APPENDIX C: APPROACHES TO ANALYZING COSTS ASSOCIATED WITH SECTION 1115 DEMONSTRATIONS FOR BENEFICIARIES WITH SERIOUS MENTAL ILLNESS/SERIOUS EMOTIONAL DISTURBANCE OR SUBSTANCE USE DISORDERS

This appendix is a supplement to the evaluation design guidance for Section 1115 demonstrations that aim to ensure and improve a continuum of care for (1) beneficiaries with serious mental illness (SMI) or serious emotional disturbance (SED) or (2) beneficiaries with substance use disorders (SUD). This appendix details methods to calculate changes in total costs and examine cost drivers within the Medicaid program. The Centers for Medicare & Medicaid Services (CMS) requires states to submit cost analyses as part of both the interim and summative evaluation reports for all SMI/SED or SUD Section 1115 demonstrations.

Cost analyses assess whether the demonstrations result in higher, lower, or neutral health care spending. The direction and degree of spending change are theoretically uncertain—that is, states might not have a clear expectation that costs will change in one direction or another. Although the demonstrations aim to increase the availability and use of SMI/SED or SUD treatment services, which could increase health care spending, the use of those services could be balanced by reduced use of costly services such as inpatient stays and emergency department (ED) visits.

This document provides an approach that state evaluation teams should use to generate several interrelated cost outcomes for each demonstration. Ideally, cost outcome measures will be expressed in terms of dollars per beneficiary per month (PBPM). This guidance focuses primarily on Medicaid fee-for-service (FFS) but notes where modifications may be needed if states use capitated Medicaid managed care.

1. Levels of cost analyses

CMS suggests conducting three levels of cost analyses. Proposed data sources for each type of costs needed for each of the following levels of analysis are presented in Table C.1:

a. Total costs PBPM. This calculation should be based on claims data (inpatient [IP], outpatient [OT], pharmacy [RX], long-term care [LT], and capitated payments to managed care organizations); costs from Institutions for Mental Diseases (IMD), which some states have historically collected outside the standard claims system; and administrative costs.[1] States should calculate total costs as well as total federal costs.

Calculating total costs is likely to be the most feasible and accurate approach for states in which capitated contracts cover most services used by Medicaid beneficiaries. To conduct more granular cost analyses, such as those recommended below, states with significant capitated arrangements would need to take one of several potential

[1] Total costs should also include administrative costs associated with the demonstration, such as hiring staff or adding a vendor to manage an activity uniquely associated with the demonstration. These costs would be included only in the total cost level of analysis because they cannot be parsed into cost drivers, such as outpatient or inpatient care. However, we recommend analyses be conducted at the beneficiary level, while administrative costs may be compiled in aggregate. One work-around would be to create a per capita allocation of administrative costs.

approaches. First, states may assign costs to encounter data claims based on their 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. Alternatively, such states could conduct utilization analyses without ascribing prices.

b. Costs related to diagnosis and treatment of SMI/SED or SUD. This level of analysis identifies cost drivers by splitting out costs associated with either an SMI/SED or SUD diagnosis and/or procedure, depending on the demonstration, as identified by using IMD and claims data. Costs in this level of analysis can be categorized as follows:

- **For SMI/SED demonstrations:** Total costs = SMI/SED-IMD costs + other SMI/SED costs + non-SMI/SED costs
- **For SUD demonstrations:** Total costs = SUD-IMD costs + other SUD costs + non-SUD costs

c. Source of treatment cost drivers. This level of analysis identifies cost drivers for the target population—beneficiaries with SMI/SED or beneficiaries with SUD, depending upon the demonstration—by splitting out costs associated with different types of care by using claims data. CMS suggests separately distinguishing 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. Costs in this level of analysis can be categorized as follows:

- Total costs = IP + non-ED OT + ED OT + RX + LT

Table C.1. Types of costs and proposed data sources

Level of analysis	Type of costs	Description/data source
Total costs	Total costs	Sum of IP, OT, RX, LT, IMD costs, administrative costs
	Total federal costs	Total Medicaid costs * Federal medical assistance percentage (FMAP) for the state
Cost related to diagnosis and treatment: SMI/SED demonstrations[a]	SMI/SED-IMD costs	IMD costs reported by states with SMI/SED diagnosis and/or procedure codes
	Other SMI/SED costs	Costs with SMI/SED diagnosis and/or procedure code from T-MSIS IP, OT, RX, LT files or equivalent
	Non-SMI/SED costs	Costs without SMI/SED diagnosis and/or procedure code from T-MSIS IP, OT, RX, LT files or equivalent
Costs related to diagnosis and treatment: SUD demonstrations[a]	SUD-IMD costs	IMD costs reported by states with SUD diagnosis and/or procedure codes
	Other SUD costs	Costs with SUD diagnosis and/or procedure code from T-MSIS OT, IT, RX, LT files or equivalent
	Non-SUD costs	Costs without SUD diagnosis and/or procedure code from T-MSIS IP, OT, RX, LT files or equivalent
Source of treatment cost drivers for beneficiaries in the target population[a]	Outpatient costs, non-ED	T-MSIS OT file or equivalent
	Outpatient costs, ED	Primarily T-MSIS OT files or equivalent, may need IP files in some states
	Inpatient costs	T-MSIS IP file or equivalent
	Pharmacy costs	T-MSIS RX file or equivalent
	Long-term care costs	T-MSIS LT file or equivalent

Note: States may have underlying data configured in a variety of ways. Although this document references T-MSIS to help clarify the type of data states could use, the data may actually be housed in different ways at the state level. In particular, behavioral health services may be tracked in different data systems, even if they are ultimately fed into T-MSIS OT files.

[a]Summed costs in this level of analysis may not equal the total costs level of analysis, particularly in states that use capitated managed care or states that have added in administrative cost data.

ED = emergency department; IMD = Institutions for Mental Diseases; IP = inpatient; LT = long-term care; OT = outpatient; RX = pharmacy; SMI/SED = serious mental illness/serious emotional disturbance; SUD = substance use disorder; T-MSIS = Transformed Medicaid Statistical Information System.

2. Unit of analysis and time periods

States should conduct the analysis with the beneficiary-month as the unit of analysis, focusing on beneficiaries in the target population in claims data. Depending upon the demonstration type, the target population is beneficiaries with SMI/SED diagnosis or treatment or beneficiaries with SUD diagnosis or treatment. States should use a repeated cross-sectional approach that does not require minimum enrollment durations for beneficiaries to be included in the analysis. Beneficiaries would be included in the analysis during the first month in which a relevant SMI/SED or 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. Setting the inclusion criteria this way results in an analysis that represents the costs of serving individuals in the target population with active treatment needs.

An alternative to analyzing costs for each year would be to conduct a cohort analysis that followed a defined group of Medicaid beneficiaries with the relevant diagnosis or treatment claims over time. A cohort analysis may be preferable when assessing clinical outcomes. However, given that the Medicaid population is known to have high levels of eligibility churn, which are likely exacerbated among individuals with SMI/SED or SUD, results from a cohort analysis might not adequately reflect the full cost picture that will appear evident to the states. Including more beneficiaries with relevant diagnosis or treatment codes in the cost analysis, even if their enrollment duration is short, may have greater policy relevance.

States should consider using two pre-demonstration years of data when analyzing costs. This differs from the guidance on assessing non-cost outcomes in other sections of the evaluation design guidance for Section 1115 SMI/SED or SUD demonstrations. Using only one year of data to assess costs may not provide sufficient information about pre-period cost trends.

3. Methodology for creating Medicaid cost variables and analyzing costs to the Medicaid program

The following instructions provide a step-by-step approach for compiling and organizing data from the relevant cost data sources and for generating descriptive statistics. These instructions assume the availability of a comparison group and pre-demonstration period data and provide the necessary foundation for regression models described in the next section. States should work with their evaluator to assess the feasibility of potential comparison groups. CMS recognizes that identifying an in-state comparison group may not be feasible for a statewide demonstration; therefore, states should explore the possibility of obtaining administrative data for other states.

a. Identify Medicaid beneficiaries with SMI/SED or with SUD

Using files obtained from the state Medicaid data warehouse, including inpatient, outpatient, pharmacy, and long-term care claims, as well as any IMD data supplied to the state (or state evaluator), identify beneficiaries in the target population with the relevant diagnosis or treatment codes during each of the pre- and post-demonstration periods. Link beneficiaries with the relevant diagnosis and/or treatment during the specified time periods to Medicaid eligibility data to identify the months that a beneficiary was enrolled in Medicaid and to obtain demographic variables. The analysis should include the first month in which a relevant diagnosis or treatment claim either for SMI/SED or for SUD—depending upon the demonstration—was observed for the beneficiary and up to 11 additional months that did not include relevant claims if the beneficiary remained enrolled in Medicaid. If a beneficiary has additional claims with the relevant diagnosis or treatment code values, the observation period included in the analysis should be extended to include up to 11 additional months following the subsequent claims if the beneficiary remained enrolled in Medicaid.

b. Create the analytic file

Organize the data to create a file with an observation for each month in which a beneficiary was Medicaid-eligible, either once their claims history meets the definition of SMI/SED or once

an SUD-related claim is observed during the analysis period. For each month in which a beneficiary is enrolled, the data file should contain an observation with the beneficiary's Medicaid costs in that month, using the 10 variables specified in Table C.1 and demographic characteristics merged from the eligibility data. Include an indicator to be used in all regression modeling analyses (difference-in-differences, interrupted time series, or pre-post) equal to 1 for months on or after the start date of the demonstration and equal to 0 for the pre-demonstration months. Include another indicator to be used in all regression modeling analyses equal to 1 if the individual is in the treatment group that month.

c. Calculate and trend average monthly spending

From the individual month-level data, calculate average costs. Table C.2 presents a sample template for reporting average costs for SMI/SED demonstrations. States should construct a similar table for reporting average costs for SUD demonstrations. Means compiled in Table C.2 could be plotted to show trends visually and to verify that month-to-month variation is within expectations and does not indicate an underlying data error. States may also wish to conduct quarterly spending analyses to smooth out monthly variation in costs.

Table C.2. Template for reporting unadjusted means of Medicaid cost estimates for individuals participating in the Section 1115 demonstration, by type of cost, period, and treatment/comparison group (SMI/SED demonstration example)

		Pre-demonstration		Post-demonstration	
	Type of cost	Month 1	Month 2[a]	Month 1	Month 2[a]
Treatment group costs					
Total costs	Total costs Total federal costs				
SMI/SED cost drivers	SMI/SED-IMD costs Other SMI/SED costs Non-SMI/SED costs				
Type or source of care cost drivers	Outpatient costs, non-ED Outpatient costs, ED Inpatient costs Pharmacy costs Long-term care costs				
Comparison group costs					
Total costs	Total costs Total federal costs				
SMI/SED cost drivers	SMI/SED-IMD costs Other SMI/SED costs Non-SMI/SED costs				
Type or source of care cost drivers	Outpatient costs, non-ED Outpatient costs, ED Inpatient costs Pharmacy costs Long-term care costs				

[a]Table includes two pre-demonstration and post-demonstration months for illustrative purposes only. States should include at least one year of pre-demonstration and all post-demonstration data.

ED = emergency department; IMD = Institutions for Mental Diseases; SMI/SED = serious mental illness/serious emotional disturbance.

4. Methodology for conducting analyses to assess changes in costs to the Medicaid program

This section outlines three potential analytic approaches, in order of preference: (1) difference-in-differences analysis, which uses a treatment and comparison group; (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 develop evaluation designs that incorporate comparison populations, including for the required cost analyses.

All modeling approaches should include covariates to control for demographic characteristics, including age, race, gender, and dual eligibility status. Additional optional covariates include clinical characteristics (behavioral or physical health comorbidities), household income, and delivery system (managed care plan or FFS).

Note that for regression modeling purposes, it is more appropriate to use log costs rather than untransformed costs because costs are typically not normally distributed. Many beneficiary-months will have $0 health care spending, while other months could have very large values. States may also wish to conduct a two-part model that includes zero costs (logit model) and nonzero costs (generalized linear model [GLM]),[2] because there may be many beneficiary-months that do not include any 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 observable demographic and other factors. The output table used for such an analysis would be organized in the same way as Table C.2, but it would be more appropriate to directly compare treatment and comparison group values and to examine their trends over time. As with the unadjusted means, the adjusted means could be plotted to show trends visually, which would make the regression results more easily interpretable.

a. Preferred approach: Conduct difference-in-differences analysis

Difference-in-differences analysis compares trends for those affected by the demonstration with trends for beneficiaries not affected by the demonstration. Comparison beneficiaries could be either in the demonstration state (and thus not affected due to the demonstration's geographic

[2] A recommended approach for conducting a two-part model appears in Buntin and Zaslavsky (2004). "Too much ado about two-part models and transformation? Comparing methods of modeling Medicare expenditures." *Journal of Health Economics* 23(3): 525-542. The GLM component of the model will be of greatest interest—that is, predictors associated with any costs.

focus or phased rollout) or in other states. States are encouraged to include at least two pre-demonstration years.

The difference-in-differences model, which is specified as follows, should be run separately for each of the cost outcomes shown in Table C.1:

$$Costs = \beta 0 + \beta 1 * TREATMENT + \beta 2 * POST + \beta 3 * (TREATMENT * POST) + Bi * CONTROLS + \varepsilon,$$

where:

TREATMENT is the indicator variable that equals 1 for a beneficiary in the treatment group and 0 if in the comparison group;

POST is the indicator variable that equals 1 if the month occurs on or after the demonstration start date and 0 if the month occurs before the demonstation start date; and

CONTROLS are covariates, such as age, gender, race, dual Medicare-Medicaid enrollment, and month.

Interpretation of difference-in-differences models. The outputs generated from the difference-in-differences model demonstrate the trends in PBPM costs in the treatment and comparison groups over time from before and after the demonstration began, including whether the rate of change differs in each of the groups. If the average marginal effect of the interaction term ($\beta 3 * TREATMENT * POST$) is a positive dollar amount, then the demonstration is associated with an increase in costs relative to the comparison group trend. If the interaction term is a negative dollar amount, then the demonstration is associated with a decrease in costs relative to the comparison group trend. States should assess not only the sign of the effect, but also whether or not the effect is statistically significant from zero.

Table C.3 presents a template for states that use a difference-in-differences model to report adjusted cost outcomes, using the SMI/SED demonstration as an example. States with SUD demonstrations should adapt the table to costs related to the diagnosis and treatment of SUD.

Table C.3. Template for reporting adjusted cost outcomes: Difference-in-differences regression results (present marginal effects and standard errors), SMI/SED demonstration example

	Total costs	Total federal costs	SMI/SED-IMD costs[a]	Other SMI/SED costs[a]	Non-SMI/SED costs[a]	Outpatient costs, non-ED	Outpatient costs, ED	Inpatient costs	Pharmacy costs	Long-term care costs
Logit										
Intervention group										
Demonstration period										
Treatment group * demonstration period										
Covariates										
Constant										
GLM										
Treatment group										
Demonstration period										
Treatment group * demonstration period										
Covariates										
Constant										

[a]For analyses of SUD-related costs, these columns would reflect IMD, other, and unrelated costs for diagnoses and treatment of SUD.

ED = emergency department; GLM = generalized linear model; IMD = Institutions for Mental Diseases; SMI/SED = serious mental illness/serious emotional disturbance; SUD = substance use disorder.

b. Alternative 1: Conduct an interrupted time series analysis without a comparison group

For states that do not have a viable comparison group, one approach to assessing trends in costs over time is to use an interrupted time series model. This model can estimate different linear effects in the pre-demonstration and post-demonstration periods. States are encouraged to include at least two pre-demonstration years. States should report marginal effects and standard errors.

The interrupted time series model, which is specified as follows, should be run separately for each of the cost outcomes shown in Table C.1:

$$Costs = \beta 0 + \beta 1 * TIME + \beta 2 * POST + \beta 3 * (TIME * POST) + \mathrm{B}i * CONTROLS + \varepsilon,$$

where:

TIME is a count variable that starts with the first quarter of pre-demonstration period data and ends with the last quarter of post-demonstration period data;

POST is the indicator variable that equals 1 if the month occurred on or after the demonstration start date and 0 if the month occurred before the demonstration start date; and

CONTROLS are covariates, such as age, gender, race, dual Medicare-Medicaid enrollment, and month.

Interpretation of the interrupted time series model without a comparison group. The estimates from the interrupted time series model demonstrate the trends in PBPM costs in the treatment group. If the average marginal effect of the interaction term ($\beta 3 * TIME * POST$) is a positive dollar amount, then the costs in the post-demonstration period are higher than the costs in the pre-demonstration period. If the interaction term is a negative dollar amount, then the costs in the post-demonstration period are lower than in the pre-demonstration period. States also assess whether or not the effect is statistically significant from zero. Interrupted time series models without a comparison group cannot determine whether any observed changes are caused by the demonstration.

Table C.4 presents a template for states that use an interrupted time series model to report adjusted outcomes, using the SMI/SED demonstration as an example. States with SUD demonstrations should adapt the table to costs related to diagnosis and treatment of SUD.

Table C.4. Template for reporting adjusted cost outcomes: Interrupted time series results (present marginal effects and standard errors), SMI/SED demonstration example

	Total costs	Total federal costs	SMI/SED-IMD costs[a]	Other SMI/SED costs[a]	Non-SMI/SED costs[a]	Outpatient costs, non-ED	Outpatient costs, ED	Inpatient costs	Pharmacy costs	Long-term care costs
Logit										
Demonstration period										
Time (continuous)										
Demonstration period * time (continuous)										
Covariates										
Constant										
GLM										
Demonstration period										
Time (continuous)										
Demonstration period * time (continuous)										
Covariates										
Constant										

[a]For analyses of SUD-related costs, these variables would reflect IMD-related, other, and unrelated costs for diagnoses and treatment of SUD.

ED = emergency department; GLM = generalized linear model; IMD = Institutions for Mental Diseases; SMI/SED = serious mental illness/serious emotional disturbance; SUD = substance use disorder.

c. Alternative 2: Conduct a pre-post analysis

If a state does not have comparison group data available and does not have enough data points to conduct an interrupted times series model, a third option would be to conduct a pre-post analysis, which compares costs among those affected by the SMI/SED or SUD demonstration in a single pre-demonstration period and a single post-demonstration period. The pre-post model, which is specified as follows, should be run separately for each of the cost outcomes shown in Table C.1:

$$Costs = \beta 0 + \beta 1 * POST + \mathrm{B}i * CONTROLS + \varepsilon,$$

where:

POST is the indicator variable that equals 1 if the month occurred on or after the demonstration start date; and

CONTROLS are covariates, such as age, gender, race, dual Medicare-Medicaid enrollment, and month.

Interpretation of pre-post models. The pre-post regression results demonstrate PBPM costs in the treatment group before and after the demonstration began. If the average marginal effect of the demonstration indicator ($\beta 1 * POST$) is a positive dollar amount, then the costs in the post-demonstration period are higher than the costs in the pre-demonstration period. If the indicator is a negative dollar amount, then the costs in the post-demonstration period are lower than in the pre-demonstration period. States also should assess whether or not the effect is statistically significant from zero. Pre-post models cannot determine whether any observed changes are caused by the demonstration. Nonetheless, when comparison group data are unavailable, pre-post models will provide information about changes in costs after demonstration implementation.

Table C.5 is a sample template for states that use pre-post models to report adjusted outcomes, using the SMI/SED demonstration as an example. States with SUD demonstrations should adapt the table to costs related to the diagnosis and treatment of SUD.

Table C.5. Template for reporting adjusted cost outcomes: Pre-post regression results (present marginal effects and standard errors), SMI/SED demonstration example

	Total costs	Total federal costs	SMI/SED IMD costs[a]	Other SMI/SED costs[a]	Non-SMI/ SED costs[a]	Outpatient costs, non-ED	Outpatient costs, ED	Inpatient costs	Pharmacy costs	Long- term care costs
Logit										
Demonstration period										
Covariates										
Constant										
GLM										
Demonstration period										
Covariates										
Constant										

[a]For analyses of SUD-related costs, these variables would reflect IMD-related, other, and unrelated costs for diagnoses and treatment of SUD.

ED = emergency department; GLM = generalized linear model; IMD = Institutions for Mental Diseases; SMI/SED = serious mental illness/serious emotional disturbance; SUD = substance use disorder.

5. Challenges and limitations, including suggested solutions

Table C.6 highlights potential data challenges that states may face when conducting cost analyses for IMD and other costs.

Table C.6. Challenges with cost analyses of SMI/SED or SUD Section 1115 demonstrations and potential solutions

Issue	Challenge	Potential solution
SMI/SED or SUD costs	When identifying SMI/SED-related costs or SUD-related costs, states may not be able to achieve the same precision as when they assess overall Medicaid costs, particularly if capitated managed care contracts are used to deliver behavioral health care services.	Determine at the outset which cost data sources will be available in a timely fashion for the pre- and post-demonstration periods. Be explicit in all analyses about which costs are and are not included. Consider using encounter data and shadow pricing to determine what SMI/SED costs or SUD costs would have been under FFS. Alternatively, states may use provider paid amounts recorded on Medicaid encounters to estimate costs.
IMD data	IMDs may not submit timely and complete claims data. In addition, data accuracy and completeness may vary by facility or IMD within a state.	Determine at the outset which cost data sources will be available for the pre- and post-demonstration periods. Be explicit in all analyses about which costs are and are not included. If a state paid for IMD services out of state funds in the pre-demonstration period, IMD data may already be included in the IP files in the pre-demonstration period. IMD data should be available in the state data files in the post-demonstration period once the Section 1115 demonstration is in place.
Pre-demonstration data	Using only one year of pre-demonstration data may not be sufficient to effectively assess costs during the pre-demonstration period.	States may wish to consider using two pre-demonstration period years of data to help identify beneficiaries with SMI/SED treatment needs or beneficiaries with SUD treatment needs for whom a full year of service use and costs can be observed, prior to the demonstration's start.
Comparison group	It may be challenging for some states to identify an appropriate comparison group to conduct cost analyses (for example, if the entire state is participating in the Section 1115 demonstration) or if other states do not have similar Medicaid programs.	Conduct an interrupted time series model without a comparison group or pre-post regression model, rather than a preferred difference-in-differences model.
Data run-out	It may be difficult to demonstrate cost outcomes if there are not enough data available in the post-demonstration period, even by the end of the demonstration period, due to anticipated data lags.	Conduct descriptive analyses only. Make sure to document in findings that there may not be enough post-demonstration data to fully capture the impact of the demonstration on costs (a minimum of 6 months but preferably 12–24 months of data are needed, because changes in care-seeking behavior may take time to materialize).

FFS = fee-for-service; IMD = Institutions for Mental Diseases; SMI/SED = serious mental illness/serious emotional disturbance; SUD = substance use disorder.