Evaluation Technical Assistance Guide for Section 1115 Demonstrations: Assessing Demonstration Costs

This document provides technical assistance for assessing costs in conjunction with section 1115 demonstration goals and outcomes as part of demonstration evaluations. Estimating impacts on demonstration costs through evaluation is distinct from budget neutrality. **This guide is not applicable for assessments of budget neutrality.**

This document includes an overview of cost assessments (Section 1) and provides specific technical assistance on the research questions (Section 2 and Table 1), measures and data sources (Section 3), and methods (Section 4) for assessing demonstration cost impacts.

1. Assessing section 1115 demonstration costs

Evaluations of section 1115 demonstrations are expected to include an assessment of the (1) administrative costs of demonstration implementation and operation and (2) costs of healthcare service expenditures associated with the demonstration. Demonstrations may have additional policy-specific requirements in the STCs for assessing costs, such as examining uncompensated care costs for providers.

Demonstration cost assessments enable comparing costs to a comparison group not participating in the demonstration, exploring how costs vary among demonstration participants or over time, and explaining why costs decrease or increase and the degree to which such change is due to demonstration policies. These detailed insights are all valuable for shaping future Medicaid policy, making comprehensive cost assessment an essential part of demonstration evaluations.

Cost assessments must be designed considering the specific populations covered and benefits provided under the demonstration. States and their independent evaluators should measure cost impacts both at the demonstration level (for all policies contained in the demonstration, in total) and, to the extent possible, separately for each demonstration policy. For example, states with demonstrations that allow them to test flexibilities to improve the continuum of care for beneficiaries with substance use disorders (SUDs) and address beneficiaries' family planning needs should analyze the cost impacts of these two policies separately.

States may also consider monetary and non-monetary benefits in their comprehensive cost assessments, for example, by conducting economic evaluation analyses such as cost-benefit, cost effectiveness (or cost-utility), or return on investment analyses. Economic evaluation analyses provide a more complete assessment of the demonstration's value by considering information that goes beyond direct monetary costs of the demonstration. Economic evaluations provide a framework for comparing different healthcare options to determine which provides the best value, considering both costs and resulting health outcomes. For example, a cost-benefit analysis of a demonstration that provides home- and community-based services could consider improvements in beneficiaries' quality of life or work capacity, which would generate additional tax revenue for the state. Methods to assign a dollar value to non-monetary benefits (see Section 4.d) allow states to compare benefits to demonstration costs.

In summary, estimating the cost impacts of a section 1115 demonstration requires a clear understanding

of demonstration objectives and mechanisms for achieving those objectives. Cost assessments should not stand alone but rather should be informed by and interpreted in light of other evaluation findings on progress towards stated demonstration goals. Economic evaluations can be a powerful tool for further assessing the value of the demonstration.

2. Research questions for cost assessments

The overarching research questions are how costs changed due to the demonstration and why. States should assess demonstration impacts on costs associated with (1) implementation and ongoing operations for all demonstration policies and (2) the outcomes of those policies. Cost analyses should cover administrative costs, healthcare service spending, and other costs (Section 3 describes these costs in detail). States should analyze costs in total for each demonstration policy and by type of service to better understand cost drivers (see Section 4.a for details). States should also compare demonstration cost impacts to the benefits of the demonstration. Thus, suggested research questions ask about a range of cost impacts:[1]

Primary Research Question 1: What are the administrative costs incurred by the state to implement and operate the demonstration (policy) and how do they change during the demonstration approval period?

Primary Research Question 2: What are the short- and long-term effects of the demonstration (policy) on Medicaid healthcare service expenditures overall?

Subsidiary Research Question 2a: What are the short- and long-term effects of the demonstration (policy) on category-specific healthcare service expenditures (for example, inpatient hospitalization, outpatient visits, etc.)?

Primary Research Question 3: How do the demonstration costs (administrative costs and healthcare service expenditures) compare to demonstration benefits for demonstration participants and the Medicaid program?

Although cost assessments are largely quantitative in nature, states should consider adding a qualitative component that may be tied with the implementation evaluation of the demonstration overall and can provide context and explanation for cost impacts and potential cost drivers. The following suggested research question asks why costs have changed:

Primary Research Question 4: What specific policies implemented during the demonstration, populations affected by the demonstration, or changes exogenous to the demonstration do state officials perceive as the main drivers for the cost impacts and why?

3. Cost measures and their data sources

Administrative costs. States and their evaluators should compute administrative costs associated with demonstration startup as well as ongoing administrative costs of demonstration operations. Specific administrative costs to examine include the cost of (1) contracts or contract amendments to implement

[1] In contrast to other evaluation technical assistance guides, this guide does not include hypotheses in addition to research questions. States should formulate their own demonstration-specific hypotheses about expected changes in costs, which could include that costs will not change.

demonstration policies, as well as those for monitoring and evaluation, and (2) staff time equivalents required to implement, administer, and communicate with beneficiaries about demonstration policies. Estimates of administrative costs should include Medicaid agency staff time for those hired to support the demonstration, as well as time redirected to the demonstration from other Medicaid operations in whole or in part.

States should take a lead role in gathering information on these costs and should facilitate evaluators' access to this information. Qualitative interviews may also help systematically gather information on administrative costs, particularly to understand the allocation of state staff time required to launch and then maintain demonstration operations. Depending on the role of managed care entities in implementing the demonstration policies, states may also need to include managed care administrative costs, gathering information through interviews and potentially through secondary data sources.

Healthcare service expenditures. States should also seek to measure changes in the costs of providing healthcare services to Medicaid beneficiaries participating in the demonstration (in the aggregate and per member per month [PMPM]). Examining changes in both total service expenditures and PMPM expenditures is important, as the two measures may move in opposite directions depending on the demonstration's effects on enrollment. For example, beneficiaries with chronic conditions or higher healthcare needs may be most likely to maintain Medicaid coverage over time, and states may observe that PMPM costs rise even if total enrollment and therefore total service costs decrease. On the other hand, one of the goals of SUD and serious mental illness (SMI)/serious emotional disturbance (SED) demonstrations is to identify people with SUD and SMI/SED, which may lead to increased total service expenditures and PMPM expenditures in the short run.

The main data sources for service expenditures are claims and encounter data from the state's own Medicaid Management Information System (MMIS) and, if the evaluation involves an out-of-state comparison group, the Transformed Medicaid Statistical Information System (T-MSIS). States should ensure that evaluators have access to necessary administrative data on service expenditures.

Other costs. Depending on the demonstration policy being evaluated, states may assess impacts on costs beyond administrative costs and healthcare service expenditures that can be attributed to specific beneficiaries. Examples of these costs may include up-front payments to providers to build the necessary support for demonstration operations, or uncompensated care costs to providers.

If analyzing uncompensated care costs is appropriate given a demonstration's policies, there are additional data sources to consider. These include:

- State-specific provider surveys, which could provide information about uncompensated care costs incurred by hospital and nonhospital providers, such as federally qualified health centers.
- The Healthcare Cost and Utilization Project, State Inpatient Databases (HCUP-SID), which contain information on the source of payment (including "no charge" [charity care] and "self-pay") for hospital inpatient stays by state. [2]

[2] https://www.hcup-us.ahrq.gov/sidoverview.jsp

- The Healthcare Cost Report Information System (HCRIS), which contains relatively current and comprehensive data for Medicare-participating institutional providers.[3] HCRIS data are captured through Medicare costs reports and include charity care (uninsured and insured, separately), non-Medicare and non-reimbursable Medicare bad debts, indigent care costs (for patients covered by state or local government programs), and Medicaid shortfalls (after Medicaid supplemental payments).

Benefits. Which data sources a state uses to measure benefits depend on demonstration goals and the nature of outcomes that the state considers benefits in the context of an economic evaluation. For example, SUD demonstrations have the goal to reduce overdose mortality, so states could use vital statistics data to measure changes in mortality among demonstration participants as a key demonstration benefit. In general, states already collect data and measure outcomes related to demonstration goals, so calculating benefits does not require additional data collection. When states assess improvements in quality of life or self-reported health status as a demonstration benefit, they can rely on beneficiary surveys to measure these outcomes. To measure demonstration benefits that accrue to Medicaid or other state entities, the state can rely on administrative records and interviews (see Administrative costs).

4. Methods for cost assessments

This section includes methodological considerations on how to conduct cost analyses on different levels (in total, by service type, and for cost drivers), units and time periods of analysis, estimation methods that can support causal inference, and approaches to economic evaluations.

a. Levels of cost analyses

Depending on the demonstration policy being evaluated, states should conduct cost analyses at the following levels:

1. **Total costs PMPM.** For all demonstration policies, states should calculate how much their Medicaid program spent on services for demonstration participants PMPM. This calculation should be based on claims data (inpatient, outpatient, pharmacy, long-term care, and capitated payments to managed care organizations). States should calculate total costs to the state as well as total federal costs.

   Depending on the demonstration policy and state data systems, total costs PMPM may require incorporating data outside the Medicaid claims system. For example, for SUD and SMI/SED demonstrations, states should include costs from Institutions for Mental Diseases (IMD), for which some states have historically collected data outside the standard claims system.

2. **Costs related to demonstration-specific services.** For some demonstration policies, it is informative to split costs into those directly related to the services provided under the demonstration and other Medicaid costs, both of which could be affected by the demonstration. This cost assessment allows states to distinguish between changes in costs they expect to be directly affected by the demonstration policy as opposed to cost impacts that may be indirectly attributed to the demonstration.

[3] https://www.cms.gov/Research-Statistics-Data-and-Systems/Downloadable-Public-Use-Files/Cost-Reports/

Which costs states decide to analyze separately depends on the demonstration type. For example, the total costs of SUD and SMI/SED demonstrations can be partitioned into IMD costs, SUD or SMI/SED treatment costs not related to IMDs, and non-SUD or non-SMI/SED related service costs.

To conduct cost analyses by service type, states that deliver a significant amount of health care through capitated arrangements would need to take one of several potential approaches. First, states may assign costs to encounter data claims based on their fee-for-service (FFS) fee schedule. Second, states may impute costs based on the relative FFS cost of different services, scaled by actual capitation rates (that is, shadow pricing). Third, states may use provider paid amounts recorded on Medicaid encounters.

3. **Cost drivers.** States should analyze cost drivers for the demonstration population to answer the question of why costs changed during the demonstration period. This aspect of the cost assessment is vital because it goes beyond merely showing that costs increased or decreased. Cost driver analyses should be informed by other components of the evaluation, including hypotheses about demonstration impacts on healthcare service use. They could also benefit from qualitative implementation analyses (for example, interviews with state staff). These qualitative assessments may be particularly valuable when a demonstration has concurrent policies and their individual impact on costs cannot be easily quantified.

   Whether states classify a particular type of healthcare service as a cost driver of interest depends on the demonstration policy being evaluated. For example, for beneficiaries with SMI/SED or SUD, states could separately distinguish emergency department (ED)-related outpatient costs from other outpatient costs, given that ED services are particularly high cost and represent an important opportunity for cost savings that could be achieved with better access to SMI/SED or SUD treatment services. In this case, the cost driver analysis ties in with other evaluation activities because states evaluating their SMI/SED or SUD demonstrations already estimate impacts on ED use among demonstration participants.

b. Unit of analysis and time periods

States should conduct cost analyses with the beneficiary-month as the unit of analysis. The demonstration population depends on the demonstration policy; for example, it consists of beneficiaries with SUD diagnosis or treatment for SUD demonstrations. For most purposes, states should use a repeated cross-sectional approach that does not require minimum enrollment durations for beneficiaries to be included in the analysis. For example, beneficiaries would be included in the analysis during the first month in which a relevant SUD diagnosis or treatment claim was observed and for up to 11 additional months that did not include a relevant diagnosis or treatment claims. Once an individual has a period of one year with no relevant diagnosis or treatment claims, that beneficiary can be excluded from further analyses, unless and until they have a subsequent relevant diagnosis or treatment claim. For other demonstration policies, states should align the selection of beneficiaries and time periods during which costs are assessed with the corresponding demonstration requirements and regulations.

An alternative would be to conduct a cohort analysis that followed over time a defined group of Medicaid beneficiaries with the relevant diagnosis or treatment claims or other characteristics that make them eligible to participate in the demonstration. However, given that beneficiaries do not always remain continuously enrolled for long periods, results from a cohort analysis might not accurately reflect full cost

impacts. Including more beneficiaries with relevant characteristics in the cost analysis, even if their enrollment duration is short, may have greater policy relevance.

In analyzing cost trends over time, states should consider using at least two years of data before implementation of the demonstration. Using only one year of pre-implementation data to assess costs may not provide sufficient information about pre-demonstration period cost trends.

c. Methods to estimate changes in costs to the Medicaid program

States should use the most rigorous method that is feasible to obtain causal estimates of demonstration cost impacts for all types of costs and levels of cost analyses. Three potential analytic approaches include, in order of preference: (1) difference-in-differences analysis, which uses a demonstration and comparison group and data before and after implementation of the demonstration; (2) interrupted time series analysis without a comparison group, which allows for estimating different linear effects during the pre- and post-demonstration periods when a comparison group is not available; and (3) a pre-post analysis, which includes a single pre-demonstration period and a single post-demonstration period when a comparison group is not available and it is not feasible to conduct an interrupted time series analysis for lack of sufficient data. States should make their best efforts to incorporate comparison populations in their cost analyses, except for the analysis of administrative costs. The document *Evaluation Design and Reporting for Section 1115 Demonstrations* provides a more detailed discussion of these methods.[4]

All modeling approaches should be at the beneficiary-month level and should include covariates to control for demographic characteristics, including age, sex, and dual eligibility status. Additional covariates could include clinical characteristics (behavioral or physical health comorbidities), household income, and delivery system (managed care plan or FFS).

For regression modeling purposes, it is more appropriate to use the logarithm of costs rather than untransformed costs because costs are typically not normally distributed. Many beneficiary-months will have low or zero health care spending, while other months could have very large values, for example, due to an inpatient stay. States may also wish to conduct a two-part model that includes zero costs (logit model) and nonzero costs (generalized linear model [GLM]), because there may be many beneficiary-months with zero costs.

For reporting purposes, states should report marginal effects and standard errors to assess statistically significant changes in costs. Estimated regression coefficients can also be used to generate predicted or adjusted monthly average costs, controlling for observed demographic and other factors. Both unadjusted and adjusted means for demonstration and comparison groups could be plotted to show trends visually, which would make the regression results more easily interpretable.

d. Economic evaluation methods

There are different types of economic evaluation methods, all of which can help states enhance the evaluation of section 1115 demonstrations by providing a comprehensive view of demonstration value, considering both costs and realized outcomes. Specifically, economic evaluation methods include cost-

[4] This resource is 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.

effectiveness analysis (or cost-utility analysis), cost-benefit analysis, and return on investment analysis.[5] Policymakers can use the information obtained from economic evaluations to assess the value of specific demonstration policies and make decisions about extending or changing demonstrations.

An economic evaluation of a section 1115 demonstration shows how demonstration costs compare to demonstration benefits. The main challenge when conducting an economic evaluation is the decision of which costs and benefits to include. States should make this decision with the demonstration goals and context in mind. Costs generally include administrative costs, healthcare service expenditures, and potentially other costs (Section 3). The benefits to include in an economic evaluation depend on the demonstration type and should derive from the desired outcomes of the demonstration. For example, for SUD demonstrations, benefits may include beneficiaries' increased life expectancy (through reduced overdose events) and quality of life (due to SUD treatment). They could also include reduced losses due to falling crime in affected communities. The latter example shows that benefits can include those that accrue to people who are not demonstration participants. States should use their best judgement to decide which benefits to include. Importantly, all costs and benefits to be considered in an economic evaluation must be prespecified in the evaluation design phase. This prevents states from selecting favorable cost and benefit items when conducting the evaluation.

In a **cost-effectiveness analysis** (sometimes called a cost-utility analysis), states compare the relative costs and outcomes of two or more alternatives, without monetizing the outcomes. In the context of section 1115 demonstrations, the two alternatives could be a demonstration policy and the state's Medicaid program without the demonstration. For example, a state could conclude that a family planning demonstration reduces unintended pregnancies by 3 percent and increases costs by 1 percent, relative to baseline Medicaid expenditures, and use this information to assess the value of the demonstration. A common practice when valuing outcomes related to mortality or quality of life is to use quality-adjusted life years (QALYs) or disability-adjusted life years (DALYs).[6] States can use an estimated value of QALYs or DALYs in combination with estimates on how a demonstration affects mortality, disability, and quality of life to calculate the monetary benefits of a demonstration. Cost-effectiveness analysis where benefits are measured in QALYs or DALYs is also known as **cost-utility analysis.**

To conduct a **cost-benefit analysis**, states must assign a dollar value to demonstration impacts or benefits. This allows direct comparison of the outcomes of a demonstration to demonstration costs. How states can do this in practice depends on the type of outcome measures states use, but in general, it involves multiplying an estimated outcome effect size by the dollar value that the state assigns to the outcome being assessed. This dollar value may come from estimates available in the broader literature or from a state's direct experience.

[5] For an overview of economic evaluation methods, see: Turner, Hugo C., Rachel A. Archer, Laura E. Downey, et al. "An Introduction to the Main Types of Economic Evaluations Used for Informing Priority Setting and Resource Allocation in Healthcare: Key Features, Uses, and Limitations." *Frontiers in Public Health*, vol. 9, 2021, article 722927.

[6] For an overview of QALYs and DALYs, see Neumann, Peter J., and David D. Kim. "Cost-effectiveness Thresholds Used by Study Authors, 1990-2021." *JAMA*, vol. 329, no. 15, 2023, pp. 1312–1314.

Finally, states can use economic evaluation to calculate the **return on investment (ROI)** of the demonstration. The ROI is a simple measure that relates demonstration costs and benefits (measured in dollars) and is defined as:

$$\text{ROI} = \frac{\text{total benefits - costs}}{\text{total costs}}$$

The ROI shows what percentage of the total costs of the demonstration are returned as net benefits. Calculating the ROI of a demonstration allows states to easily communicate to CMS and other stakeholders about the estimated value of the demonstration, after accounting for how much the demonstration costs.

Table 1: Suggested comparison strategies, measures, data sources, and analytic approaches

Note: CMS expects that states will work with their evaluators and demonstration partners to choose among and adapt suggested evaluation approaches based on comparison group opportunities and data availability. Suggested approaches to answering quantitative research questions emphasize experimental and quasi-experimental approaches, like difference-in-differences regression models, because these research questions seek to establish causal inference between demonstration policies and costs.

Outcome measures	Data sources	Comparison strategy and analytic approach
Primary Research Question 1: What are the administrative costs to implement and operate the demonstration (policy) and how do they change during the demonstration approval period?		
Administrative cost of demonstration **implementation**, including cost of contracts or contract amendments and staff time equivalents required to establish demonstration policies, typically incurred in the years prior to and including the initial year of the demonstration	State and managed care administrative records Interviews with state agency staff and partner organizations	Descriptive quantitative analysis of administrative costs
Administrative cost of ongoing demonstration **operation**, including the recurring cost of contracts or contract amendments and staff time equivalents required to administer demonstration policies		
Administrative costs to state agencies partnering with Medicaid to implement and operate the demonstration		

Outcome measures	Data sources	Comparison strategy and analytic approach
Primary Research Question 2: What are the short- and long-term effects of the demonstration (policy) on Medicaid healthcare service expenditures overall?		
Total healthcare service expenditures for demonstration population	State administrative data on beneficiary-level expenditures and enrollment in the demonstration	Descriptive quantitative analysis comparing to (1) similar beneficiaries in other states that do not operate a similar demonstration or (2) within-state beneficiaries not subject to demonstration based on implementation strategy (staged by geographic area, age group, or other beneficiary characteristic) and/or eligibility criteria, with pre-period trend analysis
Change in PMPM health service expenditures		Differences-in-differences analysis comparing to (1) similar beneficiaries in other states that do not operate a similar demonstration or (2) within-state beneficiaries not subject to demonstration based on implementation strategy (staged by geographic area, age group, or other beneficiary characteristic) and/or eligibility criteria[a]
Subsidiary Research Question 2a: What are the short- and long-term effects of the demonstration (policy) on category-specific healthcare service expenditures (for example, inpatient hospitalization, outpatient visits, etc.)?		
Total expenditures for the healthcare service of interest for demonstration population	State administrative data on beneficiary-level expenditures and enrollment in the demonstration	Descriptive analysis comparing to (1) similar beneficiaries in other states that do not operate a similar demonstration or (2) within-state beneficiaries not subject to demonstration based on implementation strategy (staged by geographic area, age group, or other beneficiary characteristic) and/or eligibility criteria, with pre-period trend analysis
Change in PMPM expenditures for the healthcare service of interest		Differences-in-differences analysis comparing to (1) similar beneficiaries in other states that do not operate a similar demonstration or (2) within-state beneficiaries not subject to demonstration based on implementation strategy (staged by geographic area, age group, or other beneficiary characteristic) and/or eligibility criteria[a]
Primary Research Question 3: How do the demonstration costs (administrative costs and healthcare service expenditures) compare to demonstration benefits for demonstration participants and the Medicaid program?		
Total demonstration benefits (potentially converted into dollars) for beneficiaries, Medicaid, and the state Total demonstration costs (administrative costs and healthcare service expenditures)	State and managed care administrative records Interviews with state agency staff and partner organizations Evaluation evidence on demonstration outcomes Estimates from literature to monetize outcomes	Descriptive quantitative analysis of demonstration costs and benefits in a cost-effectiveness or cost-benefit analysis without a comparison group
Return on investment (ROI)		Descriptive quantitative analysis of demonstration costs and benefits in an ROI analysis without a comparison group

Outcome measures	Data sources	Comparison strategy and analytic approach
Primary Research Question 4: What specific policies implemented during the demonstration, populations affected by the demonstration, or changes exogenous to the demonstration do state officials perceive as the main drivers for the cost impacts and why?		
Change in PMPM healthcare service expenditures, for specific services, demonstration populations, or demonstration policies	State administrative data on beneficiary-level expenditures and enrollment in the demonstration Interviews with state agency and partner organizations staff	Comparative analysis of PMPM expenditures across services, populations, or policies, synthesized with qualitative analysis of interview notes
Policies, populations, and changes exogenous to the demonstration mentioned by interviewees as cost drivers		Descriptive qualitative analysis of interview notes

[a] If no baseline (pre-demonstration) data are available, for example because demonstration implementation coincides with a coverage expansion to the population of interest, a difference-in-differences model is not possible. However, if the state stages (rolls out) implementation based on a continuous beneficiary characteristic such as age or income, or varies policy according to a continuous beneficiary characteristic, a regression discontinuity design may be used. See the white paper "Regression Discontinuity Designs in the Evaluation of Section 1115 Demonstrations" available at https://www.medicaid.gov/medicaid/section-1115-demo/downloads/evaluation-reports/regression-discontinuity-designs.pdf for details. If no comparison group is feasible, but baseline data are available, an interrupted time series or pre-post analysis may be used.

PMPM = per member per month.