

ATTACHMENT O

DEVELOPING THE EVALUATION DESIGN

Introduction

For states that are testing new approaches and flexibilities in their Medicaid programs through section 1115 demonstrations, evaluations are crucial to understand and disseminate what is or is not working and why. The evaluations of new initiatives seek to produce new knowledge and direction for programs and inform Medicaid policy for the future. While a narrative about what happened during a demonstration provides important information, the principal focus of the evaluation of a section 1115 demonstration should be obtaining and analyzing data on the process (e.g., whether the demonstration is being implemented as intended), outcomes (e.g., whether the demonstration is having the intended effects on the target population), and impacts of the demonstration (e.g., whether the outcomes observed in the targeted population differ from outcomes in similar populations not affected by the demonstration). Both state and federal governments need rigorous quantitative and qualitative evidence to inform policy decisions.

CMS expects evaluation designs to be rigorous, incorporate baseline and comparison group assessments, as well as statistical significance testing. Technical assistance resources for constructing comparison groups, identifying causal inferences, phasing implementation to support evaluation, and designing and administering beneficiary surveys are available on Medicaid.gov: <https://www.medicaid.gov/medicaid/section-1115-demo/evaluation-reports/evaluation-designs-and-reports/index.html>. If the state needs technical assistance using this outline or developing the evaluation design, the state should contact its demonstration team.

Expectations for Evaluation Designs

All states with Medicaid section 1115 demonstrations are required to conduct an evaluation, and the Evaluation Design is the roadmap for conducting the evaluation. The roadmap begins with the stated goals for the demonstration followed by the measurable evaluation questions and quantifiable hypotheses, all to support a determination of the extent to which the demonstration has achieved its goals. When conducting analyses and developing the evaluation reports, every effort should be made to follow the approved methodology. However, the state may request, and CMS may agree to, changes in the methodology in appropriate circumstances

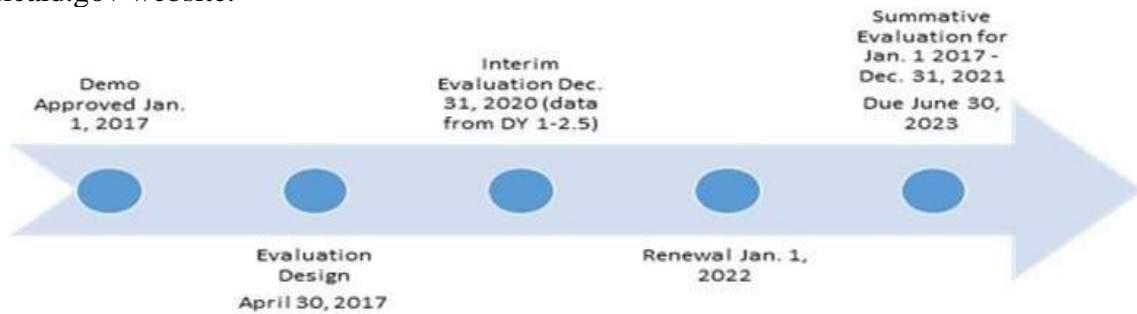
The format for the Evaluation Design is as follows:

- A. General Background Information;
- B. Evaluation Questions and Hypotheses;
- C. Methodology;
- D. Methodological Limitations;
- E. Attachments.

Submission Timelines

There is a specified timeline for the state's submission of Evaluation Design and Reports. (The graphic below depicts an example of a deliverables timeline for a 5-year demonstration). In addition, the state should be aware that section 1115 evaluation documents are public records.

The state is required to publish the Evaluation Design to the state’s website within thirty (30) calendar days of CMS approval, as per 42 CFR 431.424(e). CMS will also publish a copy to the Medicaid.gov website.



Required Core Components of All Evaluation Designs

The Evaluation Design sets the stage for the Interim and Summative Evaluation Reports. It is important that the Evaluation Design explain the goals and objectives of the demonstration, the hypotheses related to the demonstration, and the methodology (and limitations) for the evaluation. A copy of the State’s Driver Diagram (described in more detail in paragraph B2 below) should be included with an explanation of the depicted information.

A. General Background Information – In this section, the state should include basic information about the demonstration, such as:

1. The issue/s that the state is trying to address with its section 1115 demonstration and/or expenditure authorities, the potential magnitude of the issue/s, and why the state selected this course of action to address the issue/s (e.g., a narrative on why the state submitted an 1115 demonstration proposal).
2. The name of the demonstration, approval date of the demonstration, and period of time covered by the evaluation;
3. A brief description of the demonstration and history of the implementation, and whether the draft Evaluation Design applies to an amendment, extension, renewal, or expansion of, the demonstration;
4. For renewals, amendments, and major operational changes: A description of any changes to the demonstration during the approval period; the primary reason or reasons for the change; and how the Evaluation Design was altered or augmented to address these changes.
5. Describe the population groups impacted by the demonstration.

B. Evaluation Questions and Hypotheses – In this section, the state should:

1. Describe how the state’s demonstration goals are translated into quantifiable targets for improvement, so that the performance of the demonstration in

achieving these targets could be measured.

2. Include a Driver Diagram to visually aid readers in understanding the rationale behind the cause and effect of the variants behind the demonstration features and intended outcomes. A driver diagram is a particularly effective modeling tool when working to improve health and health care through specific interventions. The diagram includes information about the goal of the demonstration, and the features of the demonstration. A driver diagram depicts the relationship between the aim, the primary drivers that contribute directly to achieving the aim, and the secondary drivers that are necessary to achieve the primary drivers for the demonstration. For an example and more information on driver diagrams:
<https://innovation.cms.gov/files/x/hciatwoaimsdvrrs.pdf>
3. Identify the state's hypotheses about the outcomes of the demonstration:
 - a. Discuss how the evaluation questions align with the hypotheses and the goals of the demonstration;
 - b. Address how the research questions / hypotheses of this demonstration promote the objectives of Titles XIX and/or XXI.

C. Methodology – In this section, the state is to describe in detail the proposed research methodology. The focus is on showing that the evaluation meets the prevailing standards of scientific and academic rigor, and the results are statistically valid and reliable, and that where appropriate it builds upon other published research (use references).

This section provides the evidence that the demonstration evaluation will use the best available data; reports on, controls for, and makes appropriate adjustments for the limitations of the data and their effects on results; and discusses the generalizability of results. This section should provide enough transparency to explain what will be measured and how. Specifically, this section establishes:

1. *Evaluation Design* – Provide information on how the evaluation will be designed. For example, will the evaluation utilize a pre/post comparison? A post-only assessment? Will a comparison group be included?
2. *Target and Comparison Populations* – Describe the characteristics of the target and comparison populations, to include the inclusion and exclusion criteria. Include information about the level of analysis (beneficiary, provider, or program level), and if populations will be stratified into subgroups. Additionally discuss the sampling methodology for the populations, as well as support that a statistically reliable sample size is available.
3. *Evaluation Period* – Describe the time periods for which data will be included.
4. *Evaluation Measures* – List all measures that will be calculated to evaluate the demonstration. Include the measure stewards (i.e., the organization(s) responsible for the evaluation data elements/sets by “owning”, defining, validating; securing; and submitting for endorsement, etc.) Include numerator and denominator information.

Additional items to ensure:

- a. The measures contain assessments of both process and outcomes to evaluate the effects of the demonstration during the period of approval.
 - b. Qualitative analysis methods may be used, and must be described in detail.
 - c. Benchmarking and comparisons to national and state standards should be used, where appropriate.
 - d. Proposed health measures could include CMS's Core Set of Health Care Quality Measures for Children in Medicaid and CHIP, Consumer Assessment of Health Care Providers and Systems (CAHPS), the Initial Core Set of Health Care Quality Measures for Medicaid-Eligible Adults and/or measures endorsed by National Quality Forum (NQF).
 - e. Proposed performance metrics can be selected from nationally recognized metrics, for example from sets developed by the Center for Medicare and Medicaid Innovation or for meaningful use under Health Information Technology (HIT).
 - f. Among considerations in selecting the metrics shall be opportunities identified by the state for improving quality of care and health outcomes, and controlling cost of care.
5. *Data Sources* – Explain where the data will be obtained, and efforts to validate and clean the data. Discuss the quality and limitations of the data sources.

If primary data (data collected specifically for the evaluation) – The methods by which the data will be collected, the source of the proposed question/responses, the frequency and timing of data collection, and the method of data collection. Copies of any proposed surveys must be reviewed with CMS for approval before implementation.

6. *Analytic Methods* – This section includes the details of the selected quantitative and/or qualitative measures to adequately assess the effectiveness of the demonstration. This section should:
- a. Identify the specific statistical testing which will be undertaken for each measure (e.g., t-tests, chi-square, odds ratio, ANOVA, regression). Table A is an example of how the state might want to articulate the analytic methods for each research question and measure.
 - b. Explain how the state will isolate the effects of the demonstration (from other initiatives occurring in the state at the same time) through the use of comparison groups.
 - c. A discussion of how propensity score matching and difference-in-differences design may be used to adjust for differences in comparison populations over time (if applicable).
 - d. The application of sensitivity analyses, as appropriate, should be considered.
7. *Other Additions* – The state may provide any other information pertinent to the Evaluation Design of the demonstration.

Table A. Example Design Table for the Evaluation of the Demonstration

Research Question	Outcome measures used to address the research	Sample or population subgroups to be compared	Data Sources	Analytic Methods
Hypothesis 1				
Research question 1a	-Measure 1 -Measure 2 -Measure 3	-Sample e.g. All attributed Medicaid beneficiaries -Beneficiaries with diabetes diagnosis	-Medicaid FFS and encounter claims records	-Interrupted time series
Research question 1b	-Measure 1 -Measure 2 -Measure 3 -Measure 4	-sample, e.g., PPS patients who meet survey selection requirements (used services within the last 6 months)	-Patient survey	Descriptive statistics
Hypothesis 2				
Research question 2a	-Measure 1 -Measure 2	-Sample, e.g., PPS administrators	-Key informants	Qualitative analysis of interview material

D. Methodological Limitations – This section provides detailed information on the limitations of the evaluation. This could include the design, the data sources or collection process, or analytic methods. The state should also identify any efforts to minimize the limitations. Additionally, this section should include any information about features of the demonstration that effectively present methodological constraints that the state would like CMS to take into consideration in its review.

E. Special Methodological Considerations – CMS recognizes that there may be certain instances where a state cannot meet the rigor of an evaluation as expected by CMS. In these instances, the state should document for CMS why it is not able to incorporate key components of a rigorous evaluation, including comparison groups and baseline data analyses. Examples of considerations include: When the state demonstration is:

- a. Long-standing, non-complex, unchanged, or
- b. Has previously been rigorously evaluated and found to be successful, or
- c. Could now be considered standard Medicaid policy (CMS published regulations or guidance)

When the demonstration is also considered successful without issues or concerns that would require more regular reporting, such as:

- a. Operating smoothly without administrative changes; and
- b. No or minimal appeals and grievances; and
- c. No state issues with CMS 64 reporting or BN; and

- i. No Corrective Action Plans (CAP) for the demonstration.

F. Attachments

1) Independent Evaluator. This includes a discussion of the state's process for obtaining an independent entity to conduct the evaluation, including a description of the qualifications that the selected entity must possess, and how the state will assure no conflict of interest. Explain how the state will assure that the Independent Evaluator will conduct a fair and impartial evaluation, prepare an objective Evaluation Report, and that there would be no conflict of interest. The Evaluation Design should include a "No Conflict of Interest" statement signed by the independent evaluator.

2) Evaluation Budget. A budget for implementing the evaluation shall be provided with the draft Evaluation Design. It will include the total estimated cost, as well as a breakdown of estimated staff, administrative, and other costs for all aspects of the evaluation. Examples include, but are not limited to, the development of all survey and measurement instruments; quantitative and qualitative data collection; data cleaning and analyses; and reports generation. A justification of the costs may be required by CMS if the estimates provided do not appear to sufficiently cover the costs of the draft Evaluation Design, if CMS finds that the draft Evaluation Design is not sufficiently developed, or if the estimates appear to be excessive.

3) Timeline and Major Milestones. Describe the timeline for conducting the various evaluation activities, including dates for evaluation-related milestones, including those related to procurement of an outside contractor, if applicable, and deliverables. The Final Evaluation Design shall incorporate the Interim Evaluation Reports and Summative Evaluation. Pursuant to 42 CFR 431.424(c)(v), this timeline should also include the date by which the Final Summative Evaluation report is due.

