Evaluation Design Technical Assistance Guide For Section 1115 Demonstrations: Family Planning Demonstrations

This document provides technical assistance for evaluating section 1115 demonstrations that provide coverage for family planning (FP) or family planning-related services ("FP demonstrations"). It includes a description of the goals (Section 1), contains an example logic model (Section 2) linking demonstration activities to expected outcomes, hypotheses, and research questions (Sections 3 and 4), and presents potential data sources (Section 5) and analytic approaches (Tables 1 and 2).

The Centers for Medicare & Medicaid Services (CMS) provides evaluation technical assistance guides for several other common demonstration policies.[1] States with multiple policies in their demonstration should consult relevant policy-specific evaluation technical assistance guides to develop comprehensive demonstration evaluation designs aligned with Special Terms and Conditions (STCs) requirements.

1. Demonstration goals

States with FP demonstrations intend to test whether the demonstration has the following effects (ordered from short-term to long-term in line with the logic model discussed in Section 2):

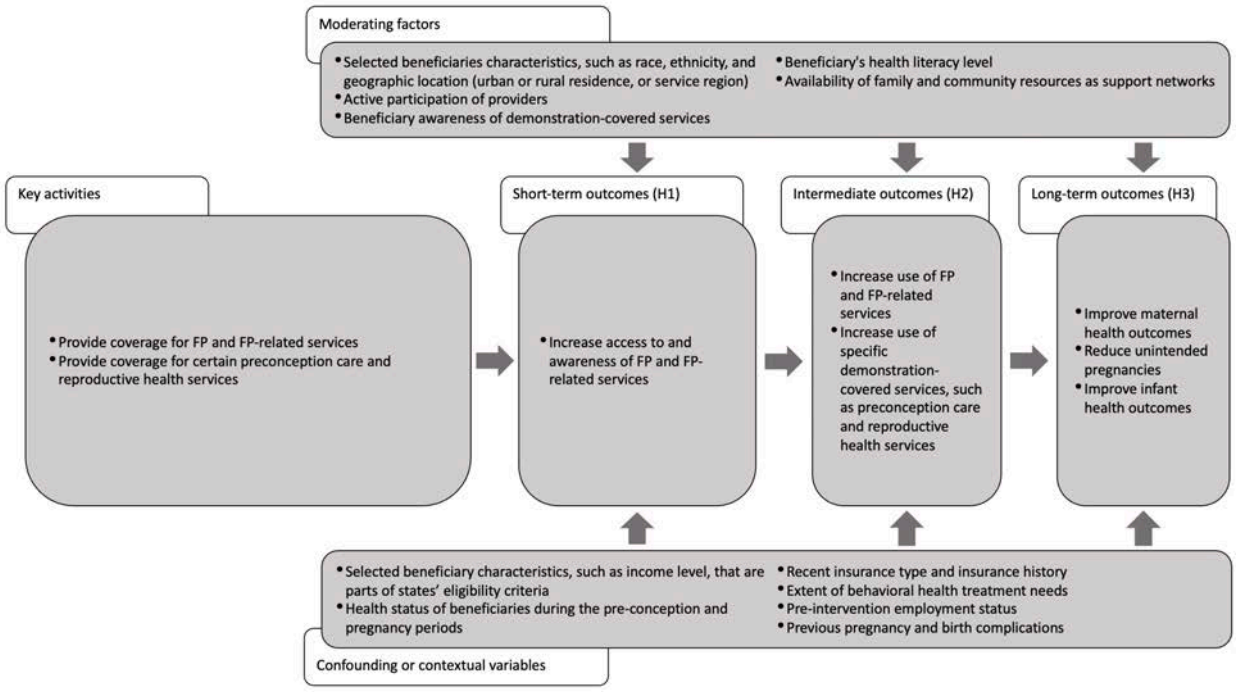
1. Enhance awareness of FP demonstrations and improve participation in programs and access to FP and other FP demonstration-covered services among eligible beneficiaries.
2. Increase the effective use of FP services or other FP demonstration-covered services among demonstration beneficiaries.
3. Improve maternal and infant health outcomes, particularly focusing on reducing unintended pregnancies.
4. States should also include any state-specific goal(s) they have identified, which could include those that address health conditions and behaviors that the demonstration seeks to impact, even if they are not directly related to FP services. For example, some FP demonstrations have goals to encourage tobacco cessation or to facilitate postpartum depression treatment.

2. Example logic model for FP demonstrations

Figure 1 shows an example logic model. It depicts how the demonstration affects outcomes, accounting for moderating and confounding or contextual variables. A state's logic model should reflect the specific FP or FP-related services its demonstration offers and any other relevant state-specific context of activities.

[1] These resources are available to states on the Medicaid.gov website: https://www.medicaid.gov/medicaid/section-1115-demonstrations/1115-demonstration-monitoring-evaluation/1115-demonstration-state-monitoring-evaluation-resources/index.html.

Figure 1. Example logic model for family planning demonstrations



Notes: Letter H and numbers in parentheses refer to the hypotheses listed in Section 4.

3. Research questions related to demonstration implementation

To understand the implementation of FP demonstrations and opportunities to improve demonstration operations, the state should specify a set of implementation questions. These questions can provide context for analyses that address hypotheses and assess progress toward demonstration goals.

Table 1 lists implementation questions along with recommended measures, data sources, and analytic approaches. To answer the implementation questions, the state can leverage data it has already presented in its monitoring reports to assess program implementation approaches, trends, and lessons learned. The state should also consider collecting additional data through interviews, surveys, and/or focus groups with key entities and beneficiaries to understand their experience with the demonstration, develop a deeper understanding of barriers and facilitators, and inform the moderating and contextual factors affecting demonstration outcomes.

Primary Implementation Question 1: What barriers and facilitators to demonstration participation do beneficiaries experience,[2] and what does this information suggest about the need for refinements to demonstration implementation or design more broadly?

Primary Implementation Question 2: What barriers and facilitators affect state implementation of the demonstration, and what strategies help address barriers?

Primary Implementation Question 3: What barriers and facilitators affect providers in delivering demonstration-covered services, and how do they address these barriers?

[2] "Participation" refers to the mechanism through which beneficiaries come into contact with the demonstration, which could include being screened, referred to services, and receiving services.

4. Hypotheses and research questions related to demonstration outcomes

The following hypotheses and research questions are consistent with CMS's expectations for evaluating the effects of FP demonstrations. Table 2 presents Hypotheses 1 through 3 and corresponding research questions, along with recommended outcome measures, measure stewards, data sources, comparison strategies, and analytic approaches. The outcome measures listed in Table 2 are recommended starting points and include example measures that may only apply to FP demonstrations that cover additional services targeting broader health areas. For example, if a demonstration includes goals related to postpartum depression treatment, the evaluation should include measures related to postpartum depression screening and antidepressant medication management; similarly, if tobacco cessation is a focus of the demonstration, relevant measures should be incorporated.

In addition to the hypotheses and research questions listed below, the state should assess demonstration costs. The document titled *Evaluation Design Technical Assistance Guide for Section 1115 Demonstrations: Assessing Demonstration Costs* provides research questions and analytic methods for cost assessments.

Hypothesis 1: The FP demonstration will maintain or increase beneficiary participation in FP programs, including awareness of and access to FP services.

Primary Research Question 1.1: To what extent are beneficiaries aware of FP services?

Primary Research Question 1.2: How does the demonstration affect participation in FP programs and access to FP services?

Primary Research Question 1.3: How do beneficiaries perceive their experience of care in FP programs?

Hypothesis 2: The ongoing access to FP services through the demonstration will maintain or increase the use of these services among beneficiaries.

Primary Research Question 2.1: How does the demonstration impact use of FP and FP-related services among beneficiaries participating in the demonstration?

Primary Research Question 2.2: How does the demonstration impact the use of other demonstration-covered services?[3]

Hypothesis 3: The ongoing provision of FP services, FP-related services, and other demonstration-covered services will lead to improvements in maternal and infant health outcomes, including a reduction in unintended pregnancies.

Primary Research Question 3.1: How does the provision of FP services, FP-related services, and other demonstration-covered services impact maternal health for eligible beneficiaries?

Primary Research Question 3.2: How does the use of FP services affect the rates of unintended pregnancies?

[3] Some FP demonstrations have covered services beyond the typical family planning and related services, such as tobacco cessation or preconception care.

Primary Research Question 3.3: How do FP services impact infant health outcomes, particularly in terms of birth weight and premature births?

5. Data sources

FP demonstration evaluations can largely rely on data sources described in the document titled Evaluation Design and Reporting for Section 1115 Demonstrations, including survey data and Medicaid administrative data. States may also consider using the following administrative data for evaluations of FP demonstrations:

- Area-level data (for example, from the Agency for Healthcare Research and Quality [AHRQ] Social Determinants of Health Database[4] and the Centers for Disease Control and Prevention [CDC] PLACES database[5]) that describe factors, such as median household income, average levels of education, and employment rates, for census tracts, zip codes, or counties of residence, that might be relevant contextual factors to consider.
- All-Payer Claims Databases (APCDs) provide a comprehensive view of all healthcare services received and can provide insight on health care use patterns before and after the demonstration. They also allow evaluators to construct appropriate comparison groups using data from individuals who transition in and out of Medicaid or are covered by commercial insurance. However, there are limitations to consider. Potential delays in data availability could impede timely analysis; there may be state-specific legal barriers to accessing these databases; and not all states have an APCD.
- The Behavioral Risk Factor Surveillance System (BRFSS) was established by the CDC but is administered by state health departments; it surveys adults ages 18 and older about their health status and is large enough to generate annual state-level estimates. States can choose to use optional BRFSS survey modules and/or state-specific questions that can inform their evaluation designs, although the lead time to add novel questions can be lengthy since the survey is administered once annually. Optional BRFSS modules pertinent to the evaluation of FP demonstrations include Module 1: Prediabetes, Module 2: Diabetes, Module 9: Depression and Anxiety, and Module 16: Clinical Breast Exam, among others. Data for each calendar year is available in the summer of the following year.

6. Methods for testing demonstration hypotheses

The research questions in Section 4 cover demonstration features typically approved by CMS in FP demonstrations. CMS expects the state to adopt the suggested research questions with appropriate modifications for the state-specific demonstration. The research questions address the evaluation hypotheses, in alignment with the anticipated goals of FP demonstrations.

[4] The AHRQ Social Determinants of Health (SDOH) Database available at https://www.ahrq.gov/sdoh/data-analytics/sdoh-data.html includes area-level SDOH variables including economic context (for example, average income, or unemployment rate), education, physical infrastructure (for example, regarding housing, crime, or transportation), and health care context (for example, rates of health insurance).

[5] The CDC PLACES database available at https://www.cdc.gov/places/index.html includes area-level variables at the county, place, census tract, and zip code tabulation area level related to health status, health outcomes, prevention practices, and rates of health risk behaviors (for example, binge drinking, smoking, insufficient sleep, insufficient physical activity).

For each of the outcome measures in Table 2, states should use the most rigorous comparison strategy and associated analytic approach feasible to obtain estimates of causal demonstration impacts. From more to less rigorous, these approaches include the following:

1. A regression model based on a randomized controlled trial comparing beneficiaries participating in the FP demonstration to beneficiaries randomized to a control group that is not participating in the demonstration
2. A difference-in-differences regression model comparing beneficiaries participating in the FP demonstration to similar beneficiaries in a state without an FP demonstration
3. An interrupted time series regression model (if multiple pre-demonstration data points about FP service use are available) or a pre-post comparison (if multiple pre-demonstration data points about FP service use are unavailable or of poor quality)
4. Descriptive trend analyses over the course of the demonstration (if data about FP service use are unavailable or are of poor quality in the pre-demonstration period and in non-demonstration states).
5. Certain analytic approaches that rely on data from the pre-demonstration period or comparison states may encounter challenges in obtaining claims data to evaluate healthcare service utilization. This is because individuals without coverage for FP or FP-related services would not be represented in the relevant Medicaid claims data. Survey or APCD data may help address this data limitation.

Table 1. Suggested measures, data sources, and analytic approaches for research questions related to demonstration implementation

Data sources	Measures and analytic approach[a]
Primary Implementation Question 1: What barriers and facilitators to participation do beneficiaries experience, and what does this information suggest about the need for refinements to demonstration implementation or design more broadly?	
• Medicaid administrative data[b] • State-administered beneficiary survey[c] • Monitoring reports	• Descriptive qualitative analysis of beneficiaries' understanding of and experience with eligibility and demonstration participation processes • Descriptive quantitative analysis of trends in the number of demonstration participants • Descriptive quantitative analysis of beneficiary-reported difficulty finding participating providers accepting new patients
Primary Implementation Question 2: What barriers and facilitators affect state implementation of the demonstration, and what strategies help address barriers?	
• Interviews with state Medicaid agency staff	• Descriptive qualitative analysis of barriers, facilitators, and strategies
Primary Implementation Question 3: What barriers and facilitators affect providers in delivering demonstration-covered services, and how do they address these barriers?	
• Interviews with health plan staff or providers	• Descriptive qualitative analysis of provider/plan barriers, facilitators, and suggestions for improvement

[a] All analytic approaches suggested in Table 1 are descriptive and do not require use of comparison groups. For more details on rigorous qualitative evaluation methods, see Conducting Robust Implementation Research for Section 1115 Demonstration Evaluations, available at https://www.medicaid.gov/sites/default/files/2021-05/Implementation-rsch.pdf.

[b] Administrative data refers to demonstration participation and other program data.

[c] States with smaller beneficiary populations or more limited resources may instead consider conducting beneficiary focus groups or interviews.

Table 2. Suggested measures, data sources, comparison strategies, and analytic approaches

Outcome measure	Measure steward, endorsement	Data source
Hypothesis 1: *The FP demonstrations will maintain or increase beneficiary participation in FP programs, including awareness of and access to FP services.*		
Primary Research Question 1.1: To what extent are beneficiaries aware of FP services?		
Beneficiary awareness of demonstration-covered FP services	None	State-administered beneficiary survey[a]
Primary Research Question 1.2: How does the demonstration impact participation in FP programs and access to FP services?		
• Percentage and total count of beneficiaries facing difficulties in healthcare access, including unmet needs or delayed care. For example: – Service-specific access: Challenges in obtaining contraceptive services (including preferred form of contraception), LARC device removal, screenings, and other covered services. • Scheduling and language services: Issues with securing timely appointments and accessing bilingual or multilingual services.	None	State-administered beneficiary survey[a]
Demonstration participation counts (or counts relative to the eligible population)	None	Medicaid administrative data[b]
Continuity of FP demonstration participation (for example, the percentage of beneficiaries remain continuously enrolled in the demonstration over a 12-month period)	None	
Primary Research Question 1.3: How do beneficiaries perceive their experience of care in FP programs?		
Consumer Assessment of Healthcare Providers and Systems (CAHPS®) Health Plan Survey 5.1H, Adult Version (Medicaid) (CPA-AD)	AHRQ, CMIT #152	State-administered beneficiary survey[a]
Beneficiary satisfaction and experience with care, including, for example, whether provider treats them with respect, whether provider prioritizes patients' values, preferences, and lived experience, whether provider communicates clearly, whether provider explains treatment options	None	
Hypothesis 2: *The ongoing access to FP services through the demonstration will maintain or increase the use of these services among beneficiaries.*		
Primary Research Question 2.1: How does the demonstration impact use of FP and FP-related services among beneficiaries participating in the demonstration?		
Percentage of beneficiaries who utilized at least one allowable service through the FP demonstration	None	Medicaid administrative data[b]
Percentage of female beneficiaries who are provided most effective (LARC devices) and moderately effective (short-acting hormonal methods) contraception (CCW-AD)	OPA, CMIT #1002	
Contraceptive Care Postpartum: • Ages 21–44 (CCP-AD) • Ages 15–20 (CCP-CH)	OPA, CMIT #166	

Outcome measure	Measure steward, endorsement	Data source
Primary Research Question 2.2: How does the demonstration impact the use of other demonstration-covered services?		
Percentage of beneficiaries tested for any sexually transmitted disease (STD)/sexual transmitted infection (STI)	None	Medicaid administrative data[c]
Cervical Cancer Screening (CCS-AD)	NCQA, CMIT #118	
Medical Assistance with Smoking and Tobacco Use Cessation (MSC-AD)[c]	NCQA, CMIT #432	
Prenatal and postpartum care (PPC2-CH and PPC2-AD)[d]	NCQA, CMIT #581	
Screening for Depression and Follow-Up Plan:[e] • Age 18 and Older (CDF-AD) • Ages 12-17 (CDF-CH)	CMS, CMIT #672	
Hypothesis 3: *The ongoing provision of FP services, FP-related services, and other demonstration-covered services will lead to improvements in maternal and infant health outcomes, including a reduction in unintended pregnancies.*		
Primary Research Question 3.1: How does the provision of FP services, FP-related services, and other demonstration-covered services impact maternal health for eligible beneficiaries?		
Severe maternal morbidity (SMM)[f]	CDC/AIM	Medicaid administrative data[c]
Antidepressant Medication Management (AMM-AD)	NCQA, CMIT #63	
Glycemic Status Assessment for Patients with Diabetes (GSD-AD)[g]	NCQA, CMIT #148	
Postpartum Depression Screening and Follow-up: Age 21 and Older (PDS-AD)	NCQA, CMIT #1781	
Perceived health status (physical and emotional)	None	State-administered beneficiary survey[d]
Primary Research Question 3.2: How does the use of FP services affect the rates of unintended pregnancies?		
Interbirth/interpregnancy interval[h] for example, percentage of beneficiaries with very short interpregnancy intervals (e.g., 6 months or less)	None	Medicaid administrative data[c] Vital statistics data
Teen birth rate		
Primary Research Question 3.3: How do FP services impact infant health outcomes, particularly in terms of birth weight and premature births?		
Percentage of infants born preterm	None	Medicaid administrative data[c] Vital statistics data
Live Births Weighing Less Than 2,500 Grams (LBW-CH)[i]	CMS, CMIT #413	
Mother's perceptions of infant health	None	State-administered beneficiary survey[d]

[a] States with smaller beneficiary populations or more limited resources may instead consider conducting beneficiary focus groups or interviews.

[b] Administrative data refers to demonstration participation and other program data and claims or encounter data; some measures may also rely on electronic health record (EHR) data or hybrid data, which refers to data that integrate claims, EHR, and potentially linkages to other administrative data such as vital records. States should review the relevant technical specifications to determine the appropriate data source for each measure used. In general, Medicaid claims data tend to be more standardized across states. If a measure requires data beyond claims, such as EHR or linked data, evaluating changes in such a measure using an out-of-state comparison group can be more difficult.

[c] Medical Assistance with Smoking and Tobacco Use Cessation (MSC-AD) refers to medical assistance provided to adult beneficiaries who use tobacco and receive counseling, medication, or other support for smoking and tobacco use cessation. This measure applies only if the service is covered under the demonstration.

[d] Factors such as participation continuity, late participation in the demonstration during pregnancy, and the use of bundled payment codes are important to consider when assessing the feasibility of measuring the timeliness of prenatal care.

[e] Adjustment of the specification is needed for the postpartum population. This measure typically excludes individuals with depression and/or bipolar disorder. However, the standard of care is that a screening should occur in the postpartum period regardless of prior diagnosis.

[f] SMM can be measured separately during the delivery hospitalization and six weeks postpartum. SMM is typically measured during the delivery hospitalization but may occur during the immediate postpartum period.

[g] This measure may require hybrid reporting or supplemental clinical data if administrative claims lack lab result values. Acceptable numerator value includes HbA1c/GMI lab results (e.g., LOINC with values) and, when available, CPT II codes.

[h] States may consider measuring the healthy timing and spacing of pregnancies through interpregnancy intervals. This could be achieved by analyzing Medicaid claims data, or by linking Medicaid claims with birth records to improve accuracy. For the calculation of interpregnancy intervals, a measurement period of at least 12 months is preferred. An alternative is to measure interbirth intervals, which may be calculated using only Medicaid data. For the calculation of interbirth intervals, a measurement period of at least 21 months (12-month interpregnancy interval + 9-month gestation period) is preferred.

[i] To create a clearer distinction between instances of low birth weight and preterm birth, states may also consider measuring the percentage of infants born at term with low birth weight.

AHRQ = Agency for Healthcare Research and Quality; CMIT = Centers for Medicare & Medicaid Services Measures Inventory Tool; FP = family planning; NCQA = National Committee for Quality Assurance; OPA = Office of Population Affairs; CMS = Centers for Medicare & Medicaid Services; CDC = Centers for Disease Control and Prevention; AIM = Alliance for Innovation on Maternal Health.