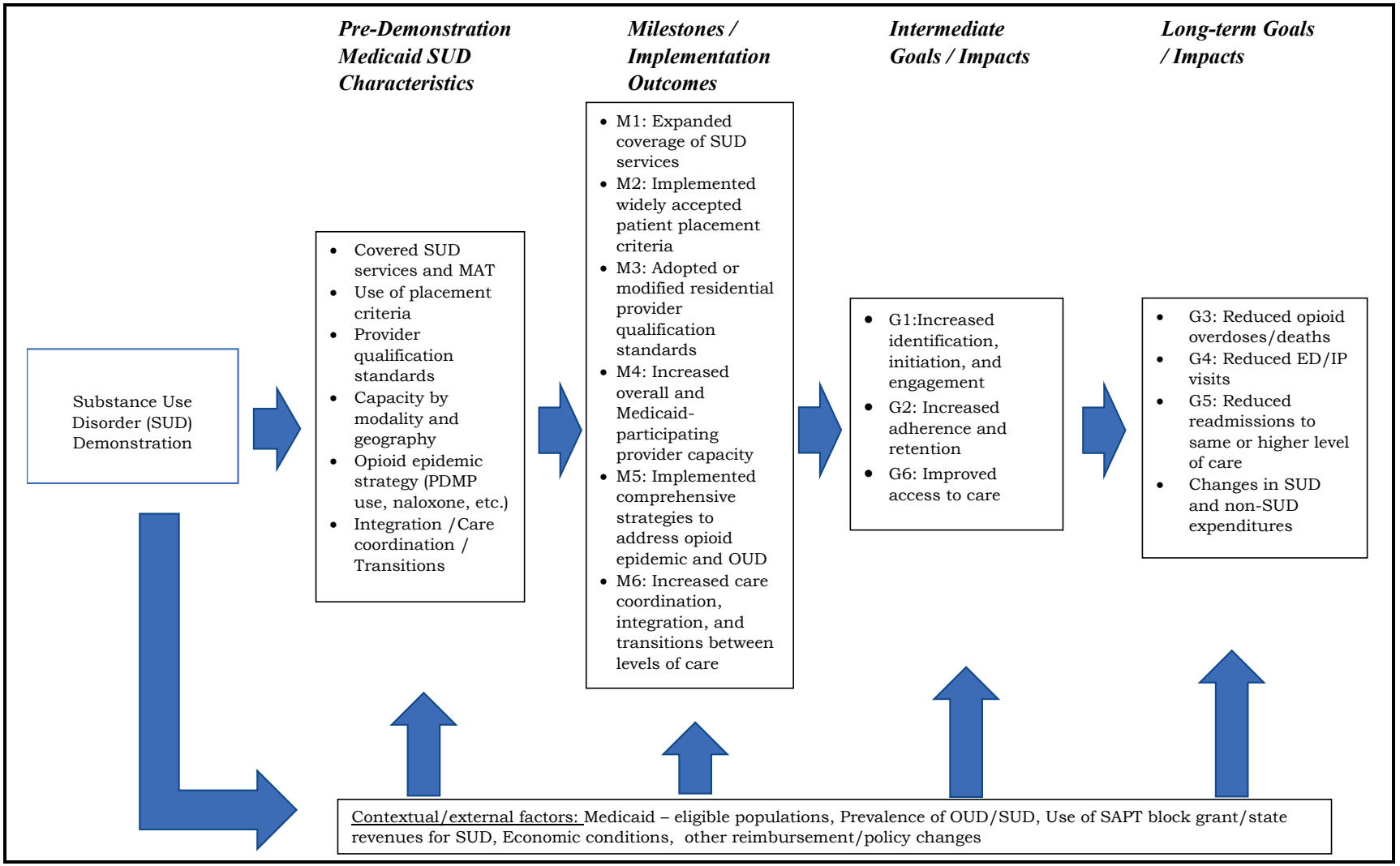
(G4) **Reduce preventable or medically inappropriate utilization of EDs and inpatient hospital settings for treatment by improving access to other continuum of care services.** Increasing the availability of SUD treatment and better coordinating treatment for SUD and comorbid conditions will increase the likelihood that individuals will receive appropriate treatment and reduce overdoses, injuries, and withdrawal symptoms as a result of their SUD, thereby avoiding inappropriate use of ED and inpatient hospital services.

(G5) **Reduce preventable or medically inappropriate readmissions to the same or higher level of care.** Increasing coverage of appropriate treatment, recovery support, and care coordination services will decrease the rate of relapse after discharge from a treatment episode, which will reduce readmissions to the same or higher level of care, especially residential treatment.

(G6) **Improve access to care for physical health conditions.** Many individuals with a SUD also have comorbid and chronic conditions that may go unmanaged. Increasing care coordination and access to primary care and other health services for treatment of these physical conditions will improve SUD and other health outcomes.

We will empirically test if specific milestones or combinations of milestones are associated with specific impacts.

Figure 2-1. Conceptual Model for Substance Use Disorder Demonstration Evaluation Design



ED = emergency department; IP = inpatient; MAT = medication-assisted treatment; OUD = opioid use disorder; PDMP = prescription drug monitoring program; SAPT = Substance Abuse Prevention and Treatment; SUD = substance use disorder.

In addition to impacts related to the six goals, CMS is also interested in understanding the impact of the demonstration on Medicaid costs. State evaluations are required to measure SUD demonstration impacts on Medicaid program costs for beneficiaries with a SUD, in total and disaggregated by whether the costs are related to SUD and within SUD, whether they are for IMD services, as well as by type of service (inpatient, ED, other outpatient, prescription drug, and long-term care). ***Table 2-1*** summarizes the expected impacts of the demonstration on Medicaid costs by spending category.

Table 2-1. Expected Impacts of SUD Demonstrations on Medicaid Costs

Cost Outcome	Expected Impact
Total Costs	
Federal + state	After a possible initial increase due to new coverage of services, these are expected to decrease if improved access to care for SUDs and comorbid conditions reduces complications and overall cost of treating these conditions
Federal only	Percent change will equal percent change in combined federal and state costs; absolute change will be proportional to state's federal medical assistance percentage (FMAP)
SUD Treatment Relationship	
IMD	Expected to increase
Other SUD treatment	Indeterminate; could decrease if IMD substitutes for other SUD treatment services and better access to treatment reduces SUD-related emergency department (ED) visits and inpatient admissions, and reduces the rate of relapses; could increase if access to the full range of SUD treatment services improves
Non-SUD	Indeterminate; could increase if identification of comorbid conditions and access to treatment of these conditions improves; could decrease if improved access to treatment reduces complications, ED visits, and hospital admissions
Type of Service	
Inpatient	Expected to decrease if better access to treatment for SUDs and comorbid conditions reduces hospital admissions
ED	Expected to decrease due to improved access to treatment for SUDs and comorbid conditions
Other outpatient	Indeterminate; could increase due to improved access to treatment for SUDs and comorbid conditions; could decrease if more timely treatment reduces complications
Pharmacy	Expected to increase as a result of increased use of MAT; could also increase if treatment of comorbid conditions requires prescriptions
IMD	Expected to increase as a result of IMD coverage
Long-term care	Not expected to change

Whether achieving the implementation outcomes specified by the milestones results in the demonstration impacts hypothesized by the goals depends on what the state had in place prior to the beginning of the demonstration. States that already covered a full continuum of SUD

benefits and had adequate SUD treatment capacity, for example, may be less likely to see increases in SUD treatment rates among Medicaid beneficiaries after the demonstration than are states that implemented new benefits and expanded capacity as part of the demonstration. Documenting differences across states in pre-demonstration SUD systems, Medicaid financing for SUD treatment, the adequacy of community linkages, and other contextual factors will be an important part of our analytic approach, to understand how treatment delivery is changed by the demonstration and, hence, the expected impacts.

The conceptual model also recognizes that the effects of the demonstration may be influenced by factors outside the demonstration. For example, whether the state expanded Medicaid eligibility to cover childless low-income adults determines the proportion of the population affected by the demonstration. One might expect larger changes in a state such as California with broad eligibility, compared to Virginia with more narrow eligibility. States are undertaking other initiatives, in parallel with the demonstration, that may influence outcomes and confound the attribution of observed effects to the demonstration. For example, states have used a variety of grants in different ways to address gaps in their continuum of SUD services, including the Substance Abuse and Mental Health Services Administration (SAMHSA) Substance Abuse Prevention and Treatment Block Grants, SAMHSA State Targeted Response and State Opioid Response grants, and Centers for Disease Control and Prevention (CDC) prescription drug monitoring program (PDMP) and prevention grants. States are responsible for ensuring that there is no duplication of federal payments, so additional grants can support the SUD continuum of services by extending what is provided through Medicaid. Other important moderating or confounding factors that need to be factored into the analysis include state innovation models and implementation of other payment reforms, the prevalence of OUD in the state and economic conditions. The additional challenges from COVID-19 spread and social isolation also need to be factored into the analysis, though they will theoretically affect both treatment and comparison groups for states with quasi-experimental evaluations. In summary, our conceptual model shows how the pre-demonstration structure of a state's SUD treatment system influences the amount of change needed to meet the implementation outcomes defined by demonstration milestones and how the amount of change made toward meeting the milestones then influences intermediate and long-term impacts.

2.2 Data Sources for the Meta-Evaluation

The meta-evaluation of the SUD demonstrations will examine whether differences in demonstration features impact an individual demonstration's outcomes and explain variation in outcomes observed across demonstrations. The meta-analysis will use a variety of data sources to

identify demonstration features and to measure implementation outcomes and impacts. The analysis will draw primarily on secondary data sources. Limited primary data will be collected as needed to supplement secondary data sources. We first discuss data sources for demonstration features and then data sources for implementation outcome and impact measures.

2.2.1 Data Sources for Demonstration Features

Document review and abstraction

The conceptual model describes demonstration design choices and state and Medicaid program contextual factors that may influence demonstration outcomes. As discussed earlier, implementation outcomes and SUD demonstration impacts are hypothesized to be affected by the state's pre-demonstration SUD systems, including their structures for delivering SUD treatment services, the gaps in their benefit packages and available services relative to the comprehensive treatment strategy requirement, and their adoption of other requirements such as patient placement criteria program integrity safeguards. The main data sources for identifying the demonstration features in the meta-analysis will be secondary data collected from demonstration documents, which we will use to identify demonstration features and how they change over the course of the demonstrations, as well as other relevant contextual characteristics that may affect demonstration outcomes. Some examples of these documents include:

- CMS SMDLs and other guidance to states and Medicaid beneficiaries
- State applications for section 1115 demonstrations
- CMS demonstration approval letters and special terms and conditions (STCs)
- Draft and approved state implementation plans
- Draft and approved state demonstration evaluation designs
- State quarterly and annual monitoring reports
- Interim and summative state evaluation reports

These data sources are available through CMS's Performance Metrics Database and Analytics (PMDA) and from Medicaid.gov and state Medicaid websites (see ***Table 2-2***).

We will abstract data from demonstration applications and STCs, quarterly monitoring reports, state evaluation reports, and other demonstration-related documents. We will standardize data abstraction using a structured and systematic coding scheme and we will create a state abstraction form (SAF) (***Appendix A***). During the qualitative coding process, coders will use the SAF to capture demonstration components and characteristics.

Table 2-2. Available State Demonstration Documents

State	Status	Application	Evaluation Design	Implementation Plan	Monitoring Protocol
Alaska	Approved	X	X	X	X
Arizona	Pending	X	-	-	-
California	Approved	X	-	X	-
Colorado	Pending	X	-	-	-
Delaware	Approved	X	Due 2/1/2020	X	Due 3/1/2020
District of Columbia	Approved	X	Due 5/4/2020	X	Due 5/17/2020
Idaho	Approved	X	-	-	-
Illinois	Approved	X	X	X	-
Indiana	Approved	X	X	X	X
Kansas	Approved	X	X	X	X
Kentucky	Approved	X	X	X	
Louisiana	Approved	X	X	X	X
Maine	Pending	-	-	-	-
Maryland	Approved	X	X	-	-
Massachusetts	Approved	X	X	-	-
Michigan	Approved	X	X	X	X
Minnesota	Approved	X	Due 3/1/2020	X	X
Nebraska	Approved	X	Due 3/31/2020	X	X
New Hampshire	Approved	X	X	X	X
New Jersey	Approved	X	X	X	X
New Mexico	Approved	X	X	X	Past Due (11/8/2019)
North Carolina	Approved	X	With CMS	X	X
Ohio	Approved	X	Due 3/21/2020	-	Due 2/21/2020
Pennsylvania	Approved	X	X	X	X
Rhode Island	Approved	X	X	X	X
Tennessee	Pending	X	-	-	-
Utah	Approved	X	X	X	X
Vermont	Approved	X	X	X	X
Virginia	Approved	X	-	-	X
Washington	Approved	X	X	X	X
West Virginia	Approved	X	X	X	X
Wisconsin	Approved	X	X	X	Past Due (12/1/2019)

X = RTI has received state document; - = State document not received and due date not specified in the PMDA

Stakeholder discussions and focus groups

Secondary data from demonstration documents will be supplemented with primary data collected through discussions with key stakeholders in demonstration states. Discussions with stakeholders will be used to confirm information on Medicaid policy changes and other SUD policy changes occurring as a part of the demonstration and to provide a contextual narrative on the impact of the SUD demonstrations. After abstracting demonstration features from the state waiver applications, STCs, implementation plans, evaluation reports, and monitoring reports, we will use stakeholder discussions to update and clarify information not covered in the regular reporting or not consistently reported across states. For example, changes in reimbursement for SUD services are not always reported as a part of the implementation and are critical in understanding the impact of the demonstrations. An example grid for state responses is included in ***Appendix B***.

Stakeholder discussions are also necessary to confirm drivers of change with the demonstrations and to accurately document variation among states in demonstration design elements. Information from discussions will be used as inputs to the demonstration features used in the meta-analysis and will facilitate a rich interpretation of results from quantitative models. We will also use the discussions to identify challenges, successes, and best practices in implementation as state demonstrations evolve. In addition to SUD section 1115 demonstration states, we will conduct discussions with states implementing a SPA through section 5052 under the SUPPORT Act to understand their choice of the SPA approach as it relates to challenges, successes, and the expansion of the SUD continuum of services.

To allow the full range of required conversations, we will pursue approval of our interview strategy from the Office of Management and Budget (OMB). We will conduct virtual stakeholder discussions with officials from the state Medicaid agency and the single state agency for substance abuse services for each state with an approved SUD demonstration or SPA during Years 2–5 of the meta-analysis contract. These two types of stakeholders are in the best position to describe a state's demonstration planning and implementation process, describe its strengths and weaknesses, describe challenges and alterations, confirm program features before and after demonstration or SPA implementation, and discuss potential impacts of other initiatives in the state on the demonstration. ***Table 2-3*** shows the intended discussion topics for these stakeholders and ***Appendix C*** provides a draft discussion guide for the baseline conversations. We will develop additional discussion guides for subsequent conversations. Initial discussions in Years 2 and 3 will aid in establishing baseline data, while discussions in Years 4 and 5 will collect data on the states' continued progress in implementation and in working toward demonstration

milestones and goals. Topics of inquiry are likely to shift over the course of the meta-analysis as demonstrations proceed and will be decided in collaboration with CMS. Discussion guides will be tailored to each state and to each stakeholder group within the state. States that are furthest along will be prioritized for scheduling discussions and we will work with CMS to determine the appropriate stakeholders in each state.

In addition to discussions with state-level stakeholders, we will also conduct virtual discussions and in-person site visits with managed care organizations and health systems in Years 3 and 4 of the meta-analysis contract. In states where we conduct in-person site visits, we will expand the stakeholder groups with whom we have discussions by conducting focus groups with providers. Our experience has shown that focus groups are an effective and efficient method for getting provider feedback. These additional discussions and focus groups will capture perspectives on the impacts of SUD demonstrations and SPAs among those paying for and providing SUD treatment services. ***Table 2-3*** shows the intended discussion topics for these stakeholders.

Table 2-3. SUD Demonstration Stakeholders and Discussion Topics

Stakeholder	Discussion Topic
State Medicaid agency officials	• Service coverage before and after the demonstration • Changes in reimbursement or the development of new reimbursement codes • Clarification of information in the implementation plans on major activities and initiatives • Rationale for pursing SUD section 1115 demonstration and • Challenges to adding SUD treatment and support services, including Medication Assisted Treatment, Residential treatment services, and other changes • Approach to implementing patient placement criteria, care coordination and, transitions between levels of care. • Perceptions of major changes resulting from the demonstration and effects on outcomes • Potential impact of other initiatives in the state on the demonstration

(*Continued*)

Table 2-3. SUD Demonstration Stakeholders and Discussion Topics (Continued)

Stakeholder	Discussion Topic
• Single state agency for substance abuse services officials	• Service coverage before and after the demonstration • Key structural characteristics and limitations of state SUD system (Medicaid and non-Medicaid specific) • Rationale for pursing SUD section 1115 demonstration and • Challenges to adding SUD treatment and support services, including Medication Assisted Treatment, Residential treatment services, and other changes • Approach to implementing patient placement criteria, care coordination and, transitions between levels of care • Other SUD policy changes occurring concurrently with the demonstration • Collaborations between states agencies and other stakeholders • Perceptions of major changes resulting from the demonstration and effects on outcomes • Potential impact of other initiatives in the state on the demonstration
• Local evaluators	• Evaluation design and comparison group selection • Quantitative and qualitative approaches • Data sources and data quality • Evaluation challenges • Preliminary findings
• Managed care organizations	• Impact of changes in reimbursement for IMD residential and for non-residential services • Impact of changes in required service offerings, patient placement criteria, and utilization management • Changes in provider networks • Changes in integration and care coordination
Health systems/providers/ social service agencies	• Impact of changes in reimbursement on Medicaid participation • Changes in provider capacity at each level of care resulting from the demonstration • Changes in care coordination and transitions between levels of care for patients • Perceptions of major changes resulting from the demonstration and effects on outcomes • Potential impact of other initiatives in the state on the demonstration
Providers (focus groups)	• Ability of beneficiaries to access appropriate level of care for OUD and SUDs • Ability of beneficiaries to access care for physical conditions • Assistance in care coordination and transitions between levels of care provided to beneficiaries

Preliminary framework for demonstration features

After an initial abstraction for SUD demonstration states, we have identified a preliminary set of features that will be used in the meta-analysis. These features will be modified and refined if needed as more information on SUD demonstrations and 5052 SPAs become available. The preliminary set of features, summarized in ***Table 2-4***, are as follows:

- **Coverage of the Comprehensive Continuum of SUD services.** We hypothesize that states that expand the services they cover will have larger impacts on access to SUD services. As a part of the SUD demonstrations, states are required to offer a comprehensive set of SUD treatment services within 12 to 24 months of the demonstration approval. All states cover basic outpatient services and buprenorphine, but there is variability in whether state Medicaid programs cover methadone for OUD, intensive outpatient services, residential SUD services in non-IMDs or residential services in IMDs without the use of FFP, and a range of recovery support services. States that already provide a wide array of SUD services at baseline will be limited with respect to how many additional services they can provide and, therefore, their demonstrations may have smaller impacts relative to states that begin with a less comprehensive set of services and expand SUD services through the demonstration.

- **Use of patient placement criteria, utilization management, and residential provider standards.** We hypothesize that states that must newly implement or update widely recognized patient placement criteria will have larger impacts on demonstration outcomes for treatment engagement. Standardized and consistent use of comprehensive assessments and utilization review should increase beneficiaries' appropriate placement in treatment. Requiring residential providers to offer on- or off-site access to MAT should also increase the number of Medicaid beneficiaries receiving MAT.

- **Provider capacity.** We hypothesize that states that increase the number of MAT providers and SUD providers accepting Medicaid will have larger changes in the demonstration outcomes. Expanding Medicaid coverage of SUD services or increasing reimbursement of SUD services may provide incentives for existing SUD providers to accept Medicaid clients that these providers could not previously serve. The demonstrations may also increase MAT provider capacity through increased reimbursement and outreach and recruitment efforts. States may also make licensing changes that allow for same day billing for SUD services in primary care settings, allow additional types of providers to prescribe MAT, or reduce the time or financial burden for certification. Increased availability of and access to treatment should result in larger demonstration impacts.

- **Care coordination.** We hypothesize that demonstration impacts will be larger in states that increase care coordination. As part of the demonstrations, states need to enact policies that connect beneficiaries with community-based treatment and recovery support services, including physical care and SUD services, social supports, and case management. While these policy changes are intended specifically for

beneficiaries discharged from residential and inpatient SUD treatment, states may extend these services to beneficiaries in other levels of care. As with coverage of SUD services, there is variation across states in care coordination policies and we anticipate that states that had robust care coordination policies prior to their demonstration may have less room for improvement relative to states that did not have as robust care coordination policies in place.

- **Health IT requirements**. States with approved section 1115 SUD demonstrations must submit a SUD health IT plan as a component of their implementation plans, related to Milestone #5 (Implementation of comprehensive treatment and prevention strategies to address opioid abuse and OUD). The SUD health IT plan specifically describes strategies to increase utilization and improve functionality of PDMPs. At this time, we do not have enough information to identify distinct health IT or PDMP features. We will continue to monitor state SUD health IT plans and monitoring protocols as they are made available to identify common features and core metrics that can be incorporated into the meta-evaluation. ***Appendix D*** provides a brief overview of what we have learned to date.

- **State evaluation methodology**. Aspects of a state's evaluation design may have as much impact on the results as the intervention itself. We will therefore monitor selected characteristics of the state evaluation research designs, such as whether the state used difference-in-differences or interrupted time series analysis and whether the state used any comparison group, a within-state comparison group, or an out-of-state comparison group.

- **Medicaid expansion status**. Medicaid expansion status is not a design feature of the demonstration; however, whether the state expanded Medicaid eligibility to cover childless low-income adults determines the proportion of the population affected by the demonstration. There might be larger impacts in states with broader eligibility compared to those with narrower eligibility.

2.2.2 Data Sources for Demonstration Implementation Outcomes and Impacts

State evaluation reports and monitoring metrics

The main data sources for demonstration impacts will be state evaluation reports and the main sources for the implementation outcomes will be state-reported monitoring metrics. Measures derived from these data sources will be used as outcome variables in meta-regressions and other analyses described in Section 3. We will abstract estimates of the impact of the demonstrations from the state evaluation reports, as well as trends in monitoring metrics from state quarterly and annual monitoring reports in the PMDA. We will also interview state evaluators if needed to confirm key details, strengths, and limitations of their evaluations and to identify state-specific evaluation challenges and considerations that may influence the demonstration impacts identified in the state evaluations.

Table 2-4. Preliminary Demonstration Features for a Sample of 25 States

	Alaska	DC	Illinois	Indiana	Kansas	Kentucky	Louisiana	Maryland	Massachusetts	Michigan	Minnesota	Nebraska	New Hampshire	New Jersey	New Mexico	North Carolina	Ohio	Pennsylvania	Rhode Island	Utah	Vermont	Virginia	Washington	West Virginia	Wisconsin
Milestone #1—Access to critical levels of care																									
Non-Residential SUD Benefit Changes																									
Added methadone	No	No	No	No	Yes*	Yes*	Yes*	No	No	No	No	Yes*	No	No	No	No	No	No	No	No	No	No	No	Yes*	No
Added intensive outpatient	Yes*	No	No	Yes†	No	No	No	No	No	No	Yes*	No	No	No	Yes†	No	No	No	No	No	No	No	No	No	No
Added Partial Hospitalization	Yes*	§	§	No	No	Yes†	§	No	No	No	§	No	No	No	Yes*	No	No	No	No	No	No	No	§	No	§
Added WM as a new service	Yes*	Yes†	Yes†	§	Yes†	Yes*	No	Yes†	No	-	Yes†	Yes†	No	No	Yes†	Yes†	No	No	§	Yes†	No	Yes*	No	Yes†	§
Added recovery supports	Yes*	Yes*	Yes†	Yes*	Yes†	Yes†	§	§	Yes*	Yes†	Yes†	No	No	Yes*	Yes†	Yes*	§	Yes†	§	Yes†	Yes†	Yes*	§	Yes*	§
Residential SUD Benefit Changes																									
Added coverage for IMDs	Yes*	Yes†	Yes*	Yes*	Yes*	Yes*	Yes†	Yes*	Yes†	Yes*	Yes*	Yes†	Yes†	Yes*	Yes*	Yes*	Yes*	Yes*	Yes*	Yes*	No	Yes*	Yes*	Yes*	Yes*
Added residential as a new benefit	Yes*	Yes†	No	Yes*	No	§	No	Yes*	Yes*	No	No	No	No	Yes†	Yes*	Yes†	Yes†	Yes†	No	No	No	Yes*	No	Yes*	Yes*
Had previously covered non-IMDs	No	Yes	Yes	No	Yes	Yes+	Yes	No	Yes	Yes	Yes	No	Yes	Yes+	Yes**	Yes	No	Yes**	Yes	Yes	Yes	Yes**	Yes	Yes	No
Had used "in lieu of" provision" for IMDs	No	Yes	No	Yes**	No	No	Yes	No	No	No	No	Yes+	No	No	No	No	Yes	Yes**	No	No	No	No	§	No	No
Milestone #2—Newly implemented or updated placement criteria	Yes*	No	Yes†	Yes†	Yes†	Yes†	No	Yes*	§	§	Yes*	Yes†	No	Yes†	Yes*	Yes†	Yes†	Yes*	§	Yes†	No	Yes†	Yes†	§	Yes†
Milestone #3—Program Standards for Residential Treatment Providers																									
Added or updated residential provider standards	Yes*	No	No	Yes*	Yes†	Yes*	Yes**	Yes	§	Yes*	Yes†	Yes†	Yes†	Yes†	Yes*	Yes†	Yes†	Yes*	No	Yes†	Yes†	Yes†	No	§	Yes*
Added residential MAT requirements	Yes*	Yes†	Yes*	Yes*	Yes*	Yes*	Yes*	§	§	§	Yes*	Yes*	Yes†	Yes*	§	Yes†	Yes*	No	No	§	Yes†	§	Yes*	§	Yes*
Milestone #6—Improved care coordination	Yes*	Yes†	§	§	Yes†	Yes†	§	§	§	§	§	§	Yes†	Yes†	Yes†	Yes*	Yes†	No	Yes†	§	Yes†	§	§	§	§
Other demonstration features																									
Improved co-morbid care	Yes*	Yes†	§	§	§	§	§	Yes*	§	§	§	§	§	§	Yes*	Yes*	§	§	§	§	§	Yes*	§	§	§
Reimbursement changes	Yes‡	Yes‡	No	Yes‡	No	No	No	§	No	No	Yes‡	No	No	Yes‡	Yes‡	Yes‡	No	No	No	Yes‡	Yes‡	Yes‡	Yes‡	§	Yes‡

* Green = Likely larger change

† Yellow = likely smaller change

‡ Orange = shift in the way services are financed

§ Purple = Missing or incomplete information

National data sources

We will derive metrics related to demonstration implementation outcomes and impacts from national data sources (summarized in ***Table 2-5***) as needed to supplement or augment measures available from state evaluations and monitoring metrics. National data sources may be used to supplement state evaluation findings by (1) deriving key measures that are not reported for all demonstrations; (2) incorporating comparison groups in analyses if these are not included in state evaluations; and (3) testing for differential demonstration effects by beneficiary subgroup. Differences in state evaluation methods and reporting of monitoring metrics may influence results; national data sources offer the potential for standardizing measure definitions and analyses across states. National data sources can also be used to characterize baseline conditions in demonstration and non-demonstration states to interpret demonstration impacts in the context of national trends and assess the generalizability of findings to non-demonstration states. In addition, these national data sets can be a source for relevant state-level characteristics that can be used as control variables in meta-analyses. Finally, national data sets offer information that can inform advice to states on comparison group selection if needed. Each data set also may have limitations, including the precision with which the intervention and comparison populations can be identified, or completeness of data fields that are of interest. For this reason, we may use multiple data sources to validate the primary data source on any outcome. We provide more details on each data set below.

Table 2-5. National Datasets to be Used in Supplemental Analyses

Dataset Name	Dataset Description	Expected Availability for Summative Evaluation Report
Transformed Medicaid Statistical Information System (T-MSIS)	Standardized Medicaid beneficiary enrollment and claims data reported by state Medicaid agencies to CMS and available in the Chronic Conditions Warehouse (CCW). Also include Medicaid provider files. Prior to T-MSIS, standardized Medicaid enrollment and claims files are available in the CCW as Medicaid Analytic Extract (MAX) and Alpha-MAX files. Date of conversion to T-MSIS varies by state but all states have reported T-MSIS data since 2016.	2016–2022, currently available through 2017 for some states
Treatment Episode Data Set (TEDS)	Administrative data set of specialty SUD treatment admissions and discharges. Administered by SAMHSA.	2010–2021, currently available through 2017

(continued)

Table 2-5. National Datasets to be Used in Supplemental Analyses (continued)

Dataset Name	Dataset Description	Expected Availability for Summative Evaluation Report
National Survey of Substance Abuse Treatment Services (N-SSATS)	Annual survey of all public and private SUD treatment facilities that collects information on facility characteristics, capacity, and licensure and accreditation. Maintained by SAMHSA.	2010–2021, currently available through 2018
Drug Enforcement Agency (DEA) Controlled Substances Act (CSA) Registration Records	Tracking database of providers with a buprenorphine waiver. Maintained by the DEA.	2014–2023, updated daily
Centers for Disease Control and Prevention Wide-ranging Online Data for Epidemiologic Research (CDC WONDER)	Contains two mortality databases that capture information on opioid-related overdose deaths: The Underlying Cause of Death database and the Multiple Cause of Death database. The Underlying Cause of Death database provides mortality and population counts for all U.S. states and counties. Maintained by the CDC.	2010–2021, currently available through 2017
Healthcare Cost Utilization Project (HCUP)	Inpatient and emergency department discharge records. Maintained by AHRQ.	2010–2021, currently available through 2017 (2018 for some states)
National Survey on Drug Use and Health (NSDUH)	Annual survey that captures prevalence rates of substance use and mental health-related issues. Maintained by SAMHSA.	2010–2021, currently available through 2018
Medicaid State Drug Utilization Data	Database of all prescriptions dispensed to Medicaid beneficiaries that are paid for by Medicaid. Updated quarterly. Maintained by CMS.	2010–2023, currently available through mid-2019

Transformed Medicaid Statistical Information System (T-MSIS)

Many SUD demonstration outcomes are claims-based measures that can be created from T-MSIS and predecessor Medicaid claims and enrollment data (MAX and Alpha-MAX). These include initiation and retention in SUD treatment; ED and inpatient utilization; hospital readmissions; use of primary care services; and expenditures for services, overall and SUD-specific. T-MSIS provider files can also be used to measure availability of SUD providers. Analysis of T-MSIS data can augment findings available from state monitoring reports and state evaluations if they do not include key measures, outcomes are not measured consistently across states, and impact estimates do not incorporate comparison groups.

T-MSIS Analytic Files (TAFs) have been available to users for less than a year. A number of limitations in the data files have been identified and these limitations often differ by state. Data limitations found in our assessment of 13 states' TAF files as of December 2018

included lack of geographic identifiers (zip code and county code); missing admission or discharge date for inpatient admissions; no payment information for managed care encounters; and high percentage of services in the Other Therapy (OT) file missing place of service or procedure codes. Using the 2018 T-MSIS data, a recent report by the Department of Health and Human Services Office of Inspector General assessed whether opioid prescribing in Medicaid could be monitored nationally. The report found that Medicaid beneficiaries could have multiple IDs within a state and that National Provider Identifier (NPI), diagnosis code, or quantity in the pharmacy data was missing in 32 states.[a] The T-MSIS files are being updated to correct data issues, however. The 2014 to 2016 TAFs were refreshed from July to September 2019 and are currently in "Version 3" of the files. The files for later years are also expected to be updated. Some previously identified data limitations may be addressed in the updated versions.

RTI will conduct analyses to determine whether the data quality and completeness in the updated files is sufficient to support analysis of SUD demonstration outcomes. We will conduct standard validation of every file to assess (1) rates of missingness; (2) validity of data fields (e.g., ensure that diagnosis code variables contain valid diagnosis codes in the expected format); (3) linkage rates across files (e.g., ensure that the baseline summary file links successfully to the inpatient file). For the baseline summary file, we will determine whether beneficiaries have multiple Medicaid IDs within a state, and we will calculate enrollment rates by eligibility category and dual status and compare the rates to other sources to validate the values. We will also assess whether we can accurately measure the key outcomes across states by calculating frequencies of key data elements such as SUD diagnoses and procedure codes in the data. We will look at managed care encounters specifically to assess whether managed care payment amounts are populated and valid. We anticipate that data quality will continue to improve, and we will assess the data as new updates become available over time.

Treatment Episode Data Set (TEDS)

TEDS comprises administrative records of 1.5 million to 2 million admissions to publicly funded substance abuse treatment across all 50 states and the District of Columbia. It contains patient demographic and socioeconomic characteristics, referral source, treatment modality, payment source, substances used and frequency, treatment history, and treatment plans (i.e., use of MAT). TEDS is an option for assessing pre- and post-demonstration admissions trends to address SUD Milestone #1 (Access to critical levels of care for OUD and other SUDs). Counts

[a] HHS-OIG Data SnapShot. "National Review of Opioid Prescribing in Medicaid Is Not Yet Possible." August 2019. Available at: https://oig.hhs.gov/oei/reports/oei-05-18-00480.pdf

of admissions of certain types (e.g., opioid-related admissions) are particularly important for understanding the experience of demonstration states prior to implementation as well as for identifying suitable comparison states. TEDS could also be useful for Milestone #3 (Use of nationally recognized evidence-based SUD program standards to set residential treatment provider qualifications) in assessing whether individuals treated in residential facilities are more likely to receive MAT. Preliminary analyses and limitations of TEDS data are discussed in ***Appendix E.***

National Survey of Substance Abuse Treatment Services (N-SSATS)

N-SSATS is an annual survey of facilities that have been identified by SAMHSA as providing SUD treatment. N-SSATS captures information on participating facilities including data on payment sources accepted, levels of care and services provided, clinical approaches used, and counts of patients receiving treatment on a single reference day (typically a day at the end of March). N-SSATS can be used to measure changes in pre- and post-demonstration capacity for SUD Milestone #4 (Sufficient provider capacity at each level of care). Preliminary analyses and limitations of N-SSATS data are discussed in ***Appendix E.***

Drug Enforcement Agency (DEA) Controlled Substances Act Registration Records

Once SAMHSA approves a buprenorphine waiver for a provider, this information is forwarded to the DEA. The DEA assigns a registration number that encodes the type of provider (e.g., physician, nurse practitioner) and the patient limit (25, 100, or 275). Providers must renew their waiver annually, and the DEA tracks both active providers and providers who have let their waiver expire in the Controlled Substances Act (CSA) Registration Records. The Active File consists of records of all individuals registered under the CSA, including registrants doing business under their individual name rather than a business name. The DEA also produces a list of registrants whose certification numbers have been retired or suspended from the Active File. This data file can be used to track changes in buprenorphine-waived-provider capacity.

The DEA CSA can also be used to assess pre- and post-demonstration trends in buprenorphine-waivered-provider capacity to address SUD Milestone #4 (Sufficient provider capacity at each level of care). State evaluations may use provider information in Medicaid claims data to track changes in provider capacity, but this will only identify providers who actually prescribed buprenorphine. Although we anticipate that Medicaid provider files will have comprehensive information, in the event that they do not, this data source can help to supplement. Additionally, the DEA CSA registration records data could be merged with claims

data using National Provider Identifiers to assess changes in capacity by provider specialty (e.g., nurse practitioners) and prescribing patterns among waived providers with different patient caps.

Centers for Disease Control and Prevention Wide-ranging Online Data for Epidemiologic Research (CDC WONDER)

CDC WONDER contains two mortality databases that capture information on opioid-related overdose deaths at the state and county level based on death certificates for U.S. residents: The Underlying Cause of Death database and the Multiple Cause of Death database. The type of opioid can be disaggregated into four categories: heroin, methadone, natural and semisynthetic opioids, and synthetic opioids other than methadone. CDC WONDER can be used to assess pre- and post-demonstration opioid-related overdose deaths for SUD Goal #3 (Reduce overdose deaths, particularly those due to opioids). Preliminary analyses and limitations of CDC WONDER data are discussed in ***Appendix E.***

Healthcare Cost Utilization Project (HCUP)

HCUP collects administrative discharge records from community hospitals, EDs, and ambulatory surgical centers. State-specific databases are available for approximately 30 states between 2010 and 2017, depending on the year, and 2018 data are starting to become available as of September 2019. Thirteen pending or approved SUD demonstration states have ED data and 17 have inpatient data. HCUP data can be used to address SUD Goal #4 (Reduce utilization of EDs and inpatient hospital settings for treatment where the utilization is preventable or medically inappropriate through improved access to other continuum of care services). HCUP data can be used to track opioid-related and other SUD-related admissions and ED visits, identified based on diagnosis codes/diagnosis-related group (DRG), where Medicaid is the expected primary payer. Most state-level databases available for purchase include location information enabling analyses at the full state, county, or zip code levels. These variables can be used to better identify demonstration-level beneficiaries in states implementing their demonstrations in targeted geographic areas. HCUP also has an indicator for events where the patient died, which can be used to address SUD Goal # 3 (Reduce overdose deaths, particularly those due to opioids). HCUP offers an alternative to Medicaid claims data if availability of T-MSIS is delayed or data quality problems are not resolved.

National Survey on Drug Use and Health (NSDUH)

NSDUH annually collects detailed information on substance use and mental health-related issues and can be used to track changes in the prevalence of substance use across time at the state level. We can use NSDUH to track the prevalence of opioid and other substance use for

the general population and among individuals with Medicaid across demonstration and non-demonstration states. Prevalence rates among the Medicaid population can be used as an alternative measure for identification relative to claims-based definitions under Goal #1 (Increased Rates of Identification, Initiation and Engagement in Treatment) and may also serve as an important control variable in other analyses. One limitation of NSDUH is that the sample size for individual states are small and state-level estimates are pooled across at least 2 years. Thus, it may be difficult to detect change in smaller states.

Medicaid State Drug Utilization Data

Medicaid State Drug Utilization Data track the numbers of prescriptions and dispensed units (e.g., tablets, milligrams, vials) for prescription drugs dispensed to Medicaid beneficiaries in each state in each calendar quarter, and the total payments made to dispensing providers (e.g., pharmacies) for those medicines. Data are reported by National Drug Code (NDC). There is no information about the beneficiaries who receive these medicines, or where or by whom or for what reason they were dispensed. Medicaid State Drug Utilization Data are updated quarterly and generally lag only 3 to 6 months from current dates. They are likely to be available more quickly than concurrent claims data from T-MSIS, MAX, or other sources. They are best used to observe trends in use and spending in Medicaid, at the state level, for medications used in treatment of SUD/OUD.

2.3 Meta-Evaluation Questions

The meta-evaluation will address questions related to both implementation outcomes of SUD demonstrations and demonstration impacts and section 5052 SPAs as applicable. We begin by discussing the questions related to implementation outcomes associated with demonstration milestones, followed by questions related to impacts associated with demonstration goals. In both sections, we describe outcome measures and potential data sources to address each question. We anticipate that the analyses will focus on a targeted set of outcome measures agreed on with CMS.

2.3.1 Demonstration Implementation Outcomes

One purpose of the meta-evaluation is to understand the extent to which implementation outcomes defined by demonstration milestones are achieved and whether state context or demonstration design features are associated with variation in implementation outcomes. ***Table 2-6*** outlines the questions that will be addressed in the meta-evaluation to assess changes in implementation outcomes associated with the milestones. Because state evaluations are not required to assess milestones, we rely primarily on the metrics in state monitoring reports to

assess these outcomes. We also describe supplemental data sources that may be used as alternatives if measures in state monitoring reports are incomplete or not measured consistently. In a few cases, supplemental data can be used to create measures that are not included in the monitoring metrics. In some cases, monitoring metrics parallel measures used for demonstration outcomes. We include these to address both implementation outcome and impact questions because monitoring metrics will provide an earlier look at outcomes than state evaluation findings, which will not be ready until the final years of the meta-analysis contract.

Table 2-6. Implementation Outcome Questions, Data Sources, and Measures for the SUD Demonstration Meta-Evaluation

Question	Source(s)	Measures
SUD Milestone #1: Access to critical levels of care for OUD and other SUDs		
Did demonstrations increase utilization of critical levels of care for SUD during their SUD demonstration period?	Quarterly monitoring reports Supplemental: T-MSIS / MAX claims	• Number of Medicaid beneficiaries enrolled receiving any SUD treatment service, facility claim, or pharmacy claim • Number of Medicaid beneficiaries who used early intervention services (e.g., SBIRT) • Number of Medicaid beneficiaries who use outpatient services for SUD • Number of Medicaid beneficiaries who use intensive outpatient and/or partial hospitalization services for SUD • Number of Medicaid beneficiaries who use residential and/or inpatient services for SUD • Number of Medicaid beneficiaries who use withdrawal management services • Number of beneficiaries who have a claim for medication-assisted treatment (MAT) for SUD
	Annual monitoring reports Supplemental: T-MSIS / MAX claims	• Average length of stay for Medicaid beneficiaries discharged from Institutions for Mental Diseases residential treatment for SUDs
	TEDS	• Number of admissions to SUD specialty treatment • Number of admissions with MAT planned
SUD Milestone #2: Widespread use of evidence-based, SUD-specific patient placement criteria		
Did demonstrations increase use of evidence-based, SUD-specific patient placement criteria?	N/A	• No metrics are related to Milestone #2
SUD Milestone #3: Use of nationally recognized evidence-based SUD program standards to set residential treatment provider qualifications		
Did demonstrations increase residential treatment facilities that offer MAT on-site or facilitate access off-site?	T-MSIS / MAX claims	• Number of beneficiaries in residential treatment with at least one claim for MAT • Number of residential treatment facilities with at least one claim for MAT
	TEDS	• Number of residential admissions with MAT planned

(continued)

Table 2-6. Implementation Outcome Questions, Data Sources, and Measures for the SUD Demonstration Meta-Analysis (continued)

Question	Source(s)	Measures
SUD Milestone #4: Sufficient provider capacity at critical levels of care		
Did demonstrations increase the overall number of Medicaid SUD providers at the state and sub-state level?	Annual monitoring reports	• Number of providers who were enrolled in Medicaid and qualified to deliver SUD services during the measurement period
	Supplemental: T-MSIS / MAX claims	• Number of providers with a SUD specialty taxonomy code in the Medicaid provider data
	N-SSATS	• Number of SUD facilities per 100,000 population • Number of SUD facilities that accept Medicaid per 100,000 population
Did demonstrations increase the number of SUD providers able to provide MAT for OUD at the state and sub-state level?	Annual monitoring reports	• Number of providers who were enrolled in Medicaid and qualified to deliver SUD services during the measurement period and who met the standards to provide buprenorphine or methadone as part of MAT
	Supplemental: T-MSIS / MAX claims	• Number of providers prescribing at least one MAT
	DEA CSA registration records	• Number of physicians with waivers per 100,000 population
	T-MSIS / MAX claims linked with DEA CSA registration records	• Number of Medicaid providers with waivers that allow them to provide buprenorphine for the treatment of OUD, including non-physicians (e.g., nurse practitioners) • Number of Medicaid buprenorphine prescriptions written by waived providers
Did demonstrations increase the number of residential SUD facilities?	N-SSATS	• Number of SUD residential facilities per 100,000 population • Number of SUD residential facilities that accept Medicaid per 100,000 population
Did demonstrations increase the number of outpatient SUD facilities?	N-SSATS	• Number of SUD outpatient facilities per 100,000 population • Number of SUD outpatient facilities that accept Medicaid per 100,000 population • Number of SUD outpatient facilities offering intensive outpatient treatment per 100,000 population • Number of SUD outpatient facilities offering intensive outpatient treatment that accept Medicaid per 100,000 population

(continued)

Table 2-6. Implementation Outcome Questions, Data Sources, and Measures for the SUD Demonstration Meta-Analysis (continued)

Question	Source(s)	Measures
		• Number of SUD outpatient facilities offering partial hospitalization per 100,000 population • Number of SUD outpatient facilities offering partial hospitalization that accept Medicaid per 100,000 population
Did demonstrations increase the number of facilities offering MAT?	N-SSATS	• Number of facilities offering MAT per 100,000 population • Number of facilities offering MAT that accept Medicaid per 100,000 population
SUD Milestone #5: Implementation of comprehensive treatment and prevention strategies to address opioid abuse and OUD		
Did demonstrations increase initiation and engagement of alcohol and other drug dependence (AOD) treatment?	Annual monitoring report Supplemental: T-MSIS / MAX claims	• Percentage of beneficiaries who initiated treatment through an inpatient AOD admission, outpatient visit, intensive outpatient encounter or partial hospitalization, telehealth, or MAT within 14 days of the diagnosis • Percentage of beneficiaries who initiated treatment and who had two or more additional AOD services or MAT within 34 days of the initial visit
Did demonstrations reduce use of opioids at high dosage and/or from multiple providers in persons without cancer?	Annual monitoring report Supplemental: T-MSIS / MAX claims	• Rate per 1,000 beneficiaries age 18 and older without cancer who received prescriptions for opioids with a daily dosage greater than 120 morphine milligram equivalents for 90 consecutive days or longer (excluding patients in hospice) • Rate per 1,000 beneficiaries without cancer who received prescriptions for opioids from four or more prescribers or four or more pharmacies* • Rate per 1,000 beneficiaries without cancer who received prescriptions for opioids greater than 120 mg morphine equivalent dose (MED) for 90 consecutive days or longer, from four or more prescribers or four or more pharmacies*
Did demonstrations reduce concurrent use of opioids and benzodiazepines?	Annual monitoring report Supplemental: T-MSIS / MAX claims	• Percentage of beneficiaries age 18 and older with concurrent use of prescription opioids and benzodiazepines (patients with a cancer diagnosis or in hospice are excluded)
Did demonstrations increase continuity of pharmacotherapy for OUD?	Annual monitoring report Supplemental: T-MSIS / MAX claims	• Percentage of adults receiving pharmacotherapy for OUD who have at least 180 days of continuous treatment
Did demonstrations expand coverage of and access to naloxone for overdose reversal?	T-MSIS / MAX claims	• Number of prescriptions for naloxone per 1,000 beneficiaries with SUD-related diagnosis
	Medicaid state drug utilization data	• Number of prescriptions for naloxone

(continued)

Table 2-6. Implementation Outcome Questions, Data Sources, and Measures for the SUD Demonstration Meta-Analysis (continued)

Question	Source(s)	Measures
How is information technology being used to slow the rate of growth of individuals identified with SUD?	Quarterly monitoring report	• State identified metrics tracking SUD health IT plan progress
How is information technology being used to treat individuals identified with SUD effectively?	Quarterly monitoring report	• State identified metrics tracking SUD health IT plan progress
How is information technology being used to effectively monitor recovery supports and services for individuals identified with SUD?	Quarterly monitoring report	• State identified metrics tracking SUD health IT plan progress
SUD Milestone #6: Improved care coordination and transitions between levels of care		
Did demonstrations increase the percentage of patients identified with alcohol or drug use disorder who receive prescriptions for MAT or referral for treatment?	Annual monitoring report	• Patients who are identified with alcohol or drug use disorder who receive or refuse at discharge a prescription for FDA-approved medications for alcohol or drug use disorder, OR who receive or refuse a referral for addictions treatment* • Patients who are identified with alcohol or drug use disorder who receive a prescription for FDA-approved medications for alcohol or drug use disorder, OR a referral for addictions treatment*
Did demonstrations increase rates of follow-up after discharge from the ED for mental health or alcohol or other drug dependence?	Annual monitoring report Supplemental: T-MSIS / MAX claims	• Percentage of ED visits for beneficiaries who have a principal diagnosis of mental illness or AOD abuse or dependence and who had a follow-up visit for mental illness or AOD • Percentage of ED visits for mental illness for which the beneficiary received follow-up within 7 days of the ED visit • Percentage of ED visits for mental illness for which the beneficiary received follow-up within 30 days of the ED visit • Percentage of ED visits for which the beneficiary received follow-up visit for mental illness or AOD within 7 days of the ED visit • Percentage of ED visits for which the beneficiary received follow-up visit for mental illness or AOD within 30 days of the ED visit
Did demonstrations increase the percentage of beneficiaries in residential SUD facilities that receive outpatient services within 30 days following discharge?	T-MSIS / MAX claims	• Percentage of patients discharged from a residential SUD facility who receive outpatient services within 30 days following discharge

(continued)

Table 2-6. Implementation Outcome Questions, Data Sources, and Measures for the SUD Demonstration Meta-Analysis (continued)

Question	Source(s)	Measures
Did demonstrations increase the percentage of beneficiaries with inpatient SUD stays that receive outpatient services within 30 days following discharge?	T-MSIS / MAX claims	• Percentage of patients discharged from an inpatient SUD stay who receive outpatient services within 30 days following discharge

CSA = Controlled Substance Act; DEA = U.S. Drug Enforcement Agency; ED = emergency department; HCUP = Healthcare Cost and Utilization Project; MAX = Medicaid Analytic eXtract; OUD = opioid use disorder; SUD = substance use disorder (includes OUD); T-MSIS = Transformed Medicaid Statistical Information System

* Measures marked with an asterisk (*) are recommended by CMS; all other measures from the quarterly and annual monitoring reports are required

2.3.2 Demonstration Impacts

The meta-analysis questions, outcome measures, and potential data sources related to demonstration impacts are listed in ***Table 2-7***. Impact outcome measures are derived from the evaluation design technical assistance guidance provided to states by CMS. The core of the meta-analysis is focused on impacts on the six SUD demonstration goals and Medicaid costs from state evaluations and state monitoring data. Thus, our meta-analysis questions align directly with the questions the state evaluations and state monitoring reports will address. As noted in ***Table 2-7***, the state evaluation and monitoring reports are primary data sources for the meta-analysis, which we will supplement with other data sources. We anticipate that many state evaluations may not use a comparison group, which limits the rigor of the meta-analyses, and states may not report on the same metrics or define them the same way, which limits the number of outcomes we can test. Incorporating analyses of the supplemental data sources ensures that the meta-analysis can address each question and provide CMS with an understanding of the impacts of the demonstrations on each of the six SUD demonstration goals and Medicaid costs.

Table 2-7. Impact Questions, Data Sources, and Measures for the SUD Demonstration Meta-Analysis

Question	Source(s)	Measures
SUD Goal #1: Increase rates of identification, initiation, and engagement in treatment		
Did demonstrations improve rates of initiation and engagement of alcohol and other drug dependence treatment?	State evaluation reports Supplemental: T-MSIS / MAX claims	• Percentage of adults (ages 18+) diagnosed with alcohol or drug abuse or dependency (new episodes) who initiate treatment within 14 days of the index episode; treatment may be inpatient, outpatient, or partial hospitalization • Percentage of adults (ages 18+) diagnosed with alcohol or drug abuse or dependency (new episode) who initiate treatment and have two additional treatments within 30 days after the first treatment
	NSDUH	• Percentage of Medicaid-covered individuals with an SUD
Did demonstrations enhance provider and plan capabilities to screen/identify patients for engagement and intervention?	State evaluation reports Supplemental: T-MSIS / MAX Claims	• Percentage of providers indicated as providing screening, services, or referral to treatment • Percentage of enrollees with a screening claim • Percentage of enrollees who screen positive that are referred for services • Percentage of enrollees who receive services after referral
	TEDS	• Number of admissions to SUD specialty treatment • Number of Medicaid-covered admissions to SUD specialty treatment • Number of opioid-related admissions to SUD specialty treatment • Receipt of MAT
	Quarterly monitoring reports Supplemental: T-MSIS / MAX Medicaid claims	• Number of Medicaid beneficiaries with SUD diagnosis and a SUD-related service during the measurement period and/or in the 11 months before the measurement period (monthly) • Number of Medicaid beneficiaries assessed for SUD treatment needs • Number of Medicaid beneficiaries with newly initiated SUD treatment/diagnosis
	Annual monitoring reports Supplemental: T-MSIS / MAX claims	• Number of beneficiaries with SUD diagnosis and a SUD-related service during the measurement period and/or in the 12 months before the measurement period (annual) • Number of beneficiaries with a claim for residential treatment for SUD in an IMD during the reporting year

(continued)

Table 2-7. Impact Questions, Data Sources, and Measures for the SUD Demonstration Meta-Analysis (continued)

Question	Source(s)	Measures
SUD Goal #2: Increase adherence to and retention in treatment		
Did demonstrations improve continuity of pharmacotherapy for OUD	State evaluation reports Supplemental: T-MSIS / MAX claims	• Percentage of adults (ages 18+) receiving pharmacotherapy for OUD who have at least 180 days of continuous treatment
Did demonstrations increase rates of use of key treatment services?	State evaluation reports Supplemental: T-MSIS / MAX claims	• Percentage of beneficiaries with a SUD diagnosis including those with OUD who used the following services: - outpatient - intensive outpatient - MAT - residential treatment - medically supervised withdrawal management
Did demonstrations reduce the time to treatment?	State evaluation reports Supplemental: T-MSIS / MAX claims	• Average number of days between clinical assessment and first contact with a patient
Did demonstrations increase the rate of continuation of treatment?	State evaluation reports Supplemental: T-MSIS / MAX claims	• Average number of days between clinical assessment and first treatment
Did demonstrations increase patient engagement in treatment?	State evaluation reports Supplemental: T-MSIS / MAX claims	• Percentage of beneficiaries receiving treatment that completed 4+ treatment sessions within 30 days
Did demonstrations increase retention in treatment?	TEDS	• Successful treatment completion as percentage of total discharges • Average length of stay
SUD Goal #3: Reduce overdose deaths, particularly those due to opioids		
Did demonstrations reduce use of opioids at high dosage levels, excluding cancer patients?	State evaluation reports Supplemental: T-MSIS / MAX Claims	• Percentage of beneficiaries with high dosage opioid prescriptions among those with prescriptions filled on at least 2 separate dates for a minimum 15 days' supply
Did demonstrations reduce concurrent use of opioids and benzodiazepines?	State evaluation reports Supplemental: T-MSIS / MAX claims	• Percentage of beneficiaries with concurrent use of prescription opioids and benzodiazepines, among beneficiaries with 2 or more prescription claims for opioids on 2 or more days, with a supply of 15 or more days

(continued)

Table 2-7. Impact Questions, Data Sources, and Measures for the SUD Demonstration Meta-Analysis (continued)

Question	Source(s)	Measures
Did demonstrations reduce overdose deaths specifically overdose deaths due to any opioid?	State evaluation reports Annual monitoring reports Supplemental: T-MSIS / MAX claims HCUP	• Number of overdose deaths per 1,000 adult Medicaid beneficiaries • Number of overdose deaths among adult Medicaid beneficiaries
	CDC WONDER / NVSS	• Opioid-related overdose deaths per 100,000 population • Heroin-related overdose deaths per 100,000 population • Natural and semisynthetic opioid-related overdose deaths per 100,000 population • Methadone-related overdose deaths per 100,000 population • Non-methadone, synthetic opioid-related overdose deaths per 100,000 population
SUD Goal #4: Reduce preventable or medically inappropriate use of emergency department and inpatient hospital services		
Did demonstrations reduce SUD and opioid-related ED visits?	State evaluation reports Supplemental: T-MSIS / MAX claims	• Emergency department (ED) visits for SUD-related diagnoses and specifically for OUD among Medicaid beneficiaries per 1,000 member months
	HCUP	• ED visits for SUD-related diagnoses and specifically for OUD among Medicaid beneficiaries • Opioid-related ED visits per 100,000 population • Other SUD-related ED visits per 100,000 population
Did demonstrations reduce SUD and opioid-related inpatient admissions?	State evaluation reports Supplemental: T-MSIS / MAX claims	• Inpatient admissions for SUD-related diagnoses and specifically OUD among Medicaid beneficiaries per 1,000 member months
	HCUP	• Inpatient admissions for SUD-related diagnoses and specifically OUD among Medicaid beneficiaries • Opioid-related inpatient discharges per 100,000 population • Other SUD-related inpatient discharges per 100,000 population

(continued)

Table 2-7. Impact Questions, Data Sources, and Measures for the SUD Demonstration Meta-Analysis (continued)

Question	Source(s)	Measures
Did demonstrations increase follow-up after discharge from an ED visit for alcohol or other drug dependence?	State evaluation reports	• Percentage of beneficiaries with outpatient visits, intensive outpatient encounters, or partial hospitalizations with a primary diagnosis of alcohol or other drug dependence within 7 days & 30 days following ED discharge for alcohol or other drug dependence
	Supplemental: T-MSIS / MAX claims	• Percentage of discharges with an SUD-related diagnosis resulting in a medical or behavioral health encounter within 7 days • Percentage of discharges with an SUD-related diagnosis resulting in a medical or behavioral health encounter within 30 days • Percentage of discharges with an OUD-related diagnosis resulting in a medical or behavioral health encounter within 7 days • Percentage of discharges with an OUD-related diagnosis resulting in a medical or behavioral health encounter within 30 days
SUD Goal #5: Reduce preventable or medically inappropriate readmissions to the same or higher level of care		
Did demonstrations reduce the 30-day readmission rate following hospitalization for a SUD-related diagnosis and specifically for OUD	State evaluation reports Supplemental: T-MSIS / MAX Claims	• Percentage of discharges with a SUD-related diagnosis resulting in readmission within 30 days • Percentage of discharges with an OUD-related diagnosis resulting in readmission within 30 days
SUD Goal #6: Improve access to care for physical health conditions among beneficiaries		
Did demonstrations improve access to preventive and/or ambulatory health services for adult Medicaid beneficiaries with SUD/OUD?	State evaluation reports Annual monitoring reports	• Percentage of Medicaid beneficiaries with a SUD diagnosis who had an ambulatory or preventive care visit
	Supplemental: T-MSIS / MAX claims	• Percentage of beneficiaries with a SUD diagnosis who use physical health care services • Percentage of beneficiaries with an OUD diagnosis who use physical health care services
Did demonstrations improve screening of tobacco use among those with SUD/OUD?	State evaluation reports Supplemental: T-MSIS / MAX claims	• Percentage of beneficiaries with a SUD diagnosis who are screened for tobacco use • Percentage of beneficiaries with an OUD diagnosis who are screened for tobacco use
Did demonstrations improve screening of unhealthy alcohol use among those with SUD/OUD?	State evaluation reports Supplemental: T-MSIS / MAX claims	• Percentage of patients with a SUD diagnosis who are screened for unhealthy alcohol use & received brief counseling if screen was positive

(continued)

Table 2-7. Impact Questions, Data Sources, and Measures for the SUD Demonstration Meta-Analysis (continued)

Question	Source(s)	Measures
Did demonstrations improve screening of depression among those with SUD/OUD?	State evaluation reports Supplemental: T-MSIS / MAX claims	• Percentage of patients with a SUD diagnosis who are screened for depression
Expenditure Impacts		
How did demonstrations change total and SUD-related Medicaid spending change?	State evaluation reports Annual monitoring reports Supplemental: T-MSIS / MAX claims	• Total Medicaid spending per beneficiary with a SUD per month • Total federal Medicaid spending per beneficiary with a SUD per month • Medicaid spending on IMD per beneficiary with a SUD per month • Medicaid spending on other SUD treatment per beneficiary with a SUD per month • Medicaid spending on non-SUD services per beneficiary with a SUD per month • Medicaid spending on inpatient services per beneficiary with a SUD per month • Medicaid spending on ED visits per beneficiary with a SUD per month • Medicaid spending on other outpatient services per beneficiary with a SUD per month • Medicaid spending on pharmacy per beneficiary with a SUD per month • Medicaid spending on long-term care per beneficiary with a SUD per month

ED = emergency department; HCUP = Healthcare Cost and Utilization Program; MAT = medication-assisted treatment; MAX = Medicaid Analytic eXtract; NQF = National Quality Forum; NVSS = National Vital Statistics System; OUD = opioid use disorder; PQA = Pharmacy Quality Alliance; SUD = substance use disorder; TEDS = Treatment Episode Data Set; T-MSIS = Transformed Medicaid Statistical Information System

SECTION 3.
META-ANALYSIS METHODOLOGY

3.1 Qualitative Analysis of Standard Abstraction Files and Qualitative Data

To manage and analyze secondary qualitative data, we will use NVivo 11 software (www.qsrinternational.com). This software is designed for qualitative and mixed methods research and allows integration of other data sources and comparisons within and across demonstrations over time.[8-10] NVivo facilitates analysis by allowing us to compare and contrast information by research question and by data source or respondent type. After data entry and coding, the data can contribute to synthesis reports, comparative case studies, and qualitative comparative analysis (QCA). In addition to conducting thematic cross-state analysis of qualitative results, whenever possible narrative data will be converted to numeric values for use in meta-regression and stratified analysis.[11]

3.2 Meta-Analysis of State Evaluation Findings

Each state demonstration will have an independent evaluator that conducts statistical analyses of demonstration effects. The evaluation guidance recommends that states use a regression-adjusted difference-in-differences model to estimate demonstration effects. The resulting effects are adjusted demonstration versus non-demonstration state values. These state-specific effects and their standard errors, which we will abstract from the evaluation reports in each state, are the primary inputs for the meta-analysis of demonstration impacts.

Preliminary evaluation design choices have been based on review of demonstration designs already approved, their STCs, quarterly reports when available, and information in state applications for demonstrations pending approval. As additional demonstrations are approved and implementation plans and evaluation design plans become available, the meta-analysis design may be revised to account for significant changes in demonstration approaches. A main constraint to meta-analysis is the data available, including the evaluation designs that the states implement.

The SUD meta-analysis will focus on the following 4 research questions and 11 outcomes aligned to the CMS's evaluation questions in the evaluation guidance TA.

Research Question #1: Does the demonstration increase access to and utilization of SUD treatment services?

(G1): Percentage of beneficiaries who initiate and engage in MAT or other SUD treatment

(G2): Percentage of beneficiaries with 180 days of continuous pharmacotherapy use

(G2): Percentage of beneficiaries with a SUD diagnosis who used MAT

(G4): Rate of SUD-related ED visits per 1,000 member months

(G4): Rate of SUD-related inpatient hospital visits per 1,000 member months

Research Question #2: Do enrollees receiving SUD services experience improved health outcomes?

(G6): Percentage of beneficiaries with a SUD who had an ambulatory or preventive care visit

(G5): 30-day rehospitalization rate

Research Question #3: Are rates of opioid-related overdose deaths impacted by the demonstration?

(G3): Opioid overdose deaths per 1,000 beneficiaries

Research Question #4: Are total and SUD Medicaid expenditures impacted by the demonstration?

Total expenditures on a per member per month (PMPM) basis

SUD expenditures (PMPM)

Expenditures by type of service (PMPM)

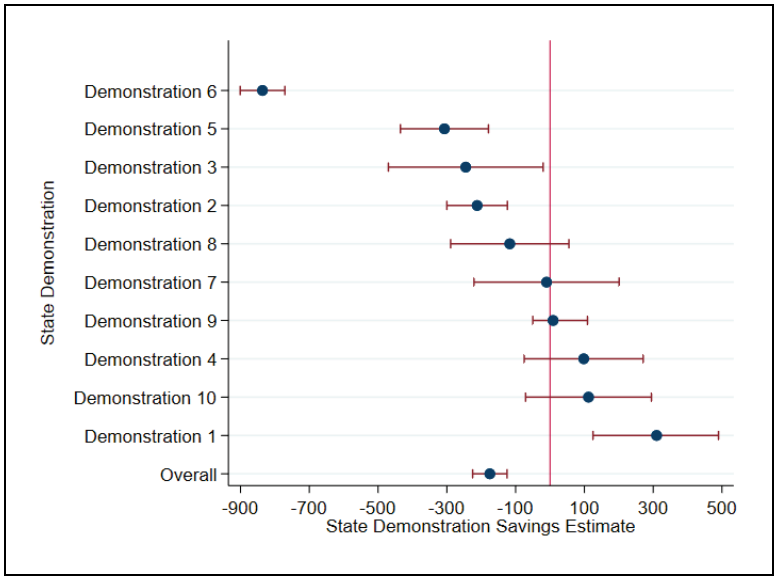
We will explore additional outcomes listed in Section 2.3.2 as feasible.

3.2.1 Descriptive Analysis

The meta-analysis for each key outcome will be performed in stages. The first stage in our meta-analysis will be descriptive. We will gather the results from the state evaluations and standardize them so they can be compared. We will then calculate a cross-state estimate of the overall demonstration effect. This mean value is derived by weighting the individual state effect sizes by the precision of each estimate so that more precise estimates from larger programs are given greater weight. The weighted mean will be tested to determine whether it differs significantly from zero.

We will then display results for individual states in the form of a forest plot, as shown in ***Figure 3-1***. These plots can illustrate the state evaluation results by feature of the demonstration. This provides a high-level sense of the difference in state results (on a particular outcome) by demonstration feature. Our experience has been that multisite meta-analyses typically yield the general pattern shown in ***Figure 3-1***, in which most interventions have insignificant, near-zero impacts with smaller numbers of successful and unsuccessful programs. We expect to see a similar pattern among the SUD demonstrations.

Figure 3-1. Forest Plot Showing State Demonstration Effect Sizes and Standard Errors for Estimated Medicaid Savings



The next stage in the meta-analysis is to determine the extent to which the variance of the overall effect represents random measurement error or whether there are systematic differences between state demonstrations. Generally, we expect that the results will be different for each state, which suggests that there is some unexplained systematic variation underlying the results. If this is the case, the analysis will then proceed to the next stage to identify the key demonstration features that affect the magnitude of different demonstration effects in different states. We describe two methods for assessing their relations with outcomes (meta-regression and qualitative case comparisons).

3.2.2 Meta-Regression

We will use meta-regression to identify the effect of key demonstration features on outcomes. In this stage, demonstration effect sizes become the dependent variables in a meta-regression model. ***Table 3-1*** describes specific hypotheses that we will address with the meta-regressions, linking the outcomes and the features that drive changes in outcomes. The explanatory variables in this model are factors that are hypothesized to influence the magnitude of the demonstration effect, including both specific demonstration components and structural, non-demonstration factors. The meta-regression will examine the influence of variations in SUD

demonstration features on the magnitude of change in each of the 11 outcomes. The demonstration features were described in Section 2.2.1.

The coding for these features will be drawn from RTI's abstraction of demonstration attributes described in Section 2 and operationalized in ***Table 3-1***.

Table 3-1. SUD Demonstration Features and Corresponding Measures for the Meta-Regressions

SUD Demonstration Feature	Potential Measures
Coverage of SUD services	• Indicator for adding coverage of residential SUD services • Indicator for adding coverage of intensive outpatient/partial hospitalization SUD services • Indicator for adding coverage of methadone for OUD and opioid treatment programs • Indicator for adding coverage recovery support services
Changes in patient placement criteria	• Indicator that the state newly implemented or updated patient placement criteria and utilization management policies
Changes in residential provider standards	• Indicator that the state newly implemented or updated residential provider standards • Indicator that the state newly required residential providers to offer MAT
Reimbursement changes	• Indicator for increasing reimbursement for SUD services
Provider capacity	• Relative change in the number of buprenorphine-waived providers • Relative change in number of residential SUD providers accepting Medicaid • Relative change in number of outpatient SUD providers accepting Medicaid
Care coordination	• Indicator for adding coverage of SUD case management • Indicator for other increases in care coordination, such as centralized care coordination
State evaluation methodology	• Indicator for using a difference-in-differences analysis • Indicator for using a comparison group
Medicaid expansion status	• Indicator for being a Medicaid expansion state

The meta-regression will be based on the following general equation:

$$Y_i = \alpha + \Sigma_j \beta_j X_{ji} + \Sigma_k \lambda_k Z_{ki} + \varepsilon_i$$

Y_i = the outcome effect size for the i-th demonstration state

α = an intercept term

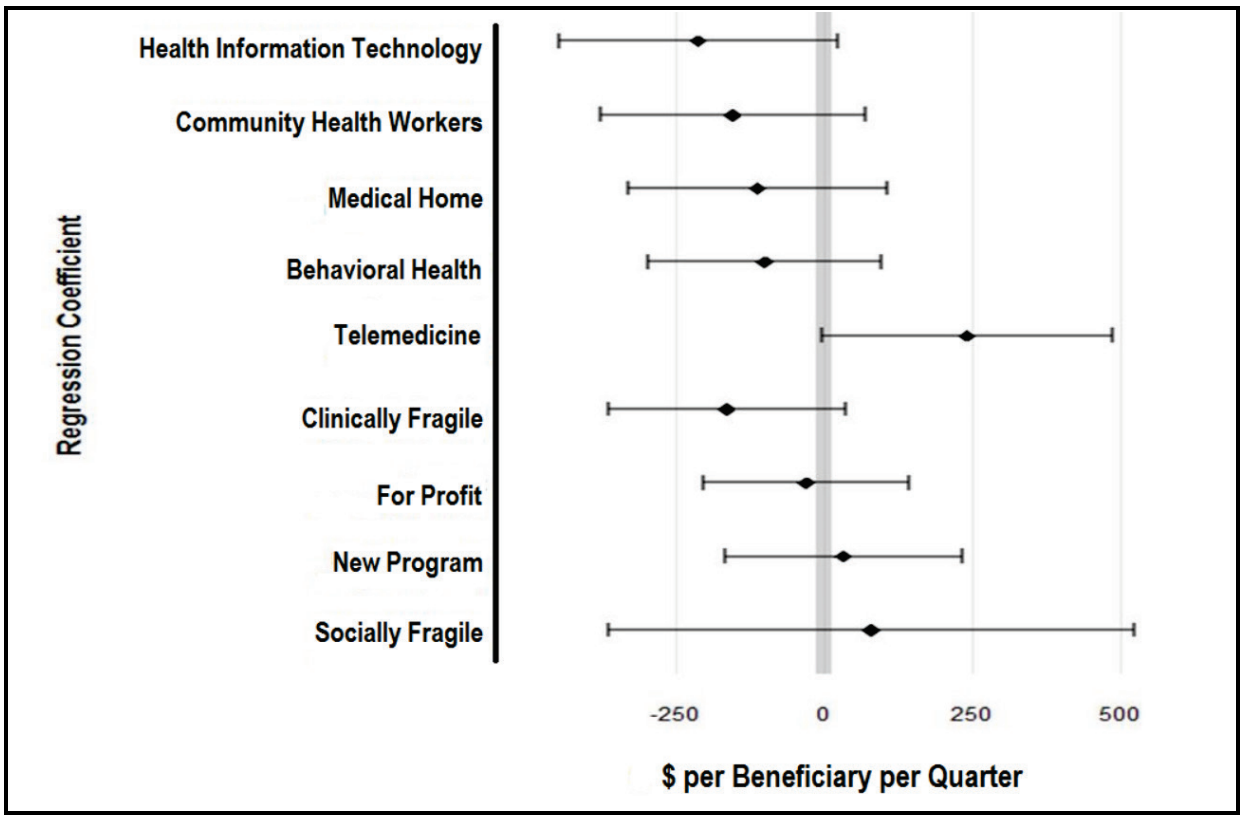
X_{ji} = a set of j components of the i-th state's demonstration implementation,

Z_{ki} = a set of k non-demonstration features of the i-th state

ε_i = an error term.

An example summary of meta-regression results is shown in ***Figure 3-2***, which is taken from RTI's meta-analysis of system delivery innovations for the Health Care Innovation Awards. This graph is also in the form of a forest plot, but here each row represents a demonstration feature showing that feature's regression coefficient (and standard error) on total cost of care.

Figure 3-2. Example Meta-Regression Results (Health Care Innovation Awards)



Because we do not expect the demonstration effect results from state evaluations to be available until late in the contract period, meta-analyses of these effects will be presented in the Summative Evaluation Report.

Limitations of meta-analysis

While meta-analysis and meta-regression are powerful tools for synthesizing SUD demonstration effects, the method may be limited in several respects in our evaluation. The primary limitation is that section 1115 demonstrations are state-based, so there will be only a comparatively small number of observations (in this case, state-specific effect sizes) available for analysis. The most significant implication of small samples is that the standard errors associated with meta-regression regression coefficients will be comparatively large. What might be considered a substantial impact (say savings of $300/beneficiary/year) for a particular

demonstration feature, may therefore not achieve statistical significance because the impact estimate is imprecise.

Like all observational studies, a second issue is that intervention effects are highly dependent on the extent to which comparison areas provide a valid counterfactual—that is, the extent to which the comparison group's outcome is a close proxy for what would have happened in the state in the absence of a section 1115 demonstration. Differences in the quality of comparison groups will inject error into the cross-state meta-analysis.

A third potential limitation is that the method assumes that intervention effects (the meta-regression outcome) are being calculated in the same way in each state. Variations in measures and analysis methods will contribute error to the meta-analysis. Some features, such as within-state versus out-of-state comparison groups can be controlled for, but other types of analytic variation that are not directly observed may produce invalid outcomes or biased estimates of the influence of program features and design elements. In this respect, CMS guidance to states regarding metrics is helpful for standardizing the computation of costs and expenditures and reducing certain measurement errors.

Because we are aware of these potential limitations, we suggest T-MSIS analyses that can supplement those provided by the states that allow us to standardize the calculation methods for key variables. ***Table 2-7*** shows outcomes that can be addressed using T-MSIS data. As discussed in ***Section 2.2.2***, the quality and completeness of T-MSIS data is evolving and we will assess its usability for specific outcomes and individual states as data are updated. We also are able to assess or create or support the creation of comparison groups to calculate treatment effects if states are unable to do this. The remaining limitation that we will try to address in early conversations with states is data availability, especially for 5052 states where evaluation is not a state requirement. We believe that states will want to learn from each other and will encourage the sharing of experiences and data.

3.3 Analysis of Quarterly and Annual Monitoring Metrics

Although they are limited because they lack comparison group data, monitoring metrics offer an earlier look at demonstration impacts than state evaluation reports, which will not be available for several years. In addition, monitoring metrics assess performance on a broader set of common outcomes than will be possible with the state evaluations, including implementation outcomes.

RTI's analysis will explore whether differences across states in performance on monitoring metrics listed in ***Tables 2-6*** and ***2-7*** are associated with state demonstration features and contextual factors, such as those described in Section 2.2.1. In addition to descriptive analyses comparing metrics across groups of states with different demonstration features, we can also use regression models to test for differences in trends in metrics between states with differing features. The association between state performance on monitoring metrics and demonstration features can be analyzed using comparative case study methods (described in Section 3.6).

3.4 Latent Trajectory Models

Latent trajectory models (LTMs) are an alternative meta-analytic method for evaluations that do not have a comparison group and cannot generate difference-in-difference estimates. LTMs use structural equation models to summarize patterns of change over time. The trajectory is generated by three latent variables—an intercept factor, an overall slope factor, and a slope factor for the intervention period. The intercept factor represents the initial level (e.g., first quarter) of an outcome, while the slope factors indicate how quickly the outcome increases or decreases over time. All factors are modeled as random coefficients, allowing for individual state demonstration differences in initial status and rates of change. The general system of equations for this model is as follows:

$$Y(t) = b_1*\text{INTERCEPT} + b_2*(t)*\text{SLOPE}$$

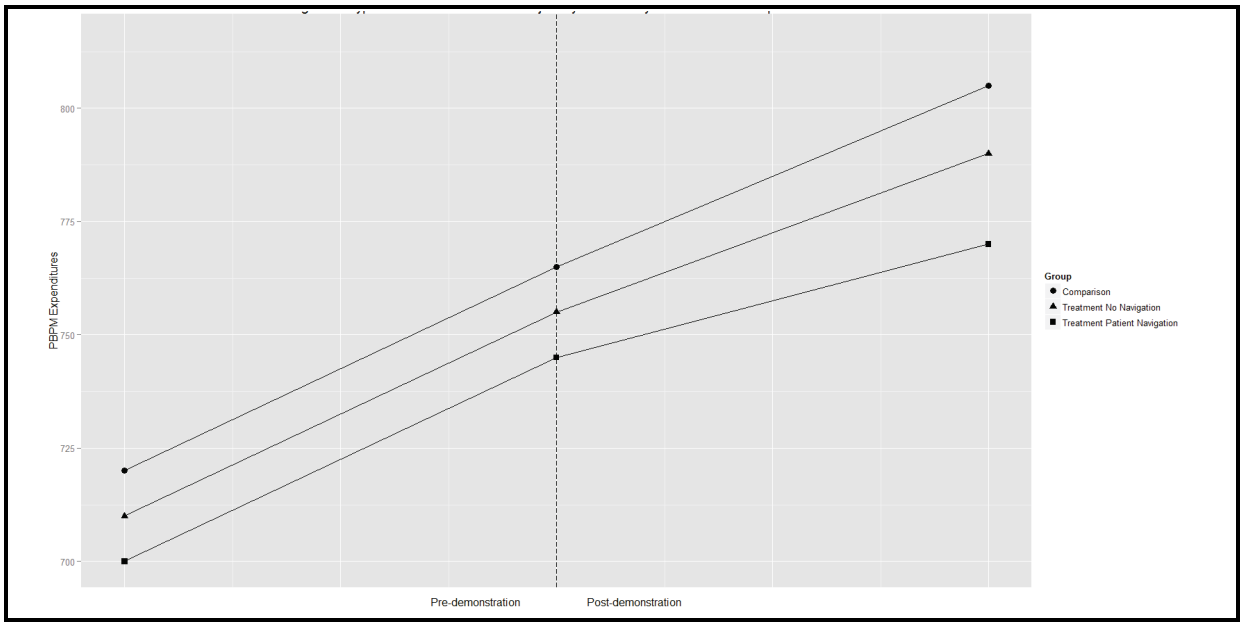
$$\text{INTERCEPT} = b_3 + b_4*X1 + b_5*X2 + b_6*X3$$

$$\text{SLOPE} = b_7 + b_8*X1 + b_9*X2 + b_{10}*X3$$

Where Y(t) is the outcome measure at time t, X1 is the binary treatment effect, X2 and X3 are covariates, and b_1 to b_{10} are regression coefficients.

Differences among state trajectories may be caused by covariates that modify the intercept or slope factors. The impact of demonstration features is estimated by their effects on the intercept and slope factors. The estimated LTM coefficients can then be used to plot predicted trajectories, as in ***Figure 3-3***, which shows hypothetical effects for health care expenditure trajectories. These results in this hypothetical example indicate that overall expenditure rates increased at a slower rate during the demonstration period, that expenditures were lower for treatment awardees than for comparison groups, and that navigation programs produced the lowest expenditures.

Figure 3-3. Hypothetical Results of Latent Trajectory Model Analysis of Health Care Expenditures



3.5 Supplemental Analyses of National Data Sources

National data sets can provide impact estimates for SUD demonstrations that complement the state evaluations. We can analyze these data sets to supplement the meta-analysis of state evaluation findings, for example to support subgroup analyses, to incorporate comparison group data if state evaluations are not able to, or to create standardized outcomes if measures used by state evaluators are not comparable.

If comparison beneficiaries can be identified, summary difference-in-differences impact estimates can be developed for these states and standard meta-analytic methods can be used to contrast state performance and identify reasons state performance may differ across outcomes. These types of data could also be used for pooled data analysis (e.g., hierarchical linear modeling, difference-in-difference-in-differences analysis) to identify subpopulations for whom the demonstrations are more or less effective, and to test if subpopulations are differentially impacted by demonstrations with specific features. Outcomes related to demonstration milestones can be analyzed in the same manner as quarterly and annual monitoring metrics.

In the event that we cannot identify a valid comparison group, we will use latent trajectory modeling approaches that allow us to statistically test for changes in milestones and outcomes without needing to meet the assumptions for difference-in-differences analyses

(described in Section 3.4). From these analyses, we could potentially infer whether we see faster growth in outcome trends for the demonstration states relative to the non-demonstration states.

3.6 Qualitative Case Comparisons

Qualitative case comparisons is an alternative approach for examining the relationship between demonstration features and outcomes. We describe two approaches to qualitative comparisons: QCA and comparative case study. We would use QCA as a complementary method to build upon qualitative and quantitative evaluation findings. We will implement a conventional exploratory sequential QCA design by using quantitative and qualitative analyses to inform the state demonstration features to include in the QCA. Subsequently will use QCA to enrich our understanding of qualitative and quantitative findings as well as identify demonstration states, features, or other contextual factors needing additional qualitative data collection.[12]

Qualitative comparative analysis

QCA is a case-oriented approach that examines relationships between explanatory factors (called "conditions" in QCA) and an outcome using set theory, which is a branch of mathematics and symbolic logic that deals with the nature and relations of sets.[13] QCA can be used with small to medium-N studies (i.e., 10 to 100 cases) and is useful for understanding causally complex phenomena. QCA is advantageous for this meta-analysis because (1) the number of state demonstrations will be relatively small, (2) QCA can accommodate qualitative and quantitative data, and (3) the causal relationships underlying the demonstration conceptual model is complex.

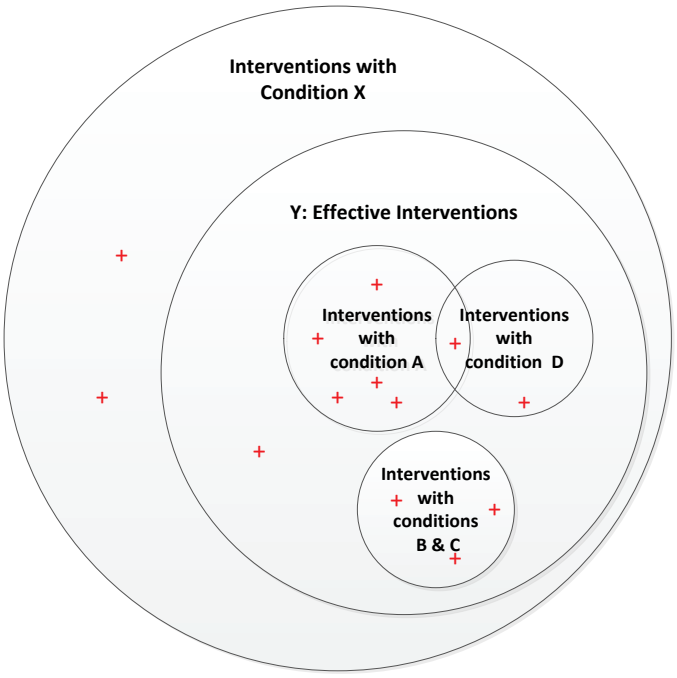
We can use QCA as an alternative to meta-regression for examining relationships between demonstration features and outcomes. Like meta-regression, use of QCA depends on variation in features and outcomes across state demonstrations. QCA will allow us to identify the combination of demonstration components, context, or other features that have a set-relationship (as opposed to a probabilistic relationship) with a specified outcome.[b] The results of a QCA analysis are statements of necessity and sufficiency, expressed as text, solution formulas, or in tabular or graphic formats.

A simplified example of how set relationships can be used to identify necessary and sufficient conditions for an outcome is depicted in ***Figure 3-4***. In this hypothetical example, five demonstration features ("condition sets") are included (A, B, C, D, and X) along with an

[b] QCA is not inferential, is not used for statistical hypothesis testing, and does not require any statistical assumptions about the underlying data as do quantitative approaches like meta-regression.

outcome set (Y). QCA determines whether superset and subset relationships between condition sets alone and in combination with each other exist with respect to the outcome set. In this example, X is a necessary condition for the outcome Y. A or D or the combination of B and C are sufficient for the outcome Y. In other words, condition X has to be present if a case is going to have any chance at achieving the outcome, but it will not guarantee the outcome. Likewise, a case can achieve the outcome in any one of three sufficient ways (having A, having D, or having both B and C).

Figure 3-4. Depiction of Simplified Set Relationships to Identify Necessary and Sufficient Conditions



Necessary conditions are supersets of an outcome set.

- Condition X is a superset of the outcome set Y.
- X is necessary for Y. However, X does not guarantee the outcome Y. A case can have X, but still be outside of the set Y.

Sufficient conditions are subsets of an outcome set.

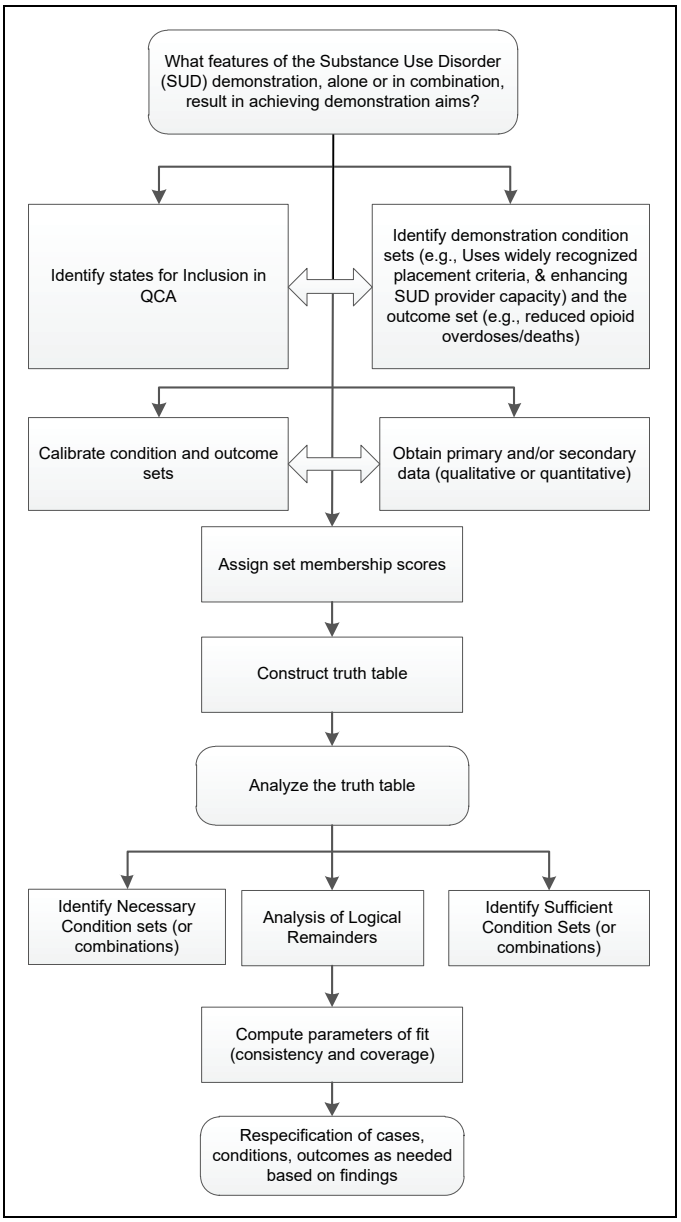
- Conditions A, D, and the combination of B and C are subsets of the outcome set Y.
- Any one (A, or D, or B & C) of the sufficient conditions is linked to the outcome Y. All cases with any one of these conditions are within the set Y.

Note: A, B, C, D and X can represent different intervention components, contextual features, or other characteristics of the population, setting, or intervention. Each "+" represents a unique case.

To conduct QCA for SUD demonstrations, we first would assess the demonstration data available for inclusion into our QCA models. Variation among demonstration conditions is required to demonstrate necessary or sufficient combinations leading to the desired outcome. After assessing available demonstration data, we would identify the research question or questions suitable for addressing using QCA. Lastly, we would identify outcomes to include in the QCA model reflective of a demonstration's success.

We would follow a standard approach to QCA based on existing best practices for the design and conduct of QCA studies.[13] As an example, ***Figure 3-5*** shows the analytic steps involved in conducting this type of analysis for the SUD demonstration.

Figure 3-5. Approach to Using QCA for SUD Demonstration



For each demonstration, the cases would include the SUD demonstration states that include specific features as a primary component of the demonstration being implemented. We will work with CMS to identify demonstration features we should consider as a primary component of the demonstration and will construct QCA conditions sets based on selected key demonstration outcomes.

In the most basic form of QCA, called "crisp set" QCA, conditions and outcomes are calibrated dichotomously and a set membership score of 0 (condition or outcome is low or absent in the case) or 1 (condition or outcome is present or high in the case) is assigned for every condition and outcome across all included cases. In "fuzzy set" QCA, conditions can be assigned set membership scores using any number between 0 and 1 to establish degrees of qualitative differences, beyond simply whether a condition is present or absent. The number of conditions used in a QCA is partly based on the underlying causal theories of the intervention, but is also limited by the number of available cases, with a recommended ratio of three or four cases per condition included.[14] The next step involves the construction of a data matrix that contains set membership scores for the conditions and outcome for all included cases. This raw data matrix is then converted into an analytic device used in formal logic known as a truth table. The truth table contains 2K number of rows,

where K represents the number of conditions included in the analysis and represents all of the logically possible combinations of conditions and outcome. However, all logically possible combinations of conditions and outcome are rarely found in empiric analyses. Some cases will fall into the same truth table row because they may share the same set membership scores across all conditions and the outcome, and other rows will be empty because no cases in the sample exhibit those precise combinations of conditions.

Once the truth table is constructed, the next step is conducting truth table analyses. This involves identifying necessary and sufficient conditions alone or in combination with each other, also known as "solutions." We will use R software to conduct the truth table analyses. Identification of necessary and sufficient conditions can be accomplished using the R QCA and Set Methods Package allowing us to apply Boolean minimization procedures that reduce logically redundant solution terms from the final solution. For example, if we find a condition that is present in some cases with the outcome and absent in other cases with the outcome, then that condition is irrelevant to the statements of necessity and sufficiency and can be dropped from the solution. Although we will most likely specify five to seven conditions for inclusion for each QCA, the number of conditions that will appear in the final solution may be considerably less because of the removal of logically irrelevant conditions.

As is the case with quantitative analyses, empiric data from complex phenomena will rarely demonstrate perfect subset and superset relationships. Thus, calculating parameters of fit (known as consistency and coverage) and conducting a type of sensitivity analysis using the empty truth table rows (also known as logical remainders) is important for evaluating the robustness and relevance of findings. Although we have characterized the steps we will follow in ***Figure 3-5*** and text as somewhat sequential, as is typical of the qualitative research tradition, an iterative process that includes respecification and reanalysis will likely occur as we seek to relate findings back to individual cases or integrate findings with other aspects of the meta-analysis.

Limitations of QCA. Despite its advantages as a method, QCA does have some limitations. A QCA model cannot accommodate a large number of conditions. The method examines every possible combination of conditions in a model; thus, adding a condition to the model translates into an exponential increase in the number of possible combinations (i.e., 5 conditions = 32 possible combinations, 6 conditions= 64 possible combinations). Too many conditions presents the problem of limited diversity (i.e., not having cases representing each of the possible combinations). To manage this challenge, we will narrow potential conditions by assessing which condition sets are higher priority to CMS and are feasible to collect across all demonstrations. We will also consider creating "superconditions" that combine two or more

conditions into one composite condition, such as baseline service comprehensiveness. Lastly, a two-step approach to QCA could be considered, allowing for only those conditions identified as relevant in the first step to be specified in the second step.

Cases also need to have some variation in the outcome that we specify. If few demonstrations have better impacts, we will not have enough variation to identify a solution. If this occurs, our efforts will focus largely on evaluating implementation effectiveness to determine conditions associated with implementation success.

Comparative case study analysis

Comparative case study analysis is an alternative approach for examining the relationships between demonstration features and demonstration outcomes. Comparative case study is a data analysis methodology that systematically examines similarities and differences across cases.[15] In our analyses, state demonstrations serve as the case through which we will make comparisons. We will create a database that includes demonstration features (or case study constructs) that we believe are relevant to the outcomes and research questions. Once in the database, we will generate a case study comparison grid similar to the example in ***Table 3-2***. Our team will use the comparison grid to systematically compare demonstrations across features to identify patterns and points of similarity and difference.[16] We will then use outcome data to identify whether relationships or patterns exist between the presence or absence of specific features and the impacts observed. Unlike QCA, comparative case study examines relationships between individual features, rather than sets of features.

Table 3-2. Example of a Comparison Case Study Grid for SUD Demonstrations

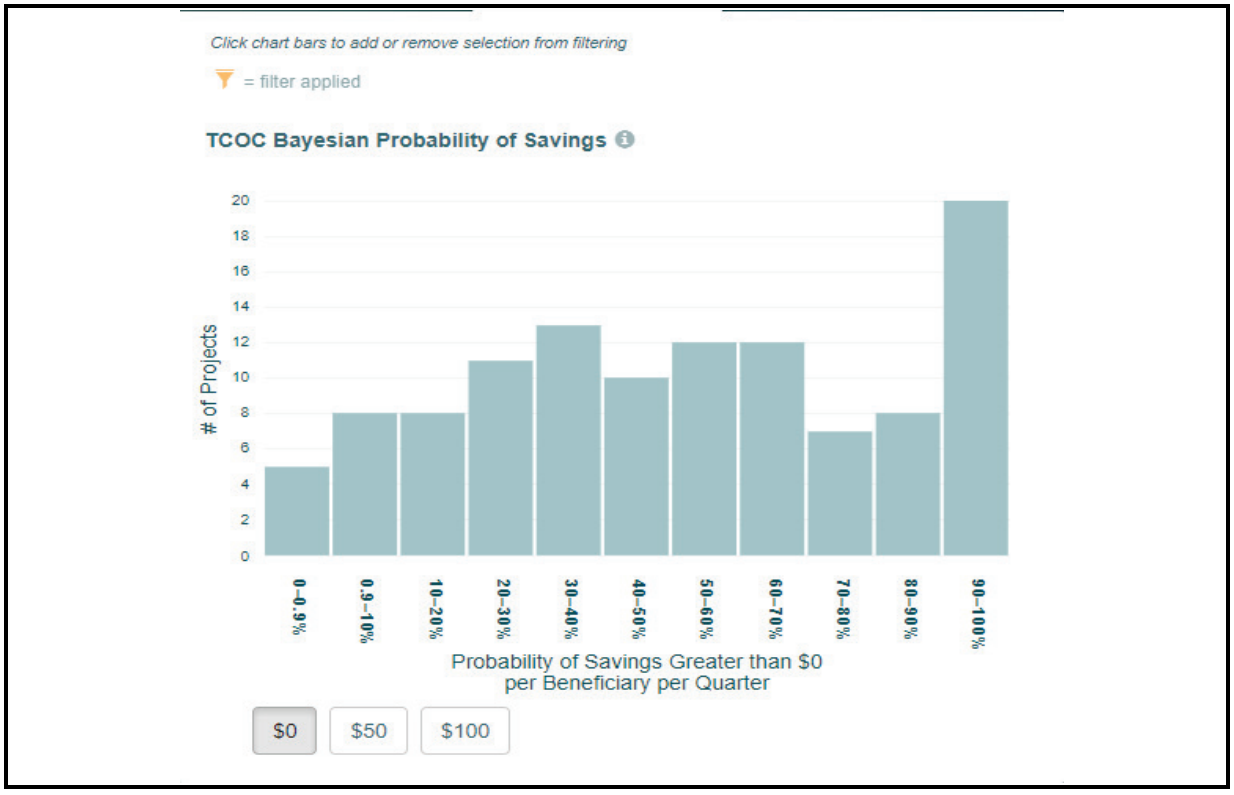
	Reduction in SUD Expenditures				
	Yes			**No**	
State Demonstration Features	**State A**	**State B**	**State C**	**State D**	**State E**
Medicaid expansion status	●	●	●	○	
Baseline service comprehensiveness	●	●	○	●	○
Change in patient placement criteria and utilization management policies	●	○	●	○	
Change in residential provider standards	●	○	●		

● = Demonstration feature is present in this state; ○ = feature is partially present in this state; empty cell = feature is not present in this state.

3.7 Bayesian Analysis

Bayesian methods provide the flexibility to go beyond answering "Are the results statistically significant?" to addressing questions such as "What is the probability of saving $10 or more per beneficiary per quarter?" and "What is the probability that employment increases by 5% over the evaluation period?" To provide easily understood evaluation results that can be used for policy decisions about the demonstrations, we can produce probabilistic estimates for the key evaluation outcomes through Bayesian estimation and related techniques. Probabilistic findings can be displayed in concise graphics, like ***Figure 3-6***.

Figure 3-6. Example of Displaying Probabilistic Findings



3.8 Additional Supplemental Analyses

3.8.1 Assessment of Baseline Trends

Section 2.2.2 described numerous national data sets that can be used to support the meta-analysis of the SUD demonstrations. RTI's initial analyses of national data sources will begin by assessing data quality for demonstration and non-demonstration states, including data

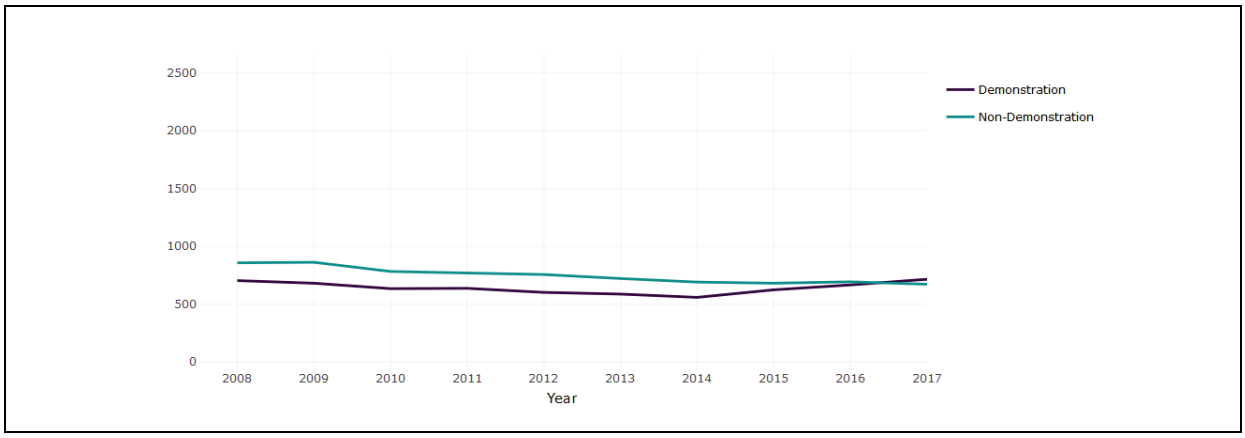
completeness and data quality. A preliminary analysis of data quality in several data sets is provided in ***Appendix E***.

We will proceed with analyses of selected data sets to characterize baseline conditions in demonstration and non-demonstration states, helping us to understand differences across demonstration states that might affect implementation of demonstration activities and outcomes. As an example, with substantial federal and state resources and policy changes to address the opioid epidemic, assessing changes among non-Medicaid-covered groups may capture important secular trends in demonstration outcomes and provide robustness checks for meta-analytic models.

We will produce annual averages for selected measures for demonstration and non-demonstration states, as well as national benchmark averages, for the national data source measures listed in the ***Table 2-6*** and ***Table 2-7***. This analysis can identify states with high potential for change (i.e., low number of Medicaid providers pre-demonstration) and less potential for change (i.e., high number of Medicaid providers pre-demonstration) and allows us to assess pre-demonstration trends that will inform the appropriate modeling approaching. As discussed in Section 2.2.2, before proceeding with baseline analyses of T-MSIS, we will conduct analyses to determine whether the data quality and completeness is sufficient to support analysis of SUD demonstration outcomes.

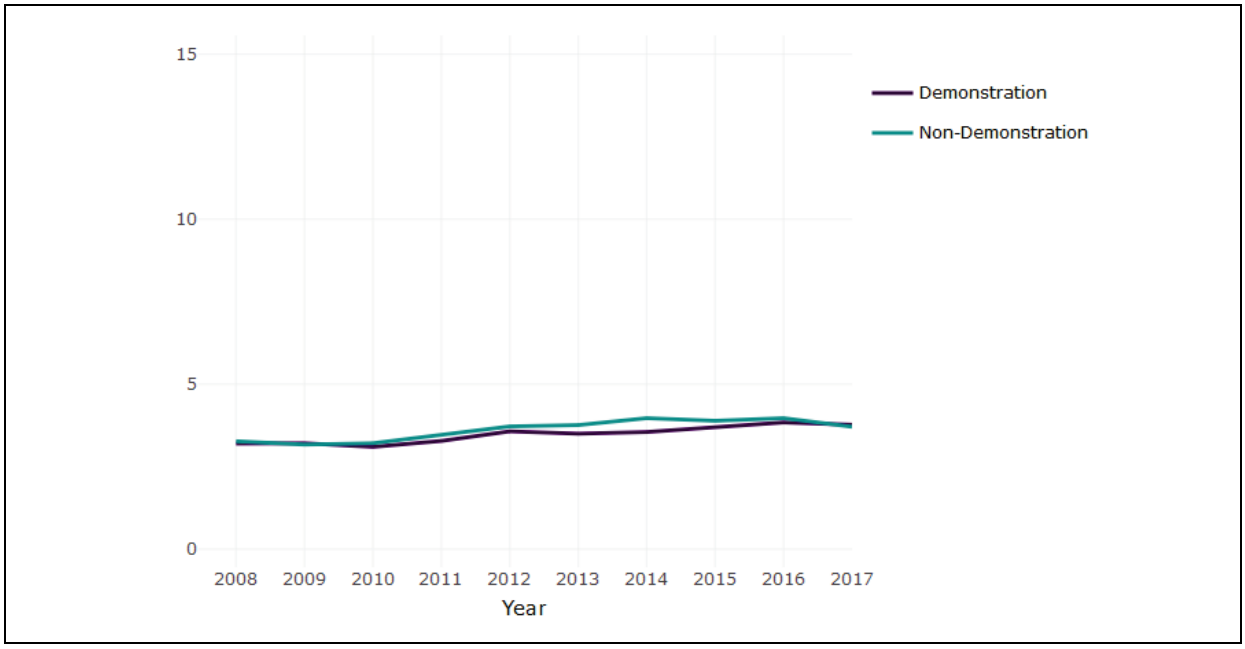
We provide briefly here a preliminary parallel trends analyses for TEDS, N-SSATS, and CDC WONDER. ***Figure 3-7*** shows average SUD admissions per 100,000 population from 2008 through 2017 using the TEDS data across demonstration and non-demonstration states—an outcome related to Milestone #1. Ideally, the lines should be parallel before the demonstrations start—California was the first state that started in 2016. The lines for the demonstration and non-demonstrations states are approximately parallel from 2008 to 2014. From 2014 to 2017, the trend line for demonstration states shifts upward and the trend for the non-demonstration states does not change. We used a regression model to assess whether there was a statistically significant difference in the trend lines and did not find evidence for a difference. While the trends pass a statistical test, this re-emphasizes that there may be potential underlying differences in demonstration and non-demonstration leading up to the start of the SUD demonstrations. For example, a number of demonstration states expanded their Medicaid programs, which could increase the number of admissions. We investigated this hypothesis and did not find expansion status to explain the rise in the trend for demonstration states.

Figure 3-7. SUD Admissions per 100,000 Population in TEDS, 2008–2017



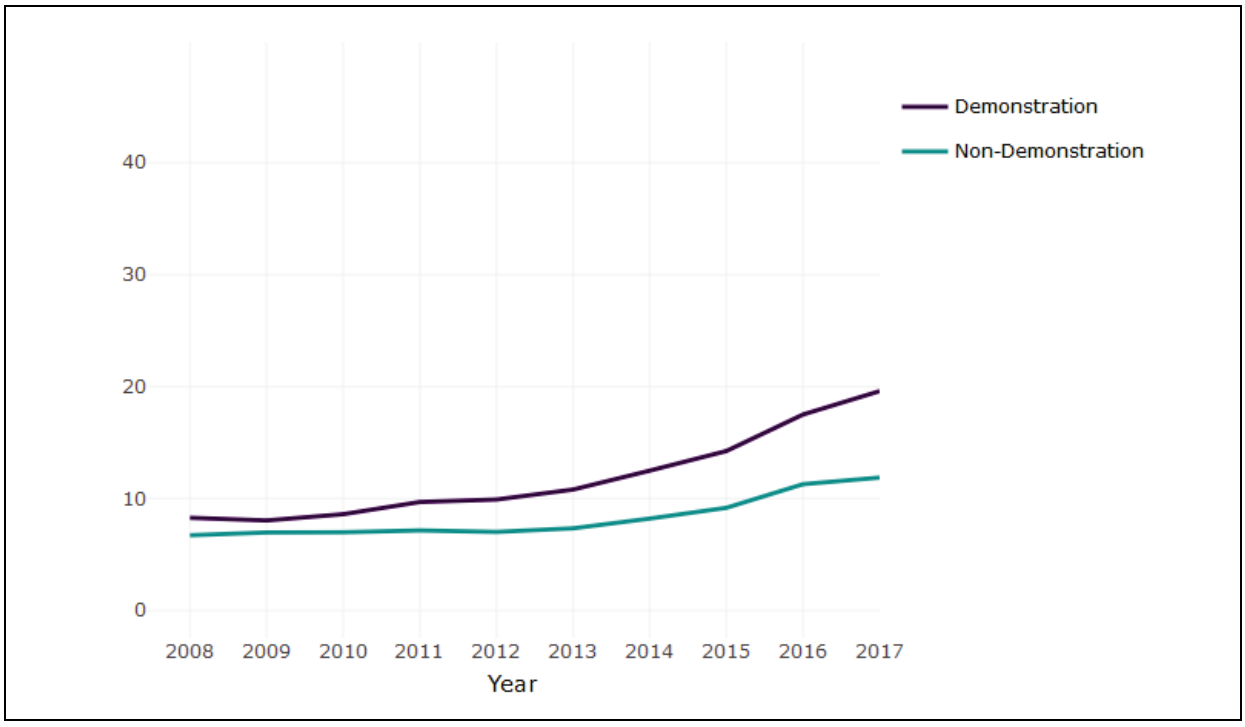
Turning to Milestone #4, in ***Figure 3-8*** we show the rates of facilities that accept Medicaid per 100,000 from 2008-2017 using the N-SSATs. The trend lines for demonstration and non-demonstration states are nearly identical leading up to 2016 when the demonstrations start and statistical tests do not provide evidence of a pre-demonstration difference in trends. Whereas the slope in ***Figure 3-7*** for SUD admissions was flat, the slope for facilities accepting Medicaid is increasing over time for both non-demonstration and demonstration states. This is not concerning from a design perspective but important context for anticipating what changes in SUD treatment capacity may look like and informing the most appropriate statistical model.

Figure 3-8. Facilities Accepting Medicaid per 100,000 Population in N-SSATS, 2008–2017



In ***Figure 3-9*** we show opioid overdose mortality from 2008 through 2017 using mortality records from CDC WONDER. Unlike the prior two figures, there is clear visual evidence that the trend line for demonstration states diverges from the non-demonstration states around 2010 and statistical tests confirm difference in the slope of trend lines are statistically significant. This signals a selection bias—states with higher opioid overdose mortality were more likely to apply for the SUD demonstration—and suggests at-face value that non-demonstration states are a less than ideal comparison group for this outcome. For overdoses, we may need to instead rely on the variation in timing of demonstration implementation among demonstration states. Future analyses will integrate additional control variables to confirm whether the trends are still different once accounting for other factors at the state level.

Figure 3-9. Opioid-Related Overdose Deaths per 100,000 Population in CDC WONDER, 2008–2017



3.8.2 Comparison Group Identification

State evaluations are strongly encouraged to include a comparison group to assess demonstration impacts on outcomes. State evaluations with more rigorous comparison strategies will provide better inputs for the federal meta-analysis. CMS has provided guidance to states on comparison group selection but in some cases states and their evaluators may require more individualized assistance. If requested to do so by the state and CMS, we will provide data and

methods-related expertise to facilitate comparison group selection. We have also developed a tool called the Health Evaluation Engine that can support state efforts to identify an appropriate comparison group if in-state comparisons are sufficient (***Appendix F***).

We expect that a within-state comparison group will not be appropriate for most demonstrations and the comparison group will be drawn from other states. One methodology for selecting comparison states that are good matches to the demonstration state is based on distance scores, which are calculated using data derived from national data sets.

A distance score is a summary measure of the difference, or "distance," between the characteristics of a demonstration state and a potential comparison state. Distance scores essentially quantify the distance between the demonstration and comparison options along the variables chosen. They make it possible to select the comparison with the nearest match on important characteristics. The smaller the distance score, the more similar the two states are. After rank-ordering potential comparisons by their distance scores, evaluators can then select one or more closely matched comparisons for analysis. Distance scores can be used to suggest potential comparison states prior to selecting comparison areas or to corroborate choices that have already been made.

A wide range of state-level characteristics can be used to compute distance scores, but the emphasis is usually on factors associated with demonstration outcomes such as Medicaid expenditures or utilization rates. Demonstration-related features and geographic proximity can also be included.

The scoring process begins by subtracting the target state's value for a characteristic from the comparison state's value and dividing this difference by a population standard deviation (SD) to create a standardized difference. For example, if mean Medicaid expenditures are $7,000 per year in the comparison state, $6,000 per year in the target state and the SD is $2,500, then the standardized difference is (7000–6000)/2500 = 0.40. Euclidean distance scores are then computed by squaring each standardized difference and summing across all characteristics. The Euclidean approach gives equal weight to each state characteristic.

Table 3-3 shows a hypothetical example of distance scores for a particular demonstration state with potential comparison states (A-E) rank-ordered by the size of their distance scores. In this example, the closest matches tend to be comparisons that are also Medicaid expansion states in the same region with similar Medicaid spending levels.

Table 3-3. Hypothetical Example of Distance Scores

State	Distance Score	Medicaid Expansion	Southeast Region	Medicaid Spending per Beneficiary	% Population Covered by Medicaid	% Change in Baseline Employment
Target	0	1	1	6,000	20	+20
A	0.78	1	1	6,200	19	+17
B	1.12	1	1	5,900	23	+22
C	2.04	1	0	6,500	17	-23
D	2.49	0	0	5,000	30	+25
E	7.59	0	0	7,000	12	-29

SECTION 4.
REPORTING

Timely, useful reporting of findings and dissemination to key stakeholder audiences is central to our evaluation approach. We will apply data visualization, storyboarding, and design services to the development of report templates, visualizations, and dissemination products so they are accessible to their intended audience. Dissemination plans will be coordinated with CMS and subject to CMS approval. The content of reports will reflect the various available results from the analyses, with subsequent reports building upon earlier findings. We have eliminated the Interim Evaluation Report, which is due in Option Year 2, because we do not expect results from state evaluations to be available in time for that report. Instead, we will increase the number of Rapid Cycle Reports delivered. We have also modified the approach to the Data Collection and Analysis Report. We discuss the plan for each of the reports below. ***Table 4-1*** provides a schedule for reports. This schedule, including the number of reports, their timing, and topics, will be reviewed each year and modified as needed in coordination with CMS.

Table 4-1. Reporting Schedule

Task #	Base Year	Option Year			
		1	2	3	4
3a-Data Collection and Analysis Memos					
T-MSIS Analytic Files	N/A	Memo due	Update due	Update due	Update due
Other national data sets	Memo due	As needed, memos on additional data sets or updates on previous memos	As needed, memos on additional data sets or updates on previous memos	As needed, memos on additional data sets or updates on previous memos	As needed, memos on additional data sets or updates on previous memos
State evaluation data	N/A	Memo due Updated memo every six months if new information is available	Updated memo every six months if new information is available	Updated memo every six months if new information is available	Updated memo every six months if new information is available
Primary data collection	N/A	Discussion guides due	Discussion guides due	Discussion guides due	Discussion guides due
4a- Rapid Cycle Reports (RCRs)					
Interim performance reports	N/A	1 RCR due*	1 RCR due	1 RCR due	N/A
Case study reports	N/A	4 RCRs due**	2 RCRs due	2 RCRs due	N/A

(continued)

Table 4-1. Reporting Schedule (continued)

Task #	Base Year	Option Year			
		1	2	3	4
4b- Brief on Rapid Cycle Reports	N/A	One for each RCR	One for each RCR	One for each RCR	N/A
5a-Draft Interim Evaluation Report	N/A	N/A	N/A	N/A	N/A
5b- Final Interim Evaluation Report	N/A	N/A	N/A	N/A	N/A
5c- Briefing on Interim Evaluation Report	N/A	N/A	N/A	N/A	N/A
6a-Draft Summative Evaluation Report	N/A	N/A	N/A	N/A	Draft Summative Evaluation Report due
6b Final Summative Evaluation Report	N/A	N/A	N/A	N/A	30 days after receipt of CMS comments on draft Summative Evaluation Report
6c-Briefing on Summative Evaluation Report	N/A	N/A	N/A	N/A	Briefing due

Note: The reporting schedule will be reviewed with CMS and updated as needed at least annually.

*Submission of report during Option Year 1 is contingent on states reporting monitoring metrics by June 2020.

*Three of the case study reports planned for Option Year 1 will be submitted in Option Year 2 because of delays in primary data collection due to PRA clearance and COVID-19.

4.1 Data Collection and Analysis Report

We will produce a series of memos that address different data sources that will be used for the meta-analysis, rather than preparing a single report. As shown above in ***Table 4-1*** and discussed below, memos will be submitted throughout the contract period as new data sources become available or data sources are updated. This approach will provide CMS with more timely, digestible information that is targeted to the issues for specific data sources and the timeline for data availability and decision making about data issues. We will produce memos on four types of data, discussed below.

4.1.1 T-MSIS Analytic Files

We will assess whether the quality and completeness of TAFs is adequate to support evaluation of SUD demonstration outcomes. The types of analyses that will be used to assess the data are described in Section 2.2.2. Refreshed TAFs through 2017 that addressed some of the

problems identified in the initial TAF release became available to researchers from July to September 2019 and data files for 2018 were released for the first time in April 2020.

RTI will submit an initial memo that assesses the usability of TAF data for the SUD demonstration meta-evaluation in June 2020. This memo will identify problems with key data elements, such as missingness and reasonableness of reported values. We will submit a second memo that assesses the validity of SUD outcome measures created from TAFs. These analyses will replicate measures created for the T-MSIS SUD data book.[c] As of May 2020 RTI was waiting to receive details on the measure specifications used for the data book. The submission date for this memo is dependent on when these specifications are received. Based on the analysis presented in these memos, RTI will make a preliminary assessment of the outcomes that can be constructed using TAFs. SUD outcomes for which TAFs are a potential data source are shown in ***Tables 2-6*** and ***2-7***. We anticipate that data quality will vary across states and over time so all data quality analyses will be reported by state and year.

We will update the memo at least annually as new years of T-MSIS data become available. The timing of the updated memo will depend on the TAF release date. If CMS refreshes previously released years of data, we will update our assessment of those years. Updated memos will include results for the most current data release and will highlight any differences in findings from the assessment of previous memos.

4.1.2 Other National Data Sets

We have already submitted a memo that assessed the value of several national data sets for the SUD demonstration meta-evaluation (see ***Appendix E***). If additional national data sets are proposed, we will prepare comparable memos that address how they could be used, their strengths and weaknesses, and recommendations for whether a data set should be used in the analyses. We will also provide updated memos on data sets that have already been considered if new information becomes available, for example, if our recommendation about how or whether to use the data changes after we begin analyzing it.

4.1.3 State Evaluation Data

Approximately 3 months after we receive approved evaluation plans for half of the currently approved SUD demonstrations, we will prepare a memo that documents the states' evaluation design (e.g., whether a comparison group is used, whether the comparison group is in-

[c] https://www.medicaid.gov/medicaid/data-and-systems/downloads/macbis/sud-data-book.pdf

state or out-of-state, what outcomes will be analyzed, differences in how outcomes are specified). As of April 2020, 10 states have an approved evaluation plan. We will use information in the state evaluation plans to assess the extent to which it will be feasible to use state evaluation results in the meta-evaluation. We will update this memo as more evaluation plans are approved. When state evaluation reports become available, we will update the memo based on the actual analyses undertaken, including our assessment of the usability of state evaluation results for the meta-evaluation. We will provide updates of the state evaluation data memo every 6 months if new evaluation designs are approved or evaluation reports are submitted.

4.1.4 Primary Data Collection

We plan limited primary data collection to gather information on demonstration design and implementation that will supplement information available through demonstration implementation plans, monitoring reports, and other secondary data sources. Discussions with key stakeholders at the state Medicaid agency and the single state agency for substance abuse services in each state with an approved SUD demonstration or section 5052 SPA will be conducted annually. We planned to begin these discussions in Option Year 1, but they may be delayed until Option Year 2 because of the time required for PRA clearance and stakeholders' competing priorities due to the COVID-19 pandemic. We also may conduct discussions with managed care organizations and health systems, as well as provider focus groups, in selected states in Option Years 2 and 3. Before beginning primary data collection in each year, we will submit a draft discussion guide to CMS for review. We will revise the discussion guide in response to comments from CMS. After the discussion guide is finalized, it will be submitted for PRA clearance. A preliminary draft discussion guide for Medicaid and substance abuse services agency staff for the first round of data collection is included as ***Appendix C***.

4.2 Rapid Cycle Reports

The statement of work for the Meta-Analysis Support contract calls for one Rapid Cycle Report in each of Option Years 1, 2, and 3. An Interim Evaluation Report is also specified for Option Year 2. We do not expect state evaluations will be available in time to include meta-analyses of their findings in an Interim Evaluation Report. Therefore, after discussion with CMS, we have eliminated the requirement for an Interim Evaluation Report and instead will submit three Rapid Cycle Reports in Option Years 1, 2, and 3, for a total of nine Rapid Cycle Reports. As agreed with CMS, RTI may increase the number of reports in some years.

Each Rapid Cycle Report will contain an executive summary providing a high-level overview of the content, conclusions, and key takeaways of the report. We will submit a draft of each Rapid Cycle Report and will revise the report in response to comments from CMS. At the request of CMS, we will create a summary slide presentation for each Rapid Cycle Report. This presentation will be used to brief CMS on the findings, methods, conclusions, and implications of each of these reports. The briefings will be conducted in person or virtually, depending on CMS's preference.

We will prepare two types of Rapid Cycle Report: interim performance reports and case study reports.

4.2.1 Interim Performance Reports

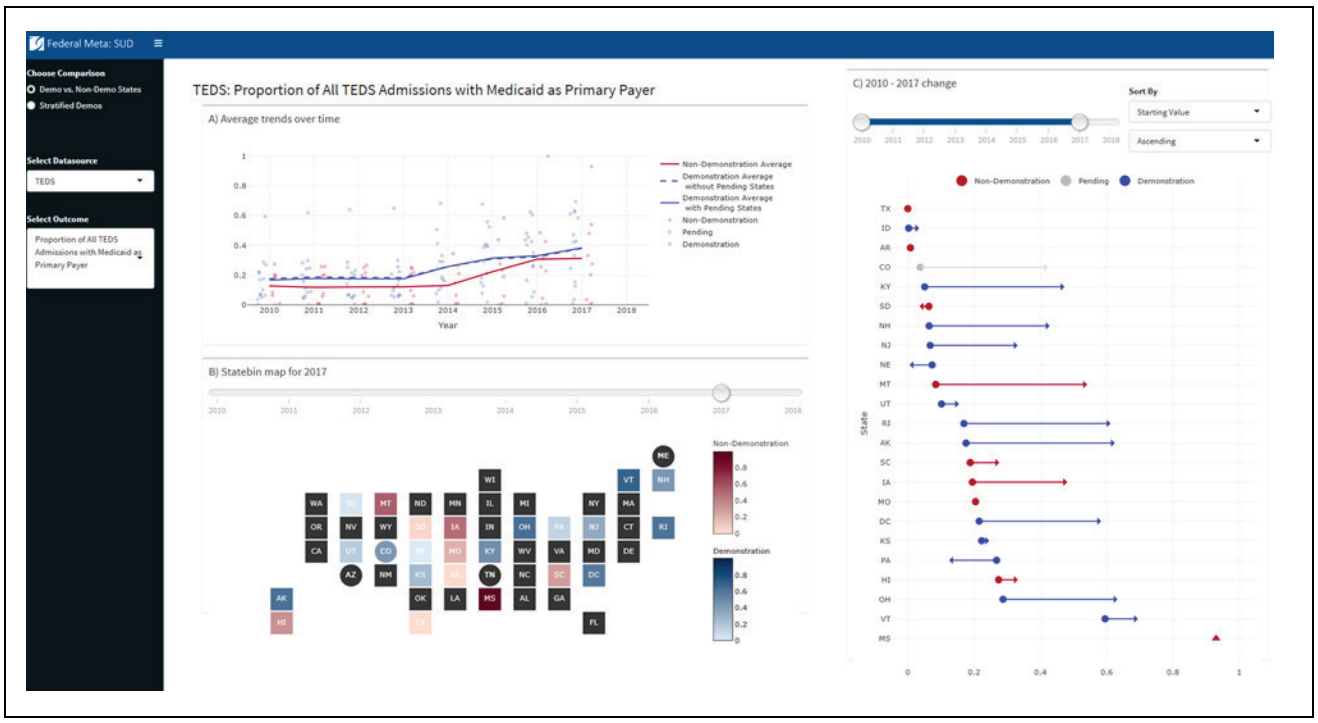
Each year, one Rapid Cycle Report will focus on analysis of quarterly and annual monitoring data. Each report will include monitoring data available as of June of that year. These reports will be based on descriptive comparisons that examine the relationship of cross-state differences in performance on monitoring metrics with state demonstration features and state contextual factors. In addition, the report will include comparisons of any monitoring metrics created with secondary data (see ***Table 2-6*** for potential metrics). The report will also include discussion of narrative findings in quarterly and annual monitoring reports that provide perspective on the monitoring metrics. Monitoring reports submitted to date have not included monitoring metrics. We expected that monitoring data would be available in time for a report at the end of Option Year 1; however, state reporting may be delayed because of the COVID-19 pandemic. We will assess reporting of monitoring metrics in June 2020 and will determine whether this report is feasible in Option Year 1. If necessary, we will delay this to Option Year 2.

Each annual Rapid Cycle Report on interim performance will be accompanied by a data dashboard. The data dashboard is an interactive tool that allows the user to access descriptive trends and maps for select outcomes. We will modify the dashboard that RTI has begun constructing using outcomes from several national datasets (TEDS, N-SSATS, and WONDER) to accommodate monitoring metrics. ***Figure 4-1*** is a screenshot of the dashboard in its current stage of development. Unlike the national data sets, monitoring metrics from quarterly reports are available for demonstration states only. Instead of comparing outcomes for demonstration and non-demonstration states, the dashboard will instead compare monitoring metrics for states by SUD demonstration feature.

As shown in ***Figure 4-1***, the current dashboard is arranged into quadrants, each with its own function. The far-left quadrant is the main control menu for variable selection and

influences what is shown in the other three quadrants. ***Figure 4-1A*** in the upper middle quadrant displays averaged time trends for demonstration versus non-demonstration states where states with a pending demonstration are included with demonstration states. The lower left quadrant (***Figure 4-1B***) displays cross-sectional maps comparing demonstration to non-demonstration states in a given year. In the left most map for demonstration states, the gray states indicate non-demonstration states or demonstration states that are missing values for that particular year and/or outcome. In the non-demonstration states map of ***Figure 4-1B***, non-demonstration states are shaded in red and states missing a value are black. ***Figure 4-1C*** in the far right quadrant displays maps that compare the change between any two years, selected by the second slider in the control menu, in demonstration, non-demonstration states, and pending states.

Figure 4-1. Example Data Dashboard



4.2.2 Case Studies

We will prepare at least two Rapid Cycle Reports each year in Option Years 1, 2, and 3 that will be case studies designed to take a deep dive into targeted demonstration design and implementation topics. Reports will draw on information available in the PMDA, primary data collected from key stakeholders, and other secondary data analyses as appropriate.

Report topics will be selected in consultation with CMS and will focus on emerging demonstration design and implementation issues. For example, report topics may be driven by narrative information in monitoring reports, findings from analyses of monitoring metrics, or information from stakeholder discussions. During the first month of each option year, RTI and CMS will discuss potential case study topics for that year. Based on these discussions, RTI will propose a set of case study topics for that year and a schedule for submitting each report. Report schedules will depend on when data required for the report is expected to be available. RTI and CMS will review the planned report topics periodically during the year in the event different priorities emerge. RTI and CMS agreed on four Rapid Cycle Reports for Option Year 1. These topics are described below. Potential topics for Rapid Cycle Reports in subsequent years might include strategies for maintaining engagement of people in SUD treatment and the role of social service, mental health, and criminal justice agencies in SUD demonstrations.

Three of the four reports identified for Option Year 1 rely on primary data collected through discussions with key stakeholders at the state Medicaid agency and the single state agency for substance abuse services. As discussed in ***Section 4.1.4***, we expect that these discussions will be delayed as result of the time required for PRA clearance and COVID-19. Therefore, RTI will submit one Rapid Cycle Report in Option Year 1 and the three remaining reports will be submitted in Option Year 2.

For Option Year 1, RTI and CMS agreed on the following report topics:

- **State baseline context for SUD demonstrations.** This report will summarize pre-demonstration population characteristics, SUD system characteristics, and select SUD-related service use outcomes among demonstration and non-demonstration states using national data sets. Baseline analyses provide important information for designing meta-analyses of demonstration implementation and impacts and for interpreting findings from these analyses. The draft of this report will be submitted to CMS in June 2020.

  The baseline context RCR will use national data sets to answer three questions:

  1. What was the prevalence of OUD and other SUDs and opioid-related overdoses in states prior to the demonstration and how did they differ from national averages and non-demonstration states?
  2. What SUD treatment capacity (e.g., number of treatment facilities or MAT providers) did demonstration states have prior to the demonstration and how did they differ from national averages and non-demonstration states?
  3. How did individuals with OUD and other SUDs in demonstration states use specialty treatment, EDs, and inpatient services prior to the demonstration and

how did their patterns of use differ from national averages and those in non-demonstration states?

- **Early implementation challenges.** The extent to which challenges delayed implementation timelines or inhibited progress toward demonstration milestones represent secondary factors that could contribute to differences in demonstration outcomes across states, beyond the major demonstration features that are central to the meta-analysis. Stakeholder assessments of factors contributing to these challenges provide information on the degree to which states' pre-demonstration context, such as geographic and other demographic characteristics, contributed to the emergence of challenges, or alternatively, facilitated progress toward demonstration milestones. Based on our understanding of pre-demonstration context, we expect that some common challenges will relate to workforce capacity and data sharing. This report relies on information from key stakeholder interviews. A draft of this report will be submitted to CMS in Option Year 2, 2 months after the stakeholder interviews are completed.

  This report will examine major implementation challenges across SUD section 1115 demonstrations focusing on the following research questions:

  1. For which operational, policy areas, and demonstration milestones are states facing the greatest implementation challenges?
  2. What factors are contributing to these challenges?
  3. How are states addressing these challenges?
  4. What is the potential impact for achieving demonstration goals and objectives?

- **SUD demonstration features.** Although all demonstration states must address a common set of six milestones, they will vary in their strategies to meet each milestone and may introduce other approaches not required for the demonstration, such as changes in reimbursement. The meta-evaluation of the SUD demonstrations will examine whether differences in demonstration features impact an individual demonstration's outcomes and explain variation in outcomes observed across demonstrations. It is important to understand how implemented strategies differ across states and across time (e.g., dates of implementation) to allow for a careful meta-analysis of the effect of the SUD waivers.

  This report will use data from demonstration applications and STCs, quarterly monitoring reports, and other demonstration-related documents to provide an overview of the key demonstration features across states. We will hold discussions with Medicaid and substance abuse services agency officials in demonstration states to augment the information drawn from these data sources, to clarify information in these documents and to fill in information that was not available. This report will be submitted 6 weeks after these discussions are completed. Information in demonstration documents is reasonably complete for some demonstration features. If CMS would like, RTI can submit a preliminary memo describing findings on features based only on the review of documents in Option Year 1.

The report will answer the following questions:

1. How did the coverage of SUD services change after the demonstration started?
2. How did the use of patient placement criteria and utilization management change after the demonstration started?
3. How did the residential provider standards change after the demonstration started?
4. How did the use of care coordination strategies and integrated physical health services for individuals with SUD change after the demonstration started?
5. How did reimbursement change after the start of the demonstration?

- **Impact on access to MAT.** As a part of meeting requirements for Milestone #3 (Use of nationally recognized, evidence-based SUD program standards to set residential treatment provider qualifications), states must ensure that within 24 months at least two FDA-approved medications for OUD are available on-site in residential treatment settings or through a direct referral. This report will describe states' activities and approaches to expand residential MAT capacity. Currently, states are pursuing several policy options to expand access to MAT across all medical settings including enhancing coverage for MAT services and medications, increasing provider reimbursement rates, and revising provider licensing or certification requirements to increase the number of certified MAT providers statewide. States are also updating or developing new MAT regulations for residential providers. To increase the number of sites providing MAT, states are also organizing training sessions to help providers comply with federal regulations pertaining to MAT prescribing, dispensing, and oversight. The monitoring metrics for the section 1115 SUD demonstrations do not include residential MAT access and thus information in this report is critical for CMS in understanding the implementation and impact of Milestone #3.

  The Rapid Cycle Report will incorporate data from national data sets and key stakeholder interviews. A draft of this report will be submitted to CMS in Option Year 2, 2 months after the stakeholder interviews are completed. The report will answer the following questions:

  1. What MAT capacity did states have in residential settings before their demonstration started?
  2. How many individuals entering residential SUD treatment planned to have MAT before their demonstration started?
  3. What steps have states taken to increase and ensure access to MAT services in residential settings and after discharge?
  4. What barriers and challenges have states experienced in increasing access to residential MAT services?

4.3 Summative Evaluation Report

In Option Year 4, we will submit a Summative Evaluation Report for the SUD demonstrations. The draft Summative Evaluation Report will be submitted in month 5 of Option Year 4 (February 2023). The report will summarize the demonstrations' accomplishments,

challenges, lessons learned (both negative and positive), findings and conclusions, and recommendations where applicable. The report will also include methods and technical details.

We will first prepare an outline containing section headings, table shells, and proposed figures for CMS review and comment. We will incorporate feedback on the report outline and submit a revised outline within 2 weeks after receiving consolidated comments from CMS. The Summative Evaluation Report will present results and data visualizations. The report will describe the findings from meta-analyses of demonstration impacts based analyses of state evaluation results and supplementary analyses of national data. The report will also discuss implementation experience across states and will synthesize this information with impact findings to further interpret demonstration outcomes. Information from Rapid Cycle Reports will be incorporated into the Summative Evaluation Report as appropriate. The report will draw on the full range of data sources proposed for the meta-evaluation, including primary and secondary data.

A critical component of dissemination is to make the results practical and valuable for stakeholders. To accomplish this, we will describe methods in nontechnical language in the body of the report and provide comprehensive technical detail in appendices. We will present key findings and takeaways at the start of chapters and sections and then provide and explain each section's results. We will close each section with a discussion of findings. The Summative Evaluation Report will include an executive summary and a closing chapter that integrates results across topics. Whenever possible, information will be presented using visual displays, infographics, or simplified tables. Where appropriate or necessary, tables or text supporting these simplified displays will be presented in appendices. To ensure findings are used in a practical and valuable way, the Summative Evaluation Report will be accompanied by a concise stand-alone document that presents key findings.

To allow the team to review interim findings, prioritize results to be reported, and develop a cohesive story for the report, we will convene a virtual 1-day meeting for report authors. By sharing content, comments, and conclusions, this meeting will enhance our ability to provide CMS an appropriate structure, design, and focus for the report.

We will submit a draft Summative Evaluation Report to CMS for review and comment. At CMS's discretion, states and state evaluators may also review the reports. Within 30 days of receipt of CMS and other stakeholder comments, our team will provide a final Summative Evaluation Report. Upon approval of the final Summative Evaluation Report, we will develop a summary slide presentation of key findings for presentation to CMS. At CMS's request we will

also develop and participate in up to three webinars to disseminate the findings. These presentations might be made at learning collaborative meetings with states and state evaluators. CMS may also request that presentations be made to CMS and collaborating stakeholders like the Center for Medicare & Medicaid Innovation and the Office of the Assistant Secretary for Planning and Evaluation.

REFERENCES

1. Centers for Disease Control and Prevention (CDC). (2011). Vital signs: Overdoses of prescription opioid pain relievers—United States, 1999–2008. *MMWR, 60*(43), 1487–1492.

2. Centers for Disease Control and Prevention (CDC). (2017). Opioid data analysis: Categories of opioids. Retrieved from https://www.cdc.gov/drugoverdose/data/analysis.html

3. Substance Abuse and Mental Health Administration. (2016). *Key substance use and mental health indicators in the United States: Results from the 2015 National Survey on Drug Use and Health (NSDUH)*. Rockville, MD: Substance Abuse and Mental Health Administration.

4. Department of Health and Human Services, National Institute on Drug Abuse. (2017). Expanding medication assisted treatment for opioid use disorders in the context of the SAMHSA Opioid STR grants (R21/R33). Retrieved from https://grants.nih.gov/grants/guide/rfa-files/RFA-DA-18-005.html

5. Jones, C. M., Campopiano, M., Baldwin, G., & McCance-Katz, E. (2015). National and state treatment need and capacity for opioid agonist medication-assisted treatment. *American Journal of Public Health, 105*(8), e55–e63. https://doi.org/10.2105/AJPH.2015.302664

6. Centers for Medicare and Medicaid Services. (2015, July). SMD # 15-003: New service delivery opportunities for individuals with a substance use disorder. Retrieved from https://www.medicaid.gov/federal-policy-guidance/downloads/smd15003.pdf

7. Centers for Medicare and Medicaid Services. (2017, November 1). SMD # 17-003: Strategies to address the opioid epidemic. Retrieved from https://www.medicaid.gov/federal-policy-guidance/downloads/smd17003.pdf

8. Bazeley, P., & Richards, L. (2000). *The NVivo Qualitative Project Book*. London: Sage.

9. Richards, L. (2009). *Using NVIVO in Qualitative Research*. London: Sage.

10. Sorensen, A. V. (2008). Media review: Use of QSR NVivo 7 qualitative analysis software for mixed methods research. *Journal of Mixed Methods Research, 2*(1).

11. Newman, I., Onwuegbuzie, A. J., & Hitchcock, J. H. (2015). Using the general linear model to facilitate the full integration of qualitative and quantitative analysis: The potential to improve prediction and theory building and testing. *General Linear Model Journal, 41*(1), 12–28.

12. Kahwati, L. C., & Kane, H. L. (2020). *Qualitative comparative analysis in mixed methods research and evaluation*. Thousand Oaks, CA: Sage.

13. Schneider, C. Q., & Wagemann, C. (2012). *Set-Theoretic Methods for the Social Sciences: A Guide to Qualitative Comparative Analysis*. Cambridge, MA: Cambridge University Press.

14. Thygeson, N. M., Peikes, D., & Zutshi, A. (2013). *Mixed methods: Fuzzy set qualitative comparative analysis and configurational comparative methods: Powerful methods to study and refine patient-centered medical home models*. Rockville, MD: Agency for Healthcare Research and Quality.

15. Goodrick, D. (2014). Comparative case studies: Methodological briefs—impact evaluation No. 9 (No. innpub754). Florence, Italy: UNICEF Office of Research—Innocenti.

16. Dinour, L. M., Kwan, A., & Freudenberg, N. (2017). Use of comparative case study methodology for US public health policy analysis: a review. *Journal of Public Health Management and Practice, 23*(1), 81–89.

APPENDIX A:
SUD ABSTRACTION EXAMPLE

State	Indiana	Kansas
Program Basics		
Demonstration name	Healthy Indiana Plan (HIP 2.0)	KanCare
Other major changes/ initiatives to code	Added OTP/Methadone just prior to demonstration	Pending Medicaid expansion
Date waiver submitted	1/31/2017	12/26/2017
Date waiver approved	2/1/2018	1/18/2018
Demonstration start date	2/1/2018	1/1/2019
Demonstration end date	12/31/2020	12/31/2023
Length of approval period	2 years, 10 months	5 years
Stand-alone SUD waiver	No	No
Medicaid expansion	Yes	No
Effective date of Medicaid expansion	2/1/2015	N/A
Pre-Waiver SUD Services and Characteristics		
Early Intervention Services (ASAM Level 0.5)	Yes	Unclear—SBIRT is covered
Outpatient Services (ASAM Level 1.0)	Yes—covered in OBOTS, and in CMHCs through the MRO	Yes
Intensive Outpatient Services (ASAM Level 2.1)	Yes—only in MRO	Yes
Partial Hospitalization (ASAM Level 2.5)	Yes	Yes
Clinically managed low-intensity residential services (ASAM Level 3.1)	No	Yes
Clinically managed population specific high-intensity residential services (ASAM Level 3.3)	Unclear—this service is not referenced in state documents	Yes
Clinically managed high-intensity residential services (ASAM Level 3.5)	No	Unclear—residential services are covered in non-IMDS but the state does not define explicit levels of care in the implementation plan. The STCs suggest this level of care is covered

State	Indiana	Kansas
Medically monitored intensive inpatient services (ASAM Level 3.7)	Yes—in lieu of	Yes
Medically managed intensive inpatient services (ASAM Level 4.0)	Yes—in lieu of	Yes
WM: Ambulatory withdrawal management *without* extended on-site monitoring (ASAM 1.0)	Unclear—the state does not differentiate between ASAM levels	No
WM: Ambulatory withdrawal management *with* extended on-site monitoring (ASAM 2.0)	Unclear—the state does not differentiate between ASAM levels	No
WM: Clinically managed withdrawal management (ASAM 3.2)	Unclear—the state does not differentiate between ASAM levels	No
WM: Medically monitored withdrawal management (ASAM 3.7)	Unclear—the state does not differentiate between ASAM levels	No
WM: Inpatient detoxification (ASAM 4.0)	Unclear—the state does not differentiate between ASAM levels	No
Covers methadone?	Yes—but the change was very recent (August 2017)	No
Changes in methadone prior authorization?	No	No
Changes in buprenorphine coverage?	No	No
Changes in buprenorphine prior authorization?	Yes—for non-Gold card providers there is a prior authorization component	No
Changes in naltrexone coverage?	No	No
Changes in naltrexone prior authorization?	No	No
MAT-related counseling required	Yes—must document referral or active counseling for buprenorphine	No
Uses widely recognized placement criteria	No	Yes—KanCare
Use of national provider standards (e.g., ASAM)	Unclear	Yes

State	Indiana	Kansas
Changes in Peer Support Worker/recovery coach coverage?	Covered only through MRO for mental health treatment	Yes
Changes in SUD case management coverage?	Covered only through MRO for mental health treatment	Yes
Changes in recovery housing/supportive housing coverage?	Unclear	No
Changes in supported employment coverage?	Unclear	No
Changes in mutual aid groups (e.g., 12 step programs)?	Unclear	No
SUD utilization review/benefit management	Yes	Yes
SUD carve-in	No	No
Post-Waiver SUD Services and Characteristics		
Early Intervention Services (ASAM Level 0.5)	No change	No change
Outpatient Services (ASAM Level 1.0)	No change	No change
Intensive Outpatient Services (ASAM Level 2.1)	Yes—state removing IOP from the MRO and allowing MCOs to bill for it	No change
Partial Hospitalization (ASAM Level 2.5)	No change	No change
Clinically managed low-intensity residential services (ASAM Level 3.1)	Yes—will be covered for both MCO and FFS beneficiaries and in both IMD and non-IMD settings	Yes—IMDs
Clinically managed population specific high-intensity residential services (ASAM Level 3.3)	No—this service is not referenced throughout the documents we have	Yes—IMDs
Clinically managed high-intensity residential services (ASAM Level 3.5)	Yes—will be covered for both MCO and FFS beneficiaries. and in both IMD and non-IMD settings	Yes—IMDs
Medically monitored intensive inpatient services (ASAM Level 3.7)	Yes—will cover FFS now as well and assessing per diem payments, must also align to 6 ASAM dimensions	Yes—IMDs

State	Indiana	Kansas
Medically managed intensive inpatient services (ASAM Level 4.0)	Yes—will cover FFS now as well and assessing per diem payments, must also align to 6 ASAM dimensions	Yes—IMDs
WM: Ambulatory withdrawal management *without* extended on-site monitoring (ASAM 1.0)	Unclear—the state does not differentiate between levels of care prior to the demonstration	Unclear—the state services will now be covered in IMDs in the STCs but did not indicate that WM services will now be covered in IMDs in the implementation plan
WM: Ambulatory withdrawal management *with* extended on-site monitoring (ASAM 2.0)	Unclear—the state does not differentiate between levels of care prior to the demonstration	No changes
WM: Clinically managed withdrawal management (ASAM 3.2)	Unclear—the state does not differentiate between levels of care prior to the demonstration	No changes
WM: Medically monitored withdrawal management (ASAM 3.7)	Unclear—the state does not differentiate between levels of care prior to the demonstration	No changes
WM: Inpatient detoxification (ASAM 4.0)	Unclear—the state does not differentiate between levels of care prior to the demonstration	No changes
Covers methadone?	Yes	Yes—planned*
Changes in methadone prior authorization?	No	Yes—planned*
Changes in buprenorphine coverage?	No change	No change
Changes in buprenorphine prior authorization?	No change	No change
Changes in naltrexone coverage?	No change	No change
Changes in naltrexone prior authorization?	No change	No change
MAT-related counseling required	Yes—must document referral or active counseling for buprenorphine.	No change
Changes in use of widely recognized patient placement criteria	Yes—ASAM is the guidepost although MCEs are not required to use it explicitly.	No change

State	Indiana	Kansas
	Some leeway to use other placement criteria so long as they cover the ASAM dimensions.	
Change in use of provider standards	Yes—creating new standards	Yes—ensuring standards are aligned to ASAM
Changes in Peer Support Worker/recovery coach coverage?	No change	No change
Changes in SUD case management coverage?	Yes—planning a SPA and working to make available in OTPs	Yes—demonstration is expanding this benefit
Changes in recovery housing/supportive housing coverage?	Yes—supportive housing for individuals transitioning to and trying to sustain housing	Yes—newly mentioned
Changes in supported employment coverage?	Unclear	Yes—demonstration is expanding this benefit.
Changes in mutual aid groups (e.g., 12 step programs)?	Unclear	Unclear
SUD utilization review/benefit management	Yes, will require assessments to align with 6 ASAM dimensions. Will be updating CANS/ANSA assessments to match Will also implement a standardized prior authorization form for services above Level 2.0	No changes
SUD carve-in	No	No

*IMD = Institution for Mental Disease; SUD = Substance Use Disorder; MAT = Medication-Assisted Treatment; ASAM = American Society of Addiction; MRO = Medicaid Rehabilitation Option; MCE = Managed Care Entities; MCO – Managed Care Organization; OTPs = Opioid Treatment Program; IOP = Intensive Outpatient Program; N/A = Not Applicable; STCs = Special Terms and Conditions; FFS = Fee For Service; CANS = Child and Adolescent Needs and Strengths; ANSA = Adult Needs and Strength Assessment; SPA = State Plan Amendment.

APPENDIX B:
EXAMPLE GRID FOR CONFIRMING SUD DEMONSTRATION POLICY CHANGES

Table 1. SUD services pre- and post-waiver for [State X]

SUD services pre- and post-demonstration for State X

The following table has been populated by the RTI Meta-Evaluation State Team based on your state's 1115 SUD Waiver special terms and conditions [INSERT HYPERLINK] and SUD demonstration Implementation Plan [INSERT HYPERLINK]. Additional sources, if used, are cited.

Instructions: *Please review the accuracy of information in this grid. Use red font to make corrections. Where "needs clarification" appears in the cell, the RTI Team explicitly requests clarifying details. We include a clarifying question with the corresponding number (#) below the table. Where policy or benefit changes took place, please enter any known effective dates of the change in the far-right column. Where an exact date is not known, specify if the change occurred before or after the start date of your state's SUD demonstration.*

SUD Services and SUD Provider Requirements	Pre-waiver Status	Changes Made as Part of Section 1115 SUD Demonstration	
	Implemented or Covered prior to demonstration (yes/no, areas of clarification)	*Added or Updated as a part of the demonstration*	*Effective date of change*
Medication Assisted Treatments			
Methadone for OUD	Yes	No	
Buprenorphine	Yes	No	
Oral naltrexone	Yes	No	
Long-acting injectable naltrexone	Yes	No	
SUD Treatment Services Covered by Medicaid State Plan or State-only Funds			
LOC 0.5: Early intervention services for SUD	Yes	Yes—new reimbursement methodology for CPEP and the CRT	
LOC 1.0: Outpatient	Yes	No changes	
LOC 2.1: Intensive outpatient	Yes	No changes	
LOC 2.5: Partial hospitalization	**Needs clarification (1)**	**Needs clarification (1)**	

SUD Services and SUD Provider Requirements	Pre-waiver Status	Changes Made as Part of Section 1115 SUD Demonstration	
	Implemented or Covered prior to demonstration (yes/no, areas of clarification)	*Added or Updated as a part of the demonstration*	*Effective date of change*
Any residential SUD treatment, LOC unspecified	Yes—covered under Medicaid through in-lieu of provision. Residential also covered through local funds	Yes	
LOC 3.1: Low-intensity residential	No—Covered by local-only funds	Yes	
LOC 3.3: High-intensity, population-specific residential	**Needs clarification (2)**	Yes	
LOC 3.5: High-intensity residential	No—Covered by local-only funds	Yes	
LOC 3.7: Medically *monitored* intensive inpatient	**Needs clarification (3)**	Yes	
LOC 4.0 Medically *managed* intensive inpatient	**Needs clarification (4)**	Yes	
Withdrawal Management (WM), LOC unspecified	Yes—covered under Medicaid through in-lieu of provision. Residential also covered through local funds	Yes	
LOC 1.0 -WM: Ambulatory without extended on-site monitoring	Yes	No	
LOC 2.0 -WM: Ambulatory with extended on-site monitoring	**Needs clarification (5)**	Yes (IMDs)	
LOC 3.2 -WM: Clinically managed	**Needs clarification (5)**	Yes (IMDs)	
LOC 3.7 -WM: Medically monitored	No	Yes (IMDs)	
LOC 4.0 -WM: Inpatient detoxification	**Needs clarification (5)**	Yes (IMDs)	
Recovery support services			
Peer support services	Yes—DBH-supported Peer-Operated Centers not covered by Medicaid	Yes	
SUD case management	Yes—through MCOs, State [X] Health Home, Other Health Home Program [X]	No changes	
Recovery housing/supportive housing coverage	No	Yes	
Supported employment coverage	No	Yes	
Patient Placement Criteria			
Widespread use of evidence-based patient placement criteria	Yes-TAP with ASAM	No—will decentralize intake, assessment, and referral system	
Use of utilization review and benefits management for SUD treatment	Yes-through QIO and MCOs	No changes	
Program Standards for Residential Treatment Providers			

SUD Services and SUD Provider Requirements	Pre-waiver Status	Changes Made as Part of Section 1115 SUD Demonstration	
	Implemented or Covered prior to demonstration (yes/no, areas of clarification)	*Added or Updated as a part of the demonstration*	*Effective date of change*
Use of widely recognized, evidence-based provider standards for SUD residential treatment	Yes—by DBH using ASAM	No changes	
Residential MAT requirements	Yes—onsite or offsite access to MAT required	Yes—updating policies	
Care Coordination: Coverage and Policies			
Policies supporting care coordination	Yes—through DBH, MCOs, State Health Home, Other Health Home Program [X], FQHC APM	No policy changes	
Policies for transitions in care	Yes—through DBH and MCOs	Yes—new requirements on psychiatric hospitals and RTCs; adds Medicaid reimbursement for TPS provided by certain BH providers	
Policies supporting integration of care	Yes—through State Health Home, Other Health Home Program [X]	Yes—new requirements on psychiatric hospitals and RTCs	
State [X] Health Home; Other Health Home Program [X]		These programs are expected to grow	

ASAM=American Society of Addiction Medicine. CTCC=Comprehensive Transitional Care Coordination. DBH=Department of Behavioral Health. MCO=Managed Care Organization. TAP=Treatment Assessment Protocol. TPS=Transition Planning Services. QIP=Quality Improvement Program. WM=Withdrawal Management. LOC=Level of Care

The numbered questions below correspond to the numbered table notes appearing in cells needing clarification. Please provide answers in your preferred format. You may write your answer below the question, or enter your answer into the respective cells:

1. Your Implementation Plan indicates that you will be making some changes to coverage and provider requirements related to intensive day treatment. Could you clarify if both ASAM LOC 2.1 and 2.5 were covered by the Medicaid State Plan (or through in-lieu-of MCO provisions) prior to the waiver, and what changes your state has made or plans to make in that regard?
2. Prior to the waiver, was ASAM LOC 3.3 covered by the State Plan for non-IMD settings? Did the "in-lieu-of" MCO provision apply to this ASAM LOC?
3. Prior to the waiver, was ASAM LOC 3.7 covered by the State Plan for non-IMD settings? Did the "in-lieu-of" MCO provision apply to this ASAM LOC?
4. Prior to the waiver, was ASAM LOC 4.0 covered by the State Plan for non-IMD settings? Did the "in-lieu-of" MCO provision apply to this ASAM LOC?

5. We understand that the District covered withdrawal management services in a wide array of settings except for ASAM LOC 3.7-WM. We would like to clarify if WM services covered every LOC as defined by ASAM, including ambulatory, clinically managed, etc.

Other document(s) cited:

Name/source/description: [INSERT HYPERLINK]
Name/source/description: [INSERT HYPERLINK]

Demonstration Characteristics Interview Questions

Thank you for making time to speak with us today. My name is [NAME] and I am here with [NAME]. We are researchers from RTI International. The Centers for Medicare & Medicaid Services contracted with RTI to conduct a federal meta-analysis of section 1115 substance use disorder demonstrations. Information gathered during this call will support the federal meta-evaluation of section 1115 SUD demonstrations.

The purpose of this call is to clarify and reconcile the information we sent to you via email in the Program Characteristics Grid for [STATE]'s section 1115 SUD demonstration. The grid was populated by the RTI meta-evaluation team after reviewing your state's section 1115 SUD demonstration special terms and conditions, Implementation Plan, Quarterly and Annual Monitoring Reports, and other information posted on your state's demonstration website. RTI submitted this program characteristics grid to you for review earlier, and you and your colleagues have provided comments and corrections in response.

Today we will focus on additional details we need to understand components of your SUD demonstration. We may need details such as the policy vehicle for the change, reimbursement increases, regulatory mandates on providers, or updates to managed care contracts.

Before we get started, I will begin by reading the PRA Disclosure Statement.

PRA Disclosure Statement

According to the Paperwork Reduction Act of 1995, no persons are required to respond to a collection of information unless it displays a valid OMB control number. The valid OMB control number for this information collection is 0938-1148 (CMS-10398 # INSERT). The time required to complete this information collection is estimated to approximately 1 hour for two people to participate in this interview, plus an average of ½ hour for two people to complete a grid prior to the interview. This includes the time to review instructions, search existing data resources, gather the data needed, and complete and review the information collection. If you have comments concerning the accuracy of the time estimate(s) or suggestions for improving this form, please write to: CMS, 7500 Security Boulevard, Attn: PRA Reports Clearance Officer, Mail Stop C4-26-05, Baltimore, Maryland 21244-1850."

Your decision to participate in this aspect of the study is voluntary. Under the Privacy Act of 1974 any personally identifying information obtained will be kept private to the extent of the law. If you do not wish to participate in this interview or answer specific questions, please let us know. We believe there are minimal risks to you from participation, and every effort will be made to protect your confidentiality. In reports to CMS, we will refer to you anonymously as a "state informants."

Your insights on the section 1115 SUD demonstrations are important and will be used by federal and state policymakers as well as other Medicaid programs in developing program policies and guidance for current SUD demonstrations and other future section 1115 demonstrations. Please note that your participation in this call is voluntary.

Finally, we would like to record our conversation, to ensure our notes from today are complete. Do I have your permission to audio record our conversation today? Do you have any questions before we begin?

We appreciate the time it has taken for you to conduct your review. This advance work allows us to focus our questions and shorten the length of this call.

Do you have any questions before we begin?

I. Medication Assisted Treatment for OUD

Let's begin with the first section of the grid on Medication Assisted Treatment for OUD. We have the following clarifying questions:

1. Of the four medications for treating OUD included in the grid, which did Medicaid cover before and after the demonstration started?
    a. Methadone for OUD?
    b. Buprenorphine?
    c. Oral naltrexone?
    d. Long-acting injectable naltrexone (or Vivitrol)
2. Of the four medications included in the grid, were there any other Medicaid policy changes made to increase access to MAT (e.g., prior authorization)
3. Could you confirm the effective dates or scheduled timeline for changes in medication coverage under the demonstration?

II. Continuum of SUD Services by Level of Care

The next section reviews SUD services by level of care as reimbursed by the Medicaid state plan prior to the demonstration. We understand that some states did not rely on a nationally recognized level of care continuum prior to the demonstration, so classification of services into such a continuum for the pre-demonstration period is not straightforward. We would appreciate your help determining if services consistent with a given level are available before or after the demonstration started. For purposes of this discussion, we will refer to levels of care based on the American Society of Addiction Medicine (ASAM) criteria that were used in the implementation plan protocol.

Based on this information, we have the following clarifying questions:

Early intervention services for SUD (LOC level 0.5) and outpatient services (LOC 1.0)

4. Did Medicaid cover SBI, SBIRT, or other early intervention services prior to the demonstration?
5. Have any changes been made as part of the SUD demonstration to expand coverage of early intervention services? For instance, were any billing codes added?

6. Have any changes been made as part of the SUD demonstration to expand coverage of outpatient services? For instance, were any billing codes added? What was the effective date of the change?

Intensive Outpatient (LOC 2.1, 2.5)

7. Did Medicaid cover intensive outpatient services and partial hospitalization services prior to the demonstration?

8. Have any changes been made as part of the SUD demonstration to expand coverage of intensive outpatient services and partial hospitalization services? For instance, were any billing codes added? What was the effective date of the change?

Inpatient and Residential Treatment (LOC 3.1-4.0, WM-3.2, WM-3.7, WM-4.0)

9. With respect to inpatient and residential treatment for SUD, which levels of care did Medicaid cover prior to the demonstration?

    a. Low-intensity residential (3.1)?

    b. High-intensity, population-specific residential (3.3)?

    c. High-intensity residential (3.5)?

    d. Medically *monitored* intensive inpatient (3.7)?

    e. Medically managed intensive inpatient (ASAM Level 4.0)?

10. Were any of these levels of care covered by Medicaid through the in-lieu-of provision for managed care plans? Were any of these levels covered for non-IMDs under the state plan?

11. Beyond the waiver of the IMD exclusion rule, have any changes been made as part of the SUD demonstration to expand coverage of inpatient and residential levels of care? What was the effective date of the change?

Withdrawal Management

12. With respect to withdrawal management, which levels of care did Medicaid cover prior to the demonstration?

    a. Ambulatory detoxification without extended on-site monitoring (1.0)?

    b. Ambulatory detoxification with extended on-site monitoring (2.0)?

    c. Clinically managed withdrawal management (WM-3.2)?

    d. Medically monitored withdrawal management (WM-3.7)?

    e. Inpatient detoxification (WM-4.0)?

13. Were any of these levels of care for withdrawal management covered by Medicaid through the in-lieu-of provision for managed care plans? Were any levels were covered for non-IMDs under the state plan?

14. Beyond the waiver of the IMD exclusion rule, have any changes been made as part of the SUD demonstration to expand coverage of withdrawal management? What was the effective date of the change?

III. Recovery Support Services

The next section covers recovery support services.

15. Which recovery support services were covered by Medicaid prior to the demonstration:

    a. Peer support services?

    b. SUD case management?

    c. Recovery housing/supportive housing coverage?

    d. Supported employment coverage?

16. Have any changes been made as part of the SUD demonstration to expand coverage of recovery support services? For instance, were any billing codes added or were services expanded for individuals with a principal diagnosis of SUD? What was the effective date of the change?

IV. Patient Placement Criteria

The next section covers use of widely recognized or evidence-based patient placement criteria.

17. Prior to the demonstration, did [STATE] have in place patient placement criteria derived from a widely recognized or evidence-based source to determine the appropriate setting for SUD services? If so, what was the evidence-based source or sources?

18. Have any changes been made as part of the SUD demonstration towards adopting or updating patient placement criteria? Could you confirm the effective dates or scheduled timeline for major changes or updates to the patient placement criteria under the demonstration?

19. Prior to the demonstration, did [STATE] have in utilization review in place for SUD services?

20. Have any changes been made as part of the SUD demonstration towards adopting or updating utilization review processes? Could you confirm the effective dates or scheduled timeline for major changes or updates to utilization review processes under the demonstration?

V. Program Standards for Residential Treatment Providers

The next section covers program standards for providers of residential treatment of SUD required for participation in the Medicaid program, including licensing and oversight.

21. Prior to the demonstration, did [STATE] have in place widely recognized, evidence-based standards for residential SUD treatment programs? If so, what was the source or sources for these standards?
22. Have any changes been made as part of the SUD demonstration towards adopting or updating standards for residential SUD treatment programs? Could you confirm the effective dates or scheduled timeline for major changes or updates to the residential treatment program standards under the demonstration?
23. Did [STATE] require residential treatment programs to offer Medication Assisted Treatment either on-site or off-site? Since the demonstration began, have new requirements for access to MAT in residential facilities become effective?

VI. Care Coordination and Transitions in Care—Policies and Coverage

The last section covers care coordination coverage and policies, policies around transitions in care, and policies supporting integration with physical health.

24. Prior to the demonstration, did Medicaid cover care coordination and transitions in care services for individuals receiving treatment for SUD?
    a. Did eligibility for these services require a principal diagnosis other than SUD?
    b. Was eligibility for these services limited to individuals with a dual diagnosis?
    c. Was eligibility limited to individuals eligible through [STATE-SPECIFIC PROGRAM NAME]?
    d. Have any changes been made as part of the SUD demonstration towards adopting or updating care coordination and transitions in care services? Could you confirm the effective dates or scheduled timeline for major changes or updates to the care coordination and transitions in care services under the demonstration?
25. Prior to the demonstration, for individuals receiving treatment for SUD, did Medicaid have policies or programs in place to improve access to other services for treatment of comorbid diagnoses, through screening or referral tools, or integration of SUD and mental health services?
26. As part of the demonstration, is the state making changes to improve access to treatment for comorbid diagnoses? Could you confirm the effective dates or scheduled timeline for major changes or updates?

CLOSING

This is all the questions we have. Thank you for taking the time to clarify your state's Medicaid policies. Your input is critical for ensuring a high-quality federal meta-analysis of SUD demonstrations**.** If there is written documentation you think would be helpful for us to have or review that would not be accessible from agency websites, we would gladly accept and review them. We will make corrections to your state's grid of program characteristics based on your input today and send a copy via email to you for your records. **You are not being asked to take any further action for this review.** However, if you have any additional clarifications or corrections you would like to make after this call, you may respond to our email or contact [RTI POINT OF CONTACT NAME] at RTI via email at POINT OF CONTACT EMAIL].

[END OF SCRIPT]

APPENDIX C:
DISCUSSION GUIDE TEMPLATE FOR RAPID CYCLE REPORTING

Protocol Questions with Instructions to Interviewers:

Thank you for making time to speak with us today. My name is [name] and I am here with [NAME]. We are researchers from RTI International, conducting a federal meta-evaluation of section 1115 substance use disorder (SUD) demonstrations. This interview will be approximately 90-minutes and will be an in-depth discussion of implementation experience, challenges, and programmatic changes.

Before we get started, I will begin by reading the PRA Disclosure Statement.

PRA Disclosure Statement

According to the Paperwork Reduction Act of 1995, no persons are required to respond to a collection of information unless it displays a valid OMB control number. The valid OMB control number for this information collection is 0938-1148 (CMS-10398 # INSERT). The time required to complete this information collection is estimated to average 90 minutes to participate in this interview. If you have comments concerning the accuracy of the time estimate(s) or suggestions for improving this form, please write to: CMS, 7500 Security Boulevard, Attn: PRA Reports Clearance Officer, Mail Stop C4-26-05, Baltimore, Maryland 21244-1850."

Your decision to participate in this aspect of the study is voluntary. Under the Privacy Act of 1974 any personally identifying information obtained will be kept private to the extent of the law. If you do not wish to participate in this interview or answer specific questions, please let us know. We believe there are minimal risks to you from participation, and every effort will be made to protect your confidentiality. In reports to CMS, we will refer to you anonymously as a "state informants."

Your insights on the section 1115 SUD demonstrations are important and will be used by federal and state policymakers as well as other Medicaid programs in developing program policies and guidance for current SUD demonstrations and other future section 1115 demonstrations.

Finally, we would like to record our conversation, to ensure our notes from today are complete. Do I have your permission to audio record our conversation today? Do you have any questions before we begin?

RATIONALE FOR SECTION 1115 SUD DEMONSTRATION

1. From your vantage point, what was the motivation for pursuing the SUD section 1115 demonstration?

   Where were the biggest gaps in service delivery or coverage prior to the demonstration?

   Gaps in provider capacity for SUD treatment?

2. Of the changes you are making through the section 1115 SUD demonstration, which do you think are most likely to have the greatest impact in your state on the following areas:

   Access to SUD services?

   Improvement in provider capacity?

   Improvement in SUD-related outcomes?

3. Are there other changes you are making outside the section 1115 SUD demonstration that you expect to have significant impact on the population targeted by the SUD demonstration?

MEDICATION-ASSISTED TREATMENT

4. [*Only for states adding methadone for OUD].* We understand from review of your demonstration's documentation that your state is adding methadone for OUD as a Medicaid-billable service. What challenges have you faced, if any, by adding methadone for OUD as a Medicaid benefit? How are you addressing them?

   Provider education and capacity

   Billing

   Prior authorization

   Stigma

5. [*Besides the challenges we just discussed*] Have you faced any challenges specific to increasing access to [*methadone/OTPs*], buprenorphine or naltrexone? How are you addressing them?

   Increasing provider capacity—outreach, recruitment, education

   Policies that allow additional types of providers to prescribe MAT

   Expanding treatment into FQHCs or CMHCs

   Billing for MAT by specific provider types for specific medications

   Prior authorization

   Stigma

RESIDENTIAL TREATMENT

6. How has the demonstration changed your state's regulations and licensing criteria for SUD residential providers?

   To align with ASAM guidelines or any other criteria?

   Monitoring mechanisms you use (e.g., accreditation, site visits, etc.)

7. What challenges have you faced adding Medicaid SUD residential services your state, if any?

   Residential capacity, licensure, and provider requirements

   Challenges specific to adding/delivering MAT in residential settings

   Billing

   Prior authorization

   Independent process review placement in residential treatment settings

   Stigma

8. To what extent do you now track which residential facilities in your state offer MAT? If you do, do you know what proportion of facilities offer MAT? When did you start tracking?

   Do you track which medications each facility uses?

   Are you aware of preferences or challenges for different medications (e.g. is there a preference of one type of medication over another)?

9. Are there other changes you made to provide better access for MAT therapy in residential settings we haven't yet discussed (e.g., regulations, licensure requirements, policies)?

   On-site?

   Off-site?

OTHER SUD TREATMENT AND RECOVERY SUPPORT SERVICES

10. What challenges have you faced, if any, by adding other SUD treatment and recovery support services [*IOP/PH/ Withdrawal management/Peer support services/Other recovery management services*]? How are you addressing them?

    Missing levels of care

    SUD provider capacity

    Billing

    Peer support services

    Supported employment

    Supportive housing

    Mutual aid and other community-based services

    Case management

Transportation and childcare

Stigma

REIMBURSEMENT FOR SUD SERVICES

11. Have you made changes in reimbursement to other SUD treatment and recovery support services or other services we've not talked about as a part of the demonstration?

    Service delivery or payment models

    Contracting arrangements

    Increases in reimbursement rates

USE OF PATIENT PLACEMENT CRITERIA

12. *[We understand that your state is making some changes to patient placement criteria under the demonstration.]* We would like to get more details about processes you are putting in place to support these changes, including changes in utilization management, and monitoring of provider and MCO use of the criteria and new tools for assessment. Please describe steps you are taking in these areas:

    Use of a multidimensional assessment or some other instrument?

    Role of MCOs/third-party administrators/prepaid inpatient health plans

    Use of the criteria for prior authorization

    State oversight and monitoring

    Provider training

    Tracking use by providers

13. What challenges have you faced, if any, in making changes in this area? How are you addressing them?

CARE COORDINATION AND TRANSITIONS BETWEEN LEVELS OF CARE

14. How has the section 1115 demonstration changed your state's approach to care coordination and managing transitions between levels of care?

    Coverage of SUD case management

    MCO use of centralized care coordinators

    Bed tracking system for SUD providers

    Tracking post-discharge follow-up using claims data

    Use of peer navigators to connect people to services

    Incorporating performance metrics into MCOs contracts

    Efforts to improve integration of MH services into SUD specialty settings

15. What challenges have you faced, if any in this area? How are you addressing them?

OTHER CHALLENGES

16. Are there other implementation challenges under the SUD section 1115 demonstration that we have not yet discussed you would like to mention?

    How are you addressing them?

APPENDIX D:
ASSESSMENT OF AVAILABLE SUD HEALTH INFORMATION TECHNOLOGY PLANS AND RELATED METRICS

States with approved section 1115 SUD demonstrations must submit a SUD health information technology (IT) plan as a component of their implementation plans, related to Milestone #5 (Implementation of comprehensive treatment and prevention strategies to address opioid abuse and OUD). The SUD health IT plan specifically describes strategies to increase utilization and improve functionality of PDMPs, including the current status and future plan for: increasing PDMP functionalities; increasing PDMP query capabilities; increasing the use of the PDMP to support clinicians with changing their office workflow or business processes; increasing the use of the master patient index (MPI) to support SUD care delivery; and supporting overall state objectives for enhancing PDMP functionality and interoperability. These plans give an overview of the status of PDMP interoperability and support of clinicians in use of the PDMP, as well as plans for enhancing PDMPs and clinicians' use of those systems.

States must report at least three metrics that measure progress on their SUD health IT plans, as noted in the instructions from CMS for the state monitoring protocol. States can choose from sample metrics or develop their own measures. There are no specifically required metrics, but states must select metrics related to each of three key IT questions:

- How is IT being used to slow down the rate of growth of individuals identified with SUD?
- How is IT being used to treat effectively individuals identified with SUD?
- How is IT being used to effectively monitor "recovery" supports and services for individuals identified with SUD?

Of 14 preliminary and approved SUD health IT plans described in the Implementation Plans and/or State Monitoring Protocols available to RTI, nearly all states expressed interest in enhanced interstate data sharing and/or connectivity with health information exchanges and other entities. There is a gap in PDMP sophistication across demonstration states. For example, some have few existing connections to other states and others have many. States also planned many different approaches to make PDMPs easier and/or more useful to clinicians, such as implementing automated warnings for patients with multiple controlled substance prescriptions, integrating PDMPs with electronic health records, or enabling PDMP access by delegates (e.g., nurses, physician assistants). Many future activities will require legislative changes, additional financing, and the cooperation of non-public entities.

Of seven State Monitoring Protocols available to RTI, we noted a wide range of SUD health IT metrics selected by states, including:

- Number of providers using the PDMP
- Percentage of prescriber "report cards" sent electronically to providers that are opened by the receiving provider
- Number of clinical alerts sent electronically by the PDMP to providers
- Use of custom reporting frameworks
- Numbers of statewide fatal drug overdoses, and
- Total numbers of telehealth/telemedicine visits for beneficiaries with an SUD diagnosis.

We will continue to monitor state SUD health IT plans and monitoring protocols as they are made available to identify common features and core metrics that can be incorporated into the meta-evaluation.

APPENDIX E:
ASSESSING QUALITY OF NATIONAL DATA SOURCES FOR SUD DEMONSTRATION AND NON-DEMONSTRATION STATES

The Federal Meta-Analysis Support project supports CMS in studying the effectiveness of the set of section 1115 substance use disorder (SUD) demonstrations, and other demonstrations requested by CMS. The meta-analyses primarily will analyze results from the evaluations conducted by the demonstration states to determine the overall effectiveness of the demonstrations, and how the effectiveness varies by key characteristics. However, the data available from state evaluations may have limitations that, in turn, limit the meta-analyses. For example, states may not collect data on key outcomes, they may not have pre-period data, or they may not have comparison groups. Other data sources can be used to supplement and address limitations in the data and evaluation findings available from states. In the Evaluation Design Report, we have proposed to use national data sources for three purposes: (1) describe baseline conditions in demonstration and non-demonstration states (baseline analysis), (2) provide key outcomes and control variables for the meta-regressions, and (3) provide supplemental analyses that are not otherwise available in a standardized form.

We explored five data sets that we identified as potentially useful for examining outcomes that are a focus of the SUD demonstrations: SUD treatment capacity, use of SUD services, and opioid use disorder (OUD) overdose deaths. This memo describes the findings from our assessment of the following data sources: Medicaid claims data (Medicaid Analytic eXtract [MAX] research files and Transformed Medicaid Statistical Information System [T-MSIS] Analytic Files [TAFs]), which can be used to understand the use of SUD services, Medicaid expenditures, and, potentially, OUD-related overdose deaths; the Treatment Episode Data Set (TEDS), which can be used to understand the effect of the demonstrations on use of specialty SUD treatment services; the National Survey of Substance Abuse Treatment Services (N-SSATS); the Centers for Disease Control and Prevention Wide-ranging Online Data for Epidemiologic Research (CDC WONDER), which can be used to understand the effect of the demonstrations on OUD-related overdose deaths; and data on buprenorphine-waived providers from the Substance Abuse and Mental Health Service Administration's (SAMHSA's) Buprenorphine Practitioner Locator (BPL), which can be used to understand how SUD treatment capacity changed as a result of the demonstrations.

As part of our assessment we reviewed technical documentation for strengths and limitations of the data. We then downloaded publicly available files, constructed outcome variables and potential covariates, assessed observation and variable-level missingness, and

conducted basic univariate and time trend analyses. The high-level findings from our assessment of these potential data sources are as follows:

- **Medicaid claims data** will be a critical data source for the meta-evaluation because they contain a complete history of SUD treatments for beneficiaries that were paid by Medicaid. However, it has a few key limitations that make it necessary to supplement the Medicaid claims with other data sources. Namely,

  - Some key data elements such as date of death and provider information may be incomplete for some states and may not be available for a long enough pre-demonstration period to examine changes over time.

  - Medicaid claims data only contain data on services paid by Medicaid. To the extent that Medicaid beneficiaries receive SUD services paid by block grants or other sources, we would need to use other data sources to capture complete information on use of SUD services.

- **TEDS** has data quality issues that limit the types of analyses we could conduct for the meta-evaluation.

  - TEDS cannot support supplemental analyses of *Medicaid-covered* SUD treatment admissions for all states, but it could support an analysis of approximately 11 demonstration states that have complete payer information.

  - Because Medicaid beneficiaries may receive SUD services not funded by Medicaid, assessing *all* SUD treatment admissions using TEDS may identify broader impacts of the SUD demonstrations that Medicaid claims-based analyses cannot.

- **N-SSATS** can be very useful for the meta-evaluation because it captures public and private specialty SUD treatment capacity consistently for all states across time.

  - Although we anticipate that Medicaid claims may have some provider information, N-SSATS is a long-standing, established source of treatment capacity and may provide more complete measures of public and private provider capacity and more reliable trends over time for all states.

  - Some local evaluators may use N-SSATS to assess provider capacity, as recommended in CMS's *SUD Evaluation Design Technical Assistance*. We could fill in data gaps for states that do not use N-SSATS to harmonize analyses.

- **CDC WONDER** public use data files can support basic analyses of opioid-related overdose deaths from all types of opioids combined at the state level. Accessing restricted data may allow us to assess changes in specific types of opioid-related overdose deaths (e.g., those due to heroin or synthetic opioids) or changes at sub-state levels.

- Medicaid claims may have limited mortality information, and it is not yet clear how well mortality information could be linked to claims with an opioid-overdose-related International Classification of Diseases (ICD) code.

- Overdose deaths in CDC WONDER cannot be limited to Medicaid beneficiaries, but CDC WONDER would allow us to assess broader impacts of the SUD demonstrations on overdose deaths.

- **SAMHSA's BPL data** are not suitable for assessing the supply of buprenorphine-waived providers, but restricted access **Drug Enforcement Agency (DEA) Controlled Substances Act (CSA) Registration Records data** could be useful for tracking changes in buprenorphine capacity.

  - SAMHSA's data are updated continuously and may not be complete because providers can let their waiver lapse and providers are not required to list their information in the SAMHSA BPL.

  - The DEA provide access to the current complete list of waived providers, including both those with active and recently expired waivers.

  - Information from the DEA data could be merged with Medicaid claims using National Provider Identifiers to assess changes in provider supply across specialties (e.g., nurse practitioners) and prescribing patterns among waived providers with different patient caps.

Based on this data quality assessment, we conclude that the Medicaid claims data could support analysis on SUD use and expenditures, provider capacity, and rates of death among beneficiaries receiving SUD treatment; N-SSATS and DEA CSA registration records data could support assessment of treatment capacity impacts; and CDC WONDER could support assessment of opioid overdoses. Because of its significant limitations, we consider TEDS to be a lower priority for the meta-evaluation; however, we will reconsider its utility later in the project. More detailed description and analyses for each national data set are provided in the following sections.

E.1 Medicaid Claims

CMS Medicaid Analytic eXtract (MAX) research files and Transformed Medicaid Statistical Information System (T-MSIS) Analytic Files (TAFs) could be broadly used to assess changes in SUD treatment utilization associated with the SUD demonstration. We would use TAFs for years they are available and MAX data for years where TAFs are not available. Each state's Medicaid Statistical Information System (MSIS) data are the source of the MAX files, and each state's T-MSIS data are the source for the TAFs. The TAF and MAX file processing adds enhancements to the source files, such as claims adjustments, creation of a national type of

service field, and state-specific quality issues corrections. T-MSIS data contain over 1,000 elements, including many data elements not previously available in MAX files. The MAX files include a person summary file, with all enrollment information and summary claims information and the TAF files include a baseline summary file with all enrollment information. Both the MAX and TAF files include four claims files: inpatient hospital, long-term care, prescription drugs, and other claims.

For this analysis, we discuss measures that Medicaid claims data (namely, MAX, TAFs, T-MSIS, and MSIS) could assess pre- and post-demonstration for SUD Milestone #1 (Access to critical levels of care for OUD and other SUDs), SUD Milestone #4 (Sufficient provider capacity at each level of care), and SUD Goal #3 (Reduce overdose deaths, particularly those due to opioids) and that may overlap with the four non-claims national data sources (namely, TEDS, N-SSATS, CDC WONDER, and the DEA's CSA registration records data). ***Table E-1*** lists the specific Medicaid claims measures that are comparable to measures in the other national data sources.

Table E-1. Medicaid Claims Measures for SUD Demonstrations Relevant to Other National Data Sources

Question	Medicaid claims measures
SUD Milestone #1: Access to critical levels of care for OUD and other SUDs	
Did states increase utilization of critical levels of care for SUD during their SUD demonstration period?	• Number of Medicaid beneficiaries enrolled receiving any SUD treatment service, facility claim, or pharmacy claim • Number of Medicaid beneficiaries who use outpatient services for SUD • Number of Medicaid beneficiaries who use intensive outpatient and/or partial hospitalization services for SUD • Number of Medicaid beneficiaries who use residential and/or inpatient services for SUD • Number of Medicaid beneficiaries who use withdrawal management services • Number of beneficiaries who have a claim for medication-assisted treatment (MAT) for SUD • Average length of stay for Medicaid beneficiaries discharged from Institutions for Mental Diseases residential treatment for SUDs
SUD Milestone #4: Sufficient provider capacity at each level of care	
Did states increase the overall number of SUD providers?	• Number of providers with a SUD specialty taxonomy code in the Medicaid provider data • Number providers prescribing at least one MAT
SUD Goal #3: Reduce overdose deaths, particularly those due to opioids	
Did states reduce opioid-related overdose deaths?	• Opioid-related overdose deaths per 100,000 Medicaid beneficiaries

MAT = medication-assisted treatment; OUD = opioid use disorder; SUD = substance use disorder.

The advantage of the claims data over other sources is that it contains a complete history of SUD treatments for Medicaid beneficiaries that were paid by Medicaid. Consequently, the Medicaid claims data will be an essential data source to assess impacts on outcomes of focus for the SUD demonstrations. However, there are key limitations that necessitate supplementing the Medicaid claims data analysis with other data sources. Specifically, the TAF processing is in early stages, so many data elements are currently missing or incomplete. TAFs vary in their availability by state, but most states have TAFs available beginning in 2014 or 2015. Some data elements of interest, such as date of death and Medicaid provider information, were not available in the MAX data, so we could not do a pre-post analysis of these outcomes for early SUD demonstration states using Medicaid claims data alone.

For overdoses, we should be able to identify these through ICD codes, and it is possible that we will have dates of death, but we have not yet been able to assess the prevalence of the overdose ICD codes and how well we can link an overdose to a death. Using claims may also miss individuals who overdose but do not have an associated claim. We can also examine SUD provider capacity by measuring changes in the number of providers with a SUD specialty code over time; however, this measure will not be accurate unless the Medicaid provider file contains complete information about providers' specialty codes.

Moreover, Medicaid claims data only contain claims for services that were paid by Medicaid. We would need other data sources to capture SUD treatment for Medicaid beneficiaries that was paid by other sources.

E.2 Treatment Episode Data Set (TEDS)

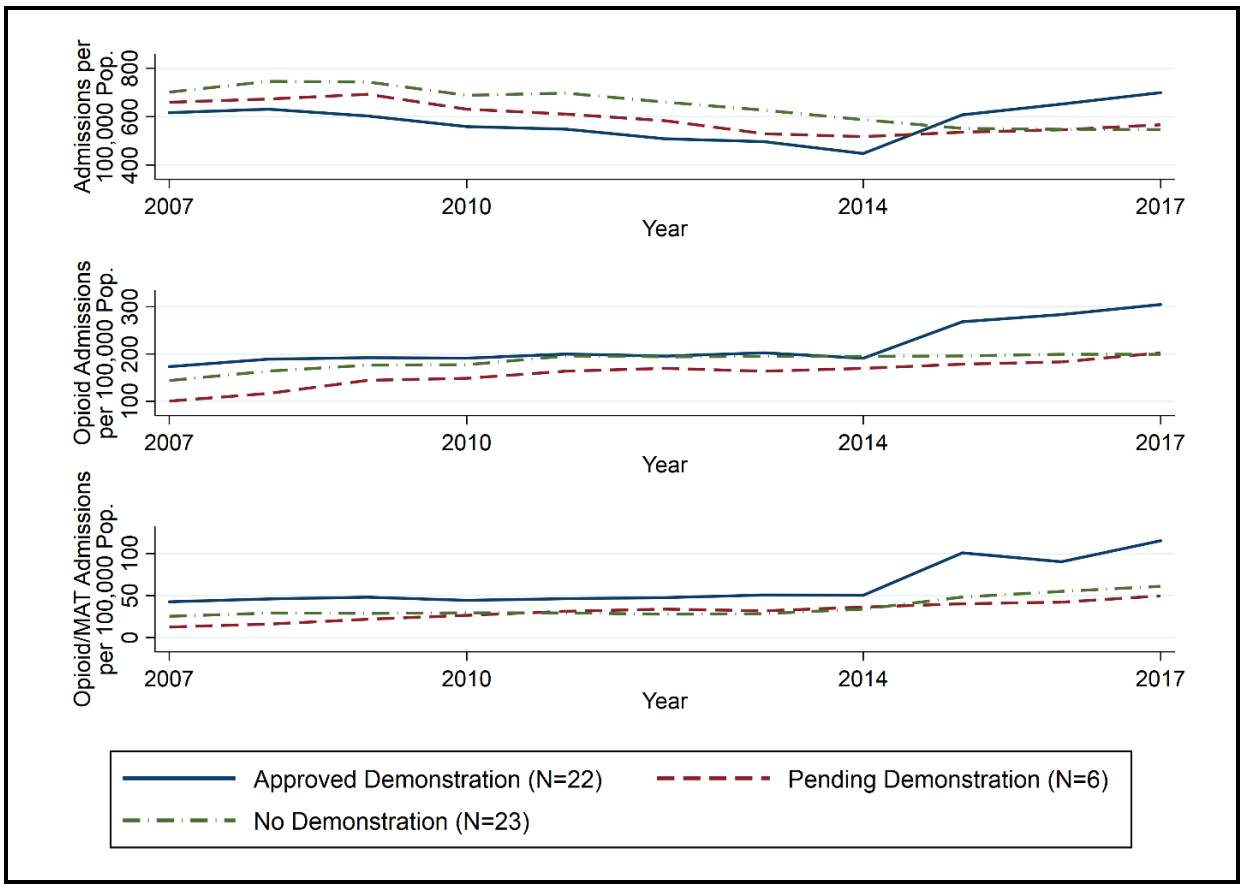
The Treatment Episode Data Set: Admissions (TEDS) comprises administrative records of 1.5 million to 2 million admissions to publicly funded substance abuse treatment across all 50 states and DC. It contains patient demographic and socioeconomic characteristics, referral source, payment source, substances used and frequency, treatment history, and treatment plans (i.e., use of medication-assisted treatment, or MAT). TEDS is an option for assessing pre- and post-demonstration admissions trends to address SUD Milestone #1 (Access to critical levels of care for OUD and other SUDs). Counts of admissions of certain types (e.g., opioid-related admissions) are particularly important for understanding the experience of demonstration states prior to implementation as well as for identifying suitable comparison states. ***Table E-2*** lists the specific TEDS measures that could be informative. Some of these measures are presented in ***Figure E-1*** stratified by demonstration and non-demonstration states.

Table E-2. TEDS Measures for the SUD Demonstrations

SUD Milestone #1: Access to critical levels of care for OUD and other SUDs	TEDS measures
Did states increase utilization of critical levels of care for SUD during their SUD demonstration period?	• Number of admissions to SUD specialty treatment • Number of admissions with MAT planned
Did states increase utilization of critical levels of care for OUD during their SUD demonstration period?	• Number of opioid-related admissions to SUD specialty treatment • Number of opioid-related admissions with MAT planned
Did states increase utilization of critical levels of care for SUDs other than OUD during their SUD demonstration period?	• Number of non-opioid-related admissions to SUD specialty treatment • Number of non-opioid-related admissions with MAT planned

MAT = medication-assisted treatment; OUD = opioid use disorder; SUD = substance use disorder; TEDS = Treatment Episode Data Set.

Figure E-1. SUD Treatment Admissions per 100,000 Population for Approved, Pending, and Non-Demonstration States



SUD = substance use disorder; MAT = medication-assisted treatment.

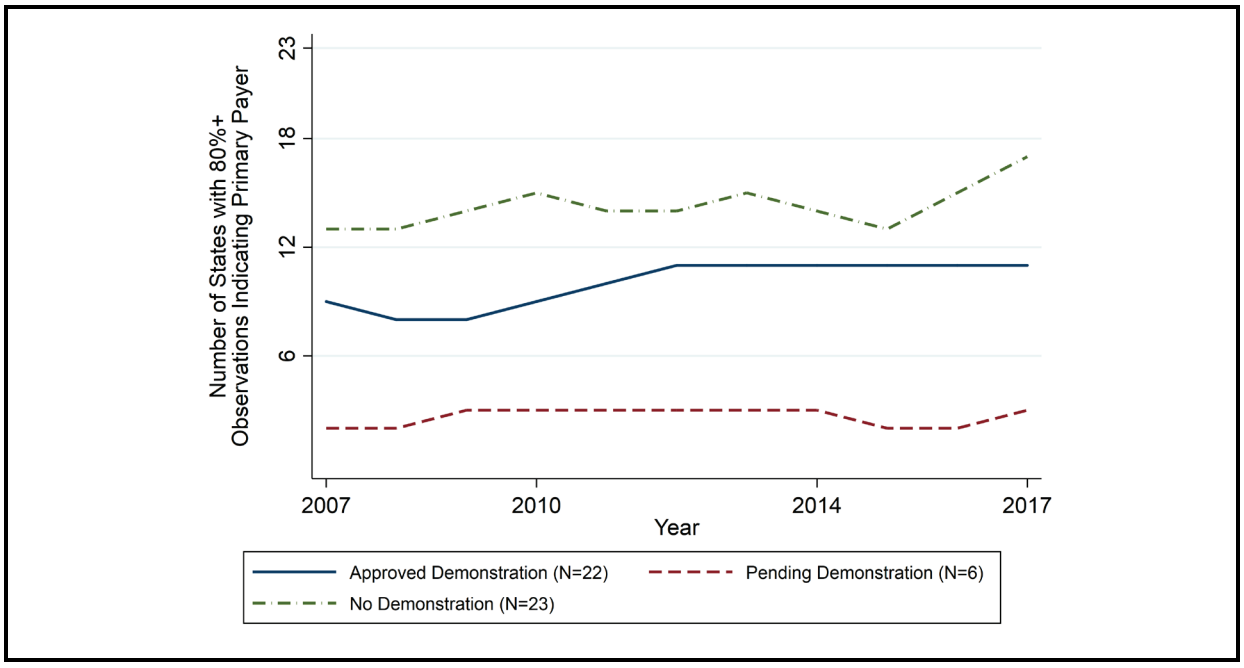
Source: RTI analysis of TEDS Public Use Files: 2007-17, retrieved from: https://wwwdasis.samhsa.gov/dasis2/teds.htm

We identified several limitations with TEDS that make it challenging to use TEDS to examine SUD Milestone #1:

- States may not report data every year.
- TEDS does not track Medicaid-specific admissions.
- There is volatility in state-level trends that suggests potential data quality problems.

Some states do not report any data for certain years: six states have failed to report in at least 1 year between 2007 and 2017. The TEDS minimum data set requirements include individual-level demographic characteristics, substances used and frequency of use, referral sources, prior treatment, and treatment services planned (i.e., medications for OUD). However, primary payer source is an optional field that many states do not report (***Figure E-2***). As shown in ***Figure E-2***, as of 2017, payment information was missing for 11 of the 22 approved SUD demonstration states; 3 of the 6 pending demonstration states; and 6 of 23 non-demonstration states. We could consider an analysis of Medicaid-specific admissions limited to the 20 states with payment source information. We could only include all states if we analyzed all SUD treatment admissions regardless of payer.

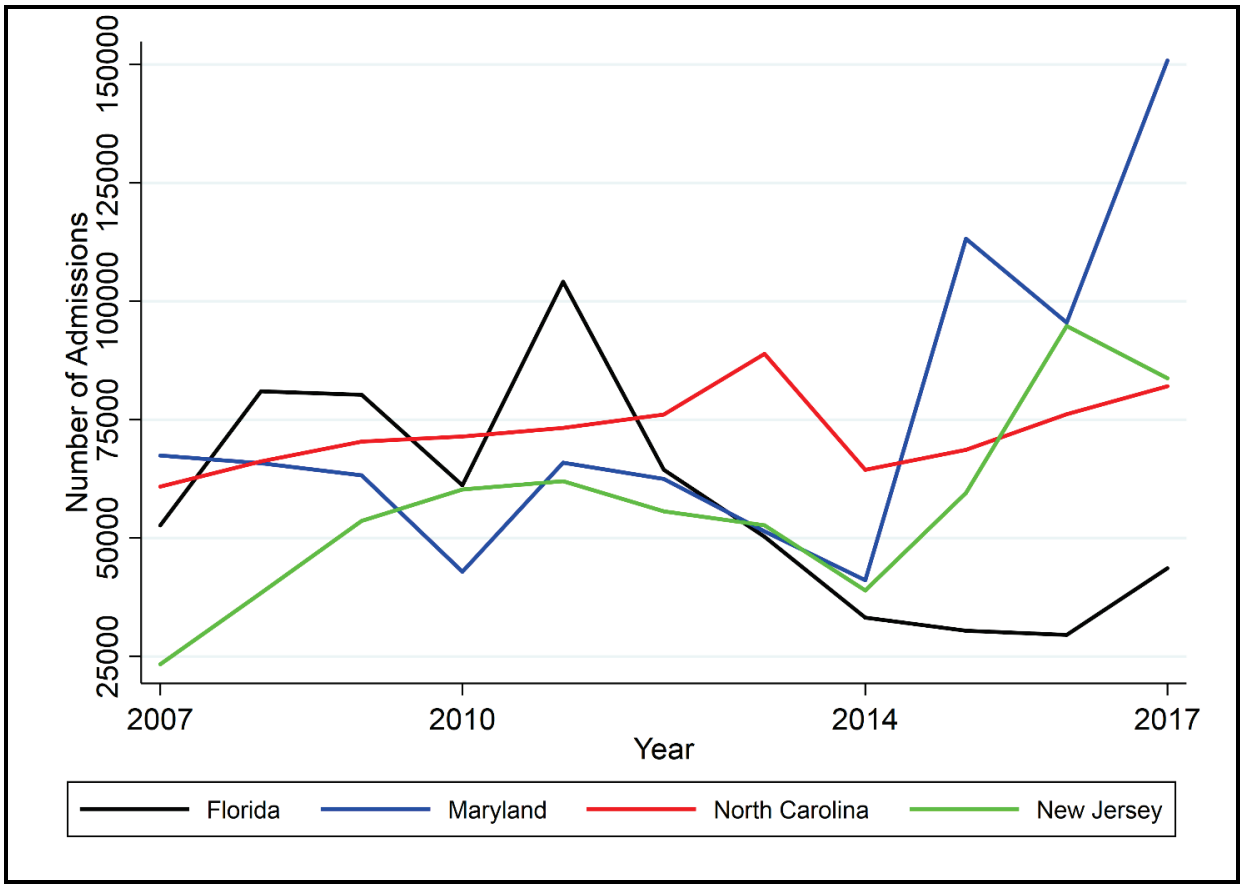
Figure E-2. Availability of Primary Payer Source for Approved, Pending, and Non-Demonstration States



Source: RTI analysis of TEDS Public Use Files: 2007-17, retrieved from: https://wwwdasis.samhsa.gov/dasis2/teds.htm

Total admissions can be volatile from year to year within a state, suggesting inconsistency in reporting. For example, reported total admissions in Illinois fell by 50% between 2010 and 2012. Inconsistencies may be caused by changes in state reporting requirements or data errors. For example, states are instructed to report one record per treatment episode, rolling up transfers to different facilities and multiple providers in a single record; this may not be done consistently each year, which leads to overcounting admissions and contributes to within-state variation over time. ***Figure E-3*** demonstrates that total admissions in some states can vary considerably from year to year. This year-to-year variability is masked in ***Figure E-1,*** which displays aggregates across states.

Figure E-3. SUD Treatment Admissions as Recorded in TEDS: Select States



SUD = substance use disorder; TEDS = Treatment Episode Data Set.

Source: RTI analysis of TEDS Public Use Files: 2007-17, retrieved from: https://wwwdasis.samhsa.gov/dasis2/teds.htm

TEDS has a few other limitations of note. Although it captures patient-level information, TEDS does not include patient identifiers, so it is not possible to identify unique patients and patients cannot be tracked across multiple admissions. Also, although TEDS includes data for all

admissions reported to the state substance abuse agency, states vary in the types of providers that are required to report. Providers that are publicly funded and/or licensed or certified by the state agency are required in most states to report and, therefore, are included in TEDS. Thus, TEDS is likely more representative of publicly funded facilities and less so of private facilities.

These issues limit the types of analyses TEDS can support for the meta-evaluation. Stratification by payment source is only possible in approximately 50%–60% of states each year. Major shifts in admission counts from year to year call into question the validity of the data in some states and may present challenges to identifying the impacts of the demonstrations. Although the outcomes cannot be restricted to Medicaid admissions in all states, if we could identify states with reliable trends, understanding whether there are broad impacts of the SUD demonstration beyond Medicaid could be informative. It is possible that Medicaid beneficiaries may receive additional SUD services financed by block grants or other funding sources external to Medicaid, and TEDS may pick up these additional services that would not appear in Medicaid claims. Supplemental analyses of Medicaid-covered admissions could also be conducted using TEDS for a smaller subset of states.

E.3 National Survey of Substance Abuse Treatment Services (N-SSATS)

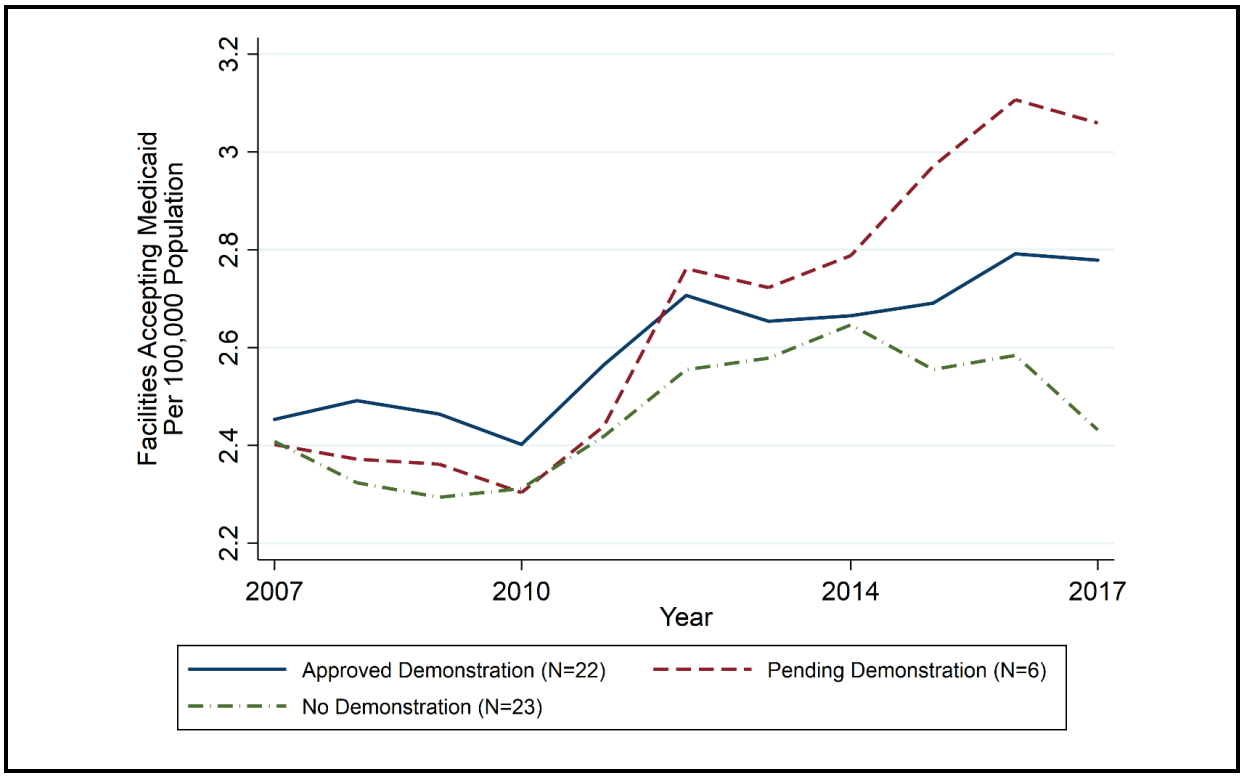
N-SSATS is an annual survey of facilities that have been identified by SAMHSA as providing SUD treatment. N-SSATS captures information on participating facilities including data on payment sources accepted, levels of care and services provided, clinical approaches used, and counts of patients receiving treatment on a single reference day (typically a day at the end of March). N-SSATS could be used to measure changes in pre- and post-demonstration capacity for SUD Milestone #4 (Sufficient provider capacity at each level of care). ***Table E-3*** lists the specific N-SSATS measures we could use. Some of these measures are presented in ***Figure E-4*** stratified by demonstration and non-demonstration states.

Table E-3. N-SSATS Measures for the SUD Demonstrations

SUD Milestone #4: Sufficient provider capacity at each level of care	N-SSATS measures
Did states increase the overall number of SUD providers?	• Number of SUD facilities per 100,000 population • Number of SUD facilities that accept Medicaid per 100,000 population
Did states increase the number of residential SUD facilities?	• Number of SUD residential facilities per 100,000 population • Number of SUD residential facilities that accept Medicaid per 100,000 population
Did states increase the number of outpatient SUD facilities?	• Number of SUD outpatient facilities per 100,000 population • Number of SUD outpatient facilities that accept Medicaid per 100,000 population
Did states increase the number of facilities offering MAT?	• Number of facilities offering MAT per 100,000 population • Number of facilities offering MAT that accept Medicaid per 100,000 population

MAT = medication-assisted treatment; N-SSATS = National Survey of Substance Abuse Treatment Services; OUD = opioid use disorder; SUD = substance use disorder.

Figure E-4. Number of Facilities That Accept Medicaid per 100,000 Population: Approved, Pending, and Non-Demonstration States



Source: RTI analysis of N-SSATS Public Use Files: 2007-17, retrieved from: https://wwwdasis.samhsa.gov/dasis2/nssats.htm

Although N-SSATS has some limitations, they are relatively minor and N-SSATS is a strong data source for the meta-evaluation. N-SSATS provides the better information nationally on the characteristics of each state's specialty SUD provider network relative to claims data sources. It has been administered reasonably consistently over time and its facility characteristic information is not available from any other publicly available source. Unlike TEDS, all publicly and privately owned facilities are included in the survey. N-SSATs has a greater than 90% response rate nationally in most years.

There are two minor limitations to the use of N-SSATS data. First, although the overall response rate is high, non-response varies by state (***Table E-4***). Second, while the survey is designed for analysis at the facility level, the unit of response depends in part on state licensing practice. For example, in some states a facility with different levels of care receive separate licenses for each, while in other states they receive a single license. Thus, states will vary in whether there are multiple responses from a single facility.

Table E-4. N-SSATS National Response Rates (%) and States with Low Response Rates, by Year, 2007–2017

Response rate measure	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017
Total Response Rate	94.5	94.1	93.4	91.4	94.0	93.1	94.4	93.7	91.6	91.4	89.2
States with 85% or Lower Response Rate			NC 84.2	DC 81.0				AR 83.9	AR 83.9	KY 85.0	DE 79.2
				NC 83.8						ME 72.6	DC 77.1
										NC 80.7	GA 82.9
											NV 82.5
											NM 83.7
											NC 77.5
											TX 82.0

N-SSATS = National Survey of Substance Abuse Treatment Services.

Source: RTI analysis of N-SSATS Annual Reports: 2007-17, retrieved from: https://wwwdasis.samhsa.gov/dasis2/nssats.htm

N-SSATS also contains limited information on single-day counts of patients in treatment, which could be used in combination with or instead of TEDS to understand treatment admissions. However, we do not think this would be a good approach. Client counts would also not be Medicaid specific and are only measured every other year (odd years), which would create a challenge in creating accurate time trends.

Although we anticipate that Medicaid claims may have some provider information, N-SSATS is a long-standing, established source of treatment capacity and may provide more

complete measures of public and private provider capacity and more reliable trends over time for all states. Local evaluators may themselves use N-SSATS to assess provider capacity. As recommended in CMS's *SUD Evaluation Design Technical Assistance*, we could fill in data gaps for states that do not use N-SSATS to harmonize analyses.

E.4 CDC WONDER

CDC WONDER contains two mortality databases that capture information on opioid-related overdose deaths at the state and county level based on death certificates for U.S. residents: The Underlying Cause of Death database and the Multiple Cause of Death database. The type of opioid can be disaggregated into four categories: heroin, methadone, natural and semisynthetic opioids, and synthetic opioids other than methadone. CDC WONDER could be used to assess pre- and post-demonstration opioid-related overdose deaths for SUD Goal #3 (Reduce overdose deaths, particularly those due to opioids). ***Table E-5*** lists the specific CDC WONDER measures we could use. These exploratory analyses are based on the public use version of CDC WONDER.

Table E-5. CDC WONDER Measures for the SUD Demonstrations

SUD Goal #3: Reduce overdose deaths, particularly those due to opioids	CDC WONDER measures
Did states reduce opioid-related overdose deaths?	• Opioid-related overdose deaths per 100,000 • Heroin-related overdose deaths per 100,000 • Natural and semisynthetic opioid-related overdose deaths per 100,000 • Methadone-related overdose deaths per 100,000 • Non-methadone, synthetic opioid-related overdose deaths per 100,000

CDC WONDER = Centers for Disease Control and Prevention Wide-ranging Online Data for Epidemiologic Research; SUD = substance use disorder.

Due to data suppression, the public use version of CDC WONDER can only be used to assess state-level opioid-related overdose deaths for all types of opioids combined. CDC WONDER suppresses data if there are fewer than 10 mortalities in a state or county and marks rates as "unreliable" when the death count is less than 20. Data suppression is not an issue for *all* opioid-related overdose deaths at the state level—only one state-year observation is suppressed from 2008 to 2017 (North Dakota, 2012). However, disaggregating the data by type of opioid at the state level is not feasible due to suppression (see ***Table E-6***). Rates of data suppression are even higher at the county level (see ***Table E-7***).

Table E-6. Counts of State-Level Suppressed Data by Drug in CDC WONDER, 2008–2017

Drug	Number of state-years with suppressed observations	Percent of state-years suppressed	Number of state-years with unreliable rates	Percent of state-years unreliable
Heroin	101	19.80%	51	10.00%
Natural and Semisynthetic Opioids	7	1.37%	20	3.92%
Methadone	54	10.59%	62	12.16%
Synthetic Opioids, Other Than Methadone	53	10.39%	62	12.16%

CDC WONDER = Centers for Disease Control and Prevention Wide-ranging Online Data for Epidemiologic Research.

Source: RTI analysis of CDC WONDER, 2008-17, retrieved from: https://wonder.cdc.gov/mcd.html

Table E-7. Counts of County-Level Suppressed Data by Drug in CDC WONDER, 2008–2017

Drug	Number of county-years with suppressed observations	Percent of county-years suppressed	Number of county-years with unreliable rates	Percent of county-years unreliable
Heroin	5,007	73.09%	839	12.25%
Natural and Semisynthetic Opioids	3,953	57.71%	1,618	23.62%
Methadone	5,996	87.53%	553	8.07%
Synthetic Opioids, Other Than Methadone	5,446	79.50%	701	10.23%

CDC WONDER = Centers for Disease Control and Prevention Wide-ranging Online Data for Epidemiologic Research.

Source: RTI analysis of CDC WONDER, 2008-17, retrieved from: https://wonder.cdc.gov/mcd.html

CDC WONDER public use files could support analyses supplemental to local evaluation reports and Medicaid claims using a difference-in-differences approach for all types of opioid-related overdose deaths. A limitation of CDC WONDER is that overdose deaths cannot be limited to Medicaid beneficiaries, unlike Medicaid data sources that states may use for their local evaluation or Medicaid claims. However, CDC WONDER's measure of overdose deaths, which includes a wider population, may allow us to assess broader impacts of the SUD demonstrations on overdose deaths. Although a date of death data element was added the Medicaid claims, it is unclear whether the date of death field will be populated in the Medicaid claims for all states. We may be able to use CDC WONDER to fill in missing mortality information for states where it is not available from Medicaid data sources.

We also recommend pursuing access to the restricted use Underlying Cause of Death and the Multiple Cause of Death databases available from the National Center for Health Statistics (NCHS) to overcome data suppression issues in WONDER. NCHS data are available through the Census Bureau's Federal Statistical Research Data Centers. The restricted data would provide access to all state- and county-level overdose death counts, in aggregate and by type of opioid. Medicaid claims may have limited mortality information, and it is not yet clear how well mortality information could be linked to claims with an opioid-overdose-related ICD code. Assessing sub-state changes and changes by type of opioid may provide richer insights into demonstration impacts.

E.5 Buprenorphine Practitioner Locator (BPL)

The Buprenorphine Practitioner Locator (BPL) is a database maintained by SAMHSA that tracks medical providers who have an approved buprenorphine waiver. This publicly available data allows consumers to identify nearby buprenorphine-waived providers and allows users to download the full list of providers. Thus, the BPL could be used to measure changes in opioid treatment capacity across states and/or counties. BPL could be used to assess pre- and post-demonstration waivered-provider trends to address SUD Milestone #4 (Sufficient provider capacity at each level of care); ***Table E-8*** lists the specific BPL measure we could use.

Table E-8. BPL Measures for the SUD Demonstrations

SUD Milestone #4: Sufficient provider capacity at each level of care	BPL measures
Did states increase the number of providers with a buprenorphine waiver?	• Number of waived physicians per 100,000 population

BPL = Buprenorphine Practitioner Locator; SUD = substance use disorder.

The SAMHSA BPL is insufficient for the analysis, so we instead propose to obtain a related DEA data file, the Controlled Substances Act (CSA) Registration Records.

- Publicly available SAMHSA data are continuously updated based on approved applications, making it challenge to extract historical data and to obtain accurate point-in-time estimates.

- Also, providers may let their approval lapse, and providers with approved waivers are not required to list their information in the SAMHSA BPL, so the data are not comprehensive.[d]

Once SAMHSA approves a waiver, the information is forwarded to the DEA. The DEA assigns a registration number, which encodes the type of provider (e.g., physician, nurse practitioner) and the patient limit (25, 100, or 275). Providers must renew their waiver annually, and the DEA tracks both active providers and providers who have let their waiver expire in the CSA registration records. The Active File consists of records of all individuals registered under the CSA, including registrants doing business under their individual name rather than a business name. The DEA also produces a list of registrants whose certification numbers have been retired or suspended from the Active File. This data file could give us the information needed to track changes in buprenorphine-waived-provider capacity.

The DEA CSA registration records data could support analyses supplemental to local evaluation reports and Medicaid claims using a difference-in-differences approach. Local evaluations may use provider information in claims data to track changes in provider capacity, but Medicaid claims data could only be used to track providers who prescribed buprenorphine. Again, although we anticipate that Medicaid claims will have comprehensive provider information, the DEA CSA registration records data may provide more reliable trends over time for all states. Additionally, the DEA CSA registration records data could be merged with claims data using National Provider Identifiers to assess changes in capacity by provider specialty (e.g., nurse practitioners) and prescribing patterns among waived providers with different patient caps.

[d] Mir M. Ali, R. Ghertner, L. Fuller, and J. Dubenitz (2019). Public Listing Status of DATA-Waivered Providers. U.S. Department of Health and Human Services, Office of the Assistant Secretary for Planning and Evaluation, Office of Disability, Aging and Long-Term Care Policy. Accessed at: https://aspe.hhs.gov/system/files/pdf/261866/DWPlisting.pdf

APPENDIX F:
HEALTH EVALUATION ENGINE DRAFT USER'S GUIDE

What is the Health Evaluation Engine?

The Health Evaluation Engine (see ***Figure F-1***) is a point-and-click user interface for a set of R routines, which conduct propensity score matching and linear regression analysis using a difference-in-differences (DiD) specification. These are the basic building blocks of many health care evaluations.

Prior to using the Health Evaluation Engine, users should have at least an introductory level of familiarity with the underlying evaluation concepts and statistical methods (i.e., propensity score matching and DiD analysis). Below we provide a brief primer on these concepts and methods. Users for whom these concepts are new may find it helpful to learn more about the statistical methods and may want to consult the references in the "Additional Readings" section.

What does the Health Evaluation Engine do?

The Health Evaluation Engine performs two statistical analyses. The first is a propensity score analysis, and the second is a regression-adjusted impact analysis.

Basics of a propensity score analysis. The goal of a health care evaluation is to measure whether changes in health care policies or procedures impact key outcomes. As an example, one might be interested in understanding whether expanding the types of behavioral health care services that are covered under a state's Medicaid program impacts key health care outcomes such as reducing the use of emergency departments. If these services are expanded in specific regions of a state, then there is an opportunity to compare health care outcomes among beneficiaries within those regions where services were expanded versus beneficiaries within those regions where services were not expanded. As evaluators, we would refer to beneficiaries within those regions where services were expanded as the "treatment group," and we would refer to beneficiaries within those regions where services were not expanded as the "control group."

In this type of study, it is important to isolate differences in outcomes that are only related to the differences in the types of behavioral health care services that are covered across the two regions. For example, imagine that beneficiaries in the intervention group are much younger than beneficiaries in the control group. This difference in age between the two regions could lead to lower rates of health care utilization—including emergency department visits—

among the intervention group relative to the comparison group even if the intervention group had not been exposed to an expansion of the types of behavioral health care services that are covered.

This is the kind of problem that propensity score methods can address. Specifically, propensity score matching can be used to identify persons from the control group that are most like persons in the intervention group in terms of characteristics such as age, gender, health status, and other observable factors. Accordingly, propensity score methods allow researchers to measure differences in health care outcomes across two populations where the only difference is that the intervention group was exposed to an intervention and the control group was not.

Propensity score matching methods. The Health Evaluation Engine implements propensity score matching. One of the most commonly applied propensity score matching methods is nearest neighbor matching, which matches persons from the control group to persons in the intervention group by finding persons who have the closest propensity score. A propensity score can be thought of as a single number that summarizes all the characteristics that are observed about the persons in your data (e.g., age, gender, health status, etc.). Accordingly, two persons with very close propensity scores will look very similar to each other in terms of the characteristics you can observe about them in your data. Other types of matching are essentially refinements of this idea.

Although there are more complex matching methods than nearest neighbor matching, there are still some important analytic choices that must be made with nearest neighbor matching (and potentially with other types of matching). For example:

- Should we consider excluding any individuals from the analysis if there are no good matches for certain persons?
- Should we only match persons in the comparison group who have propensity scores that are within a pre-defined distance from persons in the intervention group?
- Does it matter if we start by assigning matches to persons in the intervention group with the largest propensity score, the smallest propensity score, or should we start assigning matches randomly?
- Once we have matched a person from the comparison group to a person in the intervention group should that comparison subject remain available for matching to other persons in the intervention group?

For each of these questions, the Health Evaluation Engine has an option. Specifically:

- The "Discard Units" option allows the user to allow the propensity score matching algorithm to exclude persons from the treatment group, the control group, or both from the analysis if they are too dissimilar.

- The "Caliper Value" option allows the user to refine the idea above, and only choose matches that are within a pre-defined distance from persons in the intervention group.

- The "Match Order" option allows the user to specify the order in which matches are found.

- The "Replace" option allows the user to specify that control group persons can be matched more than once.

Other methods available in the Health Evaluation Engine. The following propensity score matching methods are available in the Health Evaluation Engine:

- **Subclassification.** This method uses the propensity score to form subclasses or subgroups of intervention and comparison group members who have similar propensity scores. The user must specify the number of subclasses to create, and the default number is 6.

- **Nearest Neighbor.** This method uses the propensity score to match persons from the comparison group to persons in the intervention group based on how close the propensity scores are.

- **Optimal.** This method is essentially the same as nearest neighbor except that instead of choosing the order in which matches are formed, an algorithm is implemented to form matches that minimize a global distance metric across all possible matches.

- **Full.** This method is very similar to subclassification, except that it forms subclasses in a more "optimal" way.

- **Genetic.** This method implements an algorithm that chooses matches and sets weights in such a way that distances between the observed characteristics about the persons in your data are minimized.

- **Nearest Neighbor Subclassification.** This method combines nearest neighbor with subclassification as a second stage.

Regression-adjusted impact analyses. After matches have been formed using one of the above propensity score methods, the next analysis that the Health Evaluation Engine implements is a regression-adjusted impact analysis. The specification employed is a difference-in-differences (DiD) specification and has the following form:

$$Y_{it} = \alpha + \beta Intervention_i + \Sigma_t \gamma_t Period_t + \Sigma_k \delta_k Period_k * Intervention_i + \boldsymbol{X}_{it}\boldsymbol{\theta} + \varepsilon_{it},$$

where Y_{it} denotes the outcome variable, $Intervention_i$ (= 0, 1) denotes an indicator for the intervention group (= 1) versus the comparison group (= 0), $Period_t$ (= 0, 1) denotes a set of period indicators for each period included in the dataset, and $\boldsymbol{X}_{it}$ denotes a set of regression controls that are included in the dataset. The interaction terms $Period_k * Intervention_i$ are only defined for periods after the intervention has been implemented.

How to use the Health Evaluation Engine?

Once installed, there is an example dataset that can be used to learn more about the functionality of the application.

To use the Health Evaluation Engine with your own data, you will need to identify two files for the application. The first file is the analytic data file, which contains your outcome and the covariates to include in the propensity score and impact analyses. The second file is a meta-data file, which contains data on your data that the R code underlying the Health Evaluation Engine needs to reference to know things like which variable is the outcome, which variables should be included only in the propensity score model, and so on. More detail on these files is provided below.

What types of data are expected to be included in the analytic data file? The analytic data file that you upload to the application should have at least one data period before the intervention and at least one data period after the intervention. However, the data file can have multiple periods before and multiple periods after the intervention as well. In addition, the Health Evaluation Engine assumes that the analytic data file will have each of following types of data:

- An outcome variable (e.g., emergency department visit count)
- A treatment group indicator (= 0, 1) that equals 1 for the treatment group and equals 0 for the control group
- A set of period indicators (= 0, 1) for each of the data periods included in your analytic file
- An indicator for the post-intervention period (= 0, 1) that equals 1 in all post-intervention periods and equals 0 in all periods prior to the intervention
- Any other variables that you have that characterize the persons in your data, and that you think may be related to the outcome

Figure F-1. Health Evaluation Engine Landing Page