DEPARTMENT OF HEALTH & HUMAN SERVICES
Centers for Medicare & Medicaid Services
7500 Security Boulevard, Mail Stop S2-25-26
Baltimore, Maryland 21244-1850

State Demonstrations Group

March 1, 2019

Allison Taylor
Medicaid Director
Indiana Family and Social Services Administration
402 W. Washington Street, Room W461, MS25
Indianapolis, IN 46204

Dear Ms. Taylor:

The Centers for Medicare & Medicaid Services (CMS) has completed its review of Indiana's draft evaluation design, submitted on November 27, 2018, in accordance with the Healthy Indiana Plan (HIP) section 1115(a) demonstration (Project No. 11-W-00296/5) Special Terms and Conditions (STC), in particular its Section XV. CMS reviewed the design against the applicable STCs, as well as CMS' standardized evaluation guidance (https://www.medicaid.gov/medicaid/section-1115-demo/downloads/evaluation-reports/developing-the-evaluation-design.pdf). We appreciate your submitting the evaluation design.

The state's HIP evaluation design partially addresses the requirements for evaluation described in the STCs. The attached feedback includes CMS' observations and recommendations on areas that should be better articulated and strengthened in the design, to assure that the state's independent evaluator, The Lewin Group, conducts an evaluation that is rigorous and aligned with the requirements in the STCs. For example, the design should refine some of the hypotheses and research questions as well as expand and finesse the set of measureable outcomes to meet the goals of the evaluations more fully. The design should also identify robust comparison groups and propose stronger analytic methods, including quasi-experimental approaches. It will also be important for the design to describe sufficiently mechanisms for ensuring adequate sample sizes for previous member surveys. At this time, CMS has not provided elaborate feedback on the segments of the evaluation design related to the community engagement (CE) initiative of the state's 1115 demonstration, since CMS recognizes that the state intends to expand the design addressing CE to reflect CMS' evaluation guidance targeted around this policy.

To that end, in addition to our feedback on the draft HIP evaluation design, CMS has shared, for the state's consideration, draft evaluation guidance for states that have initiated or are contemplating implementation of Eligibility and Coverage section 1115 demonstrations, including policies such as CE and non-eligibility periods. These draft guidance documents highlight CMS' recommendations regarding important hypotheses, evaluation questions, and analytic approaches for impact analyses of the 1115 demonstrations. These materials have been through a rigorous development process internal to CMS, as well as external through a series of

state advisory and research community workgroups, and is currently undergoing the final phases of internal clearance. Until further notice, these documents should be considered embargoed and not for further release.

The state will find examples of hypotheses and evaluation questions in the evaluation guidance focusing on changes in employment, income, transitions to employer sponsored insurance, health outcomes, compliance with program requirements, health services expenditures, among other areas of interest. Further, the guidance addresses incorporating comparison groups in order to understand the impact of the demonstration on the population of interest. The guidance also underscores the importance of tracking beneficiaries longitudinally, particularly after beneficiaries leave Medicaid, in order to assess the long-term goals of the demonstration. Since Indiana has policies such as CE, premiums, non-eligibility periods, and healthy behavior incentives simultaneously in implementation, and given the overall importance of understanding the cost implications of any such demonstration, CMS deems it important to consider the different demonstration components in coordination for effectively formulating the evaluation design.

The evaluation guidance documents shared with Indiana are still pre-decisional drafts and should not be distributed or attributed without CMS approval. The first set of guidance documents is slated for a wider release to the states in spring 2019. CMS believes that Indiana will find the guidance useful in evaluating adequately specific policy initiatives under the state's HIP section 1115 demonstration.

As noted in Section XV.4 of the STCs regarding the *Evaluation Design Approval and Updates*, the state is expected to submit a revised draft evaluation design within 60 days after receipt of CMS' comments. CMS is comfortable with the state submitting the revised evaluation design 60 days from the day of the state rollout of these resources, if the state finds the additional time conducive to preparing a more robust evaluation design. CMS would also be happy to review a revised design if Indiana prefers to submit one to CMS earlier than the allocated time. We are available to provide further technical assistance if the state would find that helpful. CMS also encourages the state to remain active in the CMS sponsored CE Learning Collaborative to support learning and diffusion of best practices.

If you have any questions, please contact your project officer, Ms. Shanna Janu at Shanna.Janu@cms.hhs.gov. We look forward to our continued partnership with you and your team on the Indiana HIP section 1115 demonstration.

Sincerely,

/s/

Andrea J. Casart
Director
Division of Medicaid Expansion Demonstrations

Enclosure

cc: Ruth Hughes, Associate Regional Administrator, CMS Chicago Regional Office

Healthy Indiana Plan Evaluation Design: CMS Comments
February 12, 2019

Introduction

Indiana's section 1115 demonstration authorizes the state to offer to its eligible Medicaid population the Healthy Indiana Plan (HIP) that provides a high deductible health plan paired with a health savings account, known as the Personal Wellness and Responsibility (POWER) account. Beneficiaries are required to make monthly premium contributions to their POWER account to enroll and maintain access to the enhanced HIP Plus plan. Those with income at or below the federal poverty level who do not make the monthly POWER account contributions default into a more limited plan, called the HIP Basic. Intended to encourage personal responsibility, HIP results in disenrollment and non-eligibility periods of six months for non-payment of monthly premiums for beneficiaries above the federal poverty level. The current HIP demonstration, approved by the Centers for Medicare & Medicaid Services (CMS) on February 1, 2018, was enhanced with a tiered structure to its monthly POWER account contributions to improve compliance with payments, a premium surcharge to lower tobacco usage, and community engagement (CE) policies to promote employment among beneficiaries. In addition to these newer goals associated with the more recent policy initiatives, the demonstration maintains its overarching objectives of improved health care access and health outcomes and overall member satisfaction. Indiana submitted to CMS an evaluation design for the demonstration, Healthy Indiana Plan (HIP) Final Evaluation Plan (Dated November 13, 2018; Version 6), which aims to assess whether the demonstration succeeds in achieving these key objectives.

CMS reviewed Indiana's HIP evaluation design against the demonstration's Special Terms and Conditions (STC) for the period of February 1, 2018 through December 31, 2020, dated February 1, 2018. This review led to the assessment that the evaluation design only partially addressed the requirements of the STCs. The discussion below details specific recommendations intended to strengthen and improve the evaluation design and thereby ensure complete alignment with the requirements stipulated in the STCs and CMS's general evaluation design guidance, also affixed as Attachment A with the STCs.

Recommendations

In order to ensure that the evaluation effectively assesses the impact of Indiana's HIP 1115 demonstration, CMS encourages the state to address several critical issues discussed below as the state revises the evaluation design.

CMS is developing evaluation guidance documents focusing on eligibility and coverage demonstrations, such as community engagement (CE), healthy behaviors, premiums, premium assistance, waivers of retroactive eligibility, and non-eligibility periods. These guidance documents are slated for a wider state rollout in spring 2019. However, CMS will share with Indiana the sections of the guidance documents relevant for the state's HIP demonstration in advance of the broader release because these materials will provide helpful information as the state revises the evaluation design.

As noted in Section XV.4 of the STCs regarding the *Evaluation Design Approval and Updates*, the state is expected to submit a revised draft evaluation design within 60 days after receipt of CMS' comments. However, CMS is comfortable with the state submitting the revised evaluation design 60 days from the day of the official state rollout of these resources in the spring of 2019, if the state finds the additional time beneficial to preparing a comprehensive evaluation design. CMS is also available to provide technical assistance to the state as it begins to review the materials with their evaluator.

Adherence to evaluation best practices that CMS indicates likely to be relevant for a demonstration such as HIP will support the evaluation report in providing CMS and its federal partners, the state, and other stakeholders meaningful and conclusive information.

A. Evaluation questions and hypotheses. Several hypothesis and research questions associated with the stated goals of the demonstration need further clarification and definition. For the first three goals of the demonstration, the hypotheses are aligned with the goals. However, for the fourth goal, the hypotheses do not relate to "efficient use of health care services" and the term "efficient" is not defined. For goal 5, comparing satisfaction between those who comply with required payments (HIP Plus) and those who do not (and thereby, default or transition to HIP Basic) is problematic, because beneficiaries who comply with program requirements and maintain enrollment in HIP Plus may be a self-selected group.

B. Well-defined quantitative outcome measures. The evaluation design should specify clearly quantifiable outcome measures to assess whether data sources, comparison groups, and quantitative methods are appropriate and sufficient for addressing the various research questions. For instance, it will be helpful to provide further information regarding the claims-based outcome measure, "#/% of HIP members with utilization codes related to tobacco use" (p. 31) for the new tobacco cessation component of the demonstration. The evaluation of other research questions would be enhanced by including additional measures or omitting measures that are not appropriate (see comment B1 for examples). It is also important to assess carefully the availability of measures for comparison groups for inclusion in the analysis.

Specific CMS recommendations include:

1. The evaluation design lists several health utilization and outcome measures, but does not specify in detail what these measures will be beyond broad categorization. For example, "#/% of members utilizing primary and specialty care services (claims)" or "utilization rates of chronic disease care services (claims)" (p. 15) do not offer sufficient detail. The revised design should be more specific in defining the outcomes of interest. Further, each outcome should include the endorsed measure specification that will be used and indicate the measure steward. If state-developed measures will be used, the numerator and denominator should be specified.

2. Given that beneficiaries will be enrolled for spans of different lengths, the design needs more detail about the period over which the outcome measures are defined; e.g., physician office visits within the first two/six months of enrollment. The design should also specify when the sample will be restricted to beneficiaries who are enrolled for a minimum duration of time; e.g., measure X for beneficiaries who are enrolled for at least 6 months.

3. The state should carefully identify outcomes measures that are in alignment with the objectives of the program and the hypothesis in consideration. For instance, the "# of chronic conditions (claims)" a beneficiary is diagnosed with (listed under Goal 1.2) may not be a suitable measure of health outcomes. Chronic conditions are by definition long-term and managed, not resolved, through medical treatment. The number of chronic conditions *observed* is likely to increase with regular contact with physicians, as there are more opportunities to document chronic diseases. The number of chronic conditions per beneficiary may also be a measure of selection due to other program changes related to premiums or CE.

4. The evaluation design should incorporate analysis of certain additional outcomes, which although not the intended goals of the demonstration, might nonetheless influence the interpretation of other demonstration findings. In fact, the driver diagrams should consider adding potential unintended consequences to help formulate appropriate hypothesis, research questions, and outcome measures for assessing such effects of the demonstration. For example, observed changes in health care utilization could be driven by changes in the composition of enrollees. Because those with more health conditions are likely to value their health coverage more, it is possible that health care utilization increases on a per capita basis among those who remain in the demonstration.

   For a comprehensive understanding of the demonstration's impact, the evaluation should examine the effects of the new HIP policies on total enrollment in Medicaid, measured by the number of beneficiaries enrolled relative to the population of likely-eligible persons (using the American Community Survey or other national surveys). Similarly, "lockout" or any non-eligibility periods resulting from unpaid premiums should be explored in the context of the new program design. Among the pre-decisional draft versions of the evaluation guidance documents that CMS will share with the state, one appendix segment – Indiana will find – is specifically developed to assist with formulating hypothesis and research questions associated with non-eligibility periods that might result from the programmatic design of the state's 1115 demonstration.

C. Comparison group selection. To estimate the effects of HIP on those enrolled and subject to its provisions, the evaluation requires one or more comparison groups that are similar to the treated group but exempt from the requirements of the demonstration.[1] Without valid comparison groups, the evaluation findings will not be generalizable. The evaluation design mentions benchmarking, but does not specify which measures will be benchmarked and does not present analysis of whether these benchmark populations are appropriate comparisons.

Indeed, no comparison groups are clearly defined in the state's evaluation design besides mention of drawing comparisons with Medicaid member in other comparable states using 'existing summary federal survey data' (e.g., p. 16). Even then the design does not indicate how and which comparison states will be selected, or consider quasi-experimental approaches that would use individual-level federal survey data. The proposed comparisons to summary data

[1] See "Selecting the Best Comparison Group and Evaluation Design: A Guidance Document for State Section 1115 Demonstration Evaluations" (https://www.medicaid.gov/medicaid/section-1115-demo/downloads/evaluation-reports/comparison-grp-eval-dsgn.pdf) for a detailed discussion of best practices in comparison group selection.

from the federal surveys, such as the Behavioral Risk Factor Surveillance Survey (BRFSS) and the National Health Interview Survey (NHIS) for privately insured and uninsured populations in Indiana (e.g., pp. 15-18) are not particularly instructive because these populations are likely different in observable and unobservable ways from the demonstration population.

For within-state comparison groups, the evaluation design focuses on (1) comparing HIP Basic and HIP Plus members, or (2) comparing the current demonstration period to earlier years of HIP 2.0, before the tobacco cessation component and CE requirements were implemented. The first comparison strategy may be problematic because HIP members with incomes below 100% of the poverty level choose whether to enroll in Plus or Basic, and collectively, HIP Plus and Basic members have different health statuses as documented in the evaluation of HIP 2.0's first demonstration period. The second comparison strategy may be problematic if the new HIP policies are likely to cause some individuals to disenroll, thereby changing the composition of the enrolled population in observable and unobservable ways. The state should address both these concerns directly in revising the evaluation design. The approach requires further deliberation than simply stating that confounding factors, such as health status (as indicated by a medically frail indicator) and socio-demographic factors (p. 11) will be accounted for.

CMS's specific recommendations with respect to selecting appropriate comparison groups are:

1. The evaluation design should specify comparison groups in other states, with specific focus on states that are similar to Indiana in terms of Medicaid eligibility for childless adults and population characteristics, but states that did not implement any comparable demonstrations during the period. The comparison groups can be identified or closely approximated using individual-level data in national surveys, such as the American Community Survey (ACS) or the Behavioral Risk Factor Surveillance Survey (BRFSS), or through use of Medicaid administrative data from another state. Some data sets, such as the National Health Interview Survey (NHIS) and the Current Population Survey (CPS) may not contain enough observations once restricted to specific states and population subgroups. The state should explore the sample sizes and consider the statistical power of their empirical tests that use such data. Potential comparison states that the state might wish to consider include Nevada, New Mexico, North Dakota, Ohio, Oregon, Pennsylvania, Washington, and West Virginia, among others.

2. The state should explore to identify within-state comparison groups of Medicaid beneficiaries not subject to the HIP premium requirements and non-eligibility period penalties. However, it will be important to consider whether such potential within-state comparison groups may be different in important unobserved ways after controlling for observable characteristics. Since the Section 1931 parents and caretaker-relatives and the medically frail individuals receive state plan benefits, it is probable that both groups could be candidates for regression discontinuity approaches, a methodologically rigorous analysis, discussed further below in the quantitative analysis section.

3. As discussed above, the composition of the enrolled population may change in meaningful ways in the current demonstration period relative to the previous periods, due in part to new program rules such as the premium surcharge for tobacco users and the CE requirements. For example, beneficiary disenrollment due to increased premiums, rising income above

138% FPL due to increased labor force participation, and disenrollment due to failure of reporting or complying with CE requirements will likely change the composition of the enrolled population in ways that would limit the comparability of the prior and current demonstration population as a whole. Therefore, before using the prior HIP 2.0 enrollees as comparators for current HIP enrollees, the evaluators should examine if the overall composition of the enrolled population has changed and caveat their inferences accordingly. Similarly, it would be helpful if the evaluation design provides additional details on the state's planned analyses comparing HIP Plus to HIP Basic enrollees. The design should provide, for instance, additional details on how this type of analysis would control for observable differences between the two groups, including their health status. This will be particularly useful in understanding the effects of HIP policies, and would represent an improvement relative to prior IN HIP reports where comparisons and conclusions about demonstration policies were sometimes drawn without these controls.

4. Lastly, the target population is not well defined in the evaluation design in the sense that it is unclear what groups, if any, are exempt from the new HIP rules. A more precisely defined target population will perhaps generate ideas for potential comparison groups.

D. Quantitative analytic methods. The state's evaluation design outlines, broadly, the statistical methods that will be used, but lacks analytical rigor. For example, the design does not propose any quasi-experimental approaches or discuss potential sensitivity analyses. The analytic approach should incorporate a more rigorous quantitative methodology based on available, valid, and contemporaneous comparison groups (see above) and go beyond the proposed pre-post design or comparison to summary statistics and national benchmarks.[2]

Specifically, the state should take into consideration the following recommendations:

1. The state's analysis should consider using propensity score matching or weighting to adjust for observable differences between the treatment and comparison groups. This will be particularly important in using adult Medicaid beneficiaries in other states, who are likely different in observable ways, as comparison groups. Similarly, outcomes for HIP Basic and HIP Plus beneficiaries should be presented with controls for observable characteristics or use propensity score methods, as these groups have different underlying health status, as reflected in prior evaluation reports on HIP 2.0.

2. If sharp cutoffs exist for exemptions from HIP program rules, such as for those deemed medically frail and exempt from premiums, regression discontinuity designs may be appropriate and would provide a credible estimate of a causal effect. For example, if the medical frailty scores are available, and the cutoff value for determining medical frailty exemptions is known, the regression discontinuity approach can be applied to compare beneficiaries just below the cutoff to those above, who would serve as a viable comparison group. While this would produce results specific to beneficiaries with higher than average medical needs, this is an important subgroup to explore. Similarly, the state could consider

[2] See "Best Practice in Causal Inference for Evaluations of Section 1115 Eligibility and Coverage Demonstrations" (https://www.medicaid.gov/medicaid/section-1115-demo/downloads/evaluation-reports/causal-inference.pdf) for a discussion of these aspects of the evaluation planning process.

an approach around the income cutoff between Section 1931 parents and caretaker relatives and those who qualify for the adult group.

3. Since several years of data are available, and the period when the new policies begin is clearly defined, an interrupted time series design or a difference-in-differences design, pending the availability of comparison groups, is recommended.

E. Subgroup analysis. Depending on sample size and statistical power, the evaluation design should better articulate plans to estimate the effects of the demonstration on various subgroups. While the evaluation design mentions subgroups, the rationale for exploring particular subgroups is not discussed. Therefore, in revising the evaluation design:

1. The state should consider analyzing subgroups by regional economic conditions, substance-use disorder status (especially opioid dependency), disability status, racial/ethnic group, age group, or eligibility group. Justification for why a subgroup analysis is undertaken would be helpful in motivating the analysis plan.

F. Primary data collection. Consistency between questions in previous HIP 2.0 surveys and the current HIP survey are valuable to allow for comparison. However, it may be important to be mindful of the wording in some survey questions that had particularly high rates of non-response or vague responses in the first HIP 2.0 survey. Overall, CMS has the following recommendations regarding primary data collection:

1. The beneficiary survey should use the same wording as previous surveys, except in cases where the survey questions resulted in high rates of non-responses. Questions with high rates of certain types of responses, such as "unsure" should be considered for redesign to make them more easily understood.

2. Where resources are limited or response rates are low, the surveys should focus on key groups of participants. For example, the "Previous Member Survey" should focus on those who leave the program because of income increases or who become subjected to non-eligibility periods due to non-payment. Responses from individuals who move out of state, or generally lose eligibility, may be less relevant to the evaluation of novel demonstration policies. Key informant interviews with beneficiaries (n=50) may provide little information beyond what is learned in beneficiary focus groups.

3. Telephone interviews are planned with 202 providers spread across the six regions (p. 9). It is unclear why providers will be interviewed or which research questions data from the interviews will support. Additional detail should be provided as to the value-add of these interviews.

G. Community Engagement (CE). CMS understands that the state intends to expand the evaluation design to address adequately the policies specific to CE, pending CMS's formal guidance on evaluating CE policies. To that end, the draft and pre-decisional version of the evaluation guidance documents that CMS will share with the state includes recommendations for enhancing the evaluation design pertaining to this policy area. At this time, these documents are not available for wider circulation or attribution without CMS approval. CMS anticipates

receiving a revised evaluation design from the state, inclusive of the proposed evaluation approaches specific to CE policies, sixty days from the date when CMS officially releases these evaluation guidance documents to Indiana, along with other states, in spring 2019.

In general, the research questions for Goal 2 (research questions 2.1–2.4) are on point, but omit an evaluation of the effects of CE requirements on health insurance coverage. The concerns about using outcomes among prior HIP 2.0 enrollees during 2015-2018 are especially important for CE, as labor force participation and income are in part dependent upon the state of the economy. Specific recommendations include:

1. The evaluation should include research questions that address outcomes related to insurance coverage including transitions to private coverage (employer or the individual marketplace) or to becoming uninsured.

2. Contemporaneous comparison groups are especially important for CE evaluation as the economy in Indiana is likely to have changed from 2015-2018 to the present. The use of in-state comparison groups that are not subject to CE requirements as well as Medicaid enrollees in other states that do not require CE are strongly advised.

H. Additional requirements for compliance with the STCs. The evaluation design does not provide information on some other important aspects that are critical to the completeness of the design. To remedy these deficiencies, the design should –

1. Describe the population groups impacted by the demonstration. The design, for instance, should explicitly identify specific groups among those earning less than 138% of the federal poverty line that are exempt from the CE requirements and premium non-payment penalties. This information can be drawn from the demonstration STCs.

2. Discuss potential methodological limitations the evaluation may still encounter despite a robust design, and consequent special methodological and mitigating considerations that the state and the evaluators might need to consider.

3. Describe the exact years and periods for which data will be included. The evaluation design notes that the interim evaluation will include the 2015-2017 period but does not specify the full range of timespan over which the demonstration will be evaluated. Several of the proposed data sources (ACS, CPS, BRFSS) are released more than a year after the period captured in the data collected, which must be considered when selecting the analysis period and the data used in the analysis.

4. Last, but not least, the design should incorporate as attachment an evaluation budget. It should also provide timeline for other milestones and major activities, such as primary data collection along with specifying the timespan covered by the evaluation beyond stating the expected completion dates for the interim and summative evaluations.