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

January 30, 2023

Farah Hanley
Chief Deputy for Health
Michigan Department of Health and Human Services
400 South Pine Street
Lansing, MI 48933

Dear Ms. Hanley:

The Centers for Medicare & Medicaid Services (CMS) completed its review of the Evaluation Design, which is required by the Special Terms and Conditions (STCs), specifically, STC 32, of Michigan's section 1115 demonstration, "Flint Michigan Section 1115 Demonstration" (Project No: 11-W-00302/5), effective through September 30, 2026. CMS has determined that the Evaluation Design, which was submitted on March 14, 2022 and revised on October 4, 2022, meets the requirements set forth in the STCs and our evaluation design guidance, and therefore approves the state's Evaluation Design.

CMS has added the approved Evaluation Design to the demonstration's STCs as Attachment C. A copy of the STCs, which includes the new attachment, is enclosed with this letter. In accordance with 42 CFR 431.424, the approved Evaluation Design may now be posted to the state's Medicaid website within thirty days. CMS will also post the approved Evaluation Design as a standalone document, separate from the STCs, on Medicaid.gov.

Please note that an Interim Evaluation Report, consistent with the approved Evaluation Design, is due to CMS one year prior to the expiration of the demonstration, or at the time of the extension application, if the state chooses to extend the demonstration. Likewise, a Summative Evaluation Report, consistent with this approved design, is due to CMS within 18 months of the end of the demonstration period. In accordance with 42 CFR 431.428 and the STCs, we look forward to receiving updates on evaluation activities in the demonstration monitoring reports.

We appreciate our continued partnership with Michigan on the Flint Michigan Section 1115 Demonstration. If you have any questions, please contact your CMS demonstration team.

Sincerely,

Danielle Daly -S

Digitally signed by Danielle Daly -S
Date: 2023.01.30 14:23:23 -05'00'

Danielle Daly
Director
Division of Demonstration
Monitoring and Evaluation

cc: Keri Toback, State Monitoring Lead, CMS Medicaid and CHIP Operations Group

Flint, Michigan Section 1115 Demonstration

#11W 00302/5

2021-2026 Renewal Evaluation

FIRST DRAFT: 03/14/2022
CURRENT DRAFT: 10/04/2022

A. General Background Information

1) *The Issue*

In April 2014, the water source in Flint, Michigan was changed from Lake Huron (via the Detroit Water and Sewerage Department) to the Flint River without appropriate treatment. This change caused lead to leach from the city's water lines (pipes), increasing the incidence of elevated lead levels in tap water and consequently in children's blood. After testing and discovery of the cause of the crisis, the water source was switched back to the original source, eighteen months later, on October 16, 2015. However, lead from the pipes continued to contaminate the tap water of structures served by the City of Flint Water Department and elevated blood lead levels persisted. In January 2016, President Obama declared an emergency in Flint, leveraging federal aid to support state and local response efforts. The declaration expired August 14, 2016, although some federal resources remained.

The State of Michigan's Department of Health and Human Services (MDHHS) applied for a Medicaid Section 1115 Demonstration waiver in February 2016, to expand eligibility and benefits. The demonstration was to support potentially exposed individuals who did not have the resources to manage the adverse health effects of lead exposure ("Flint, Michigan Section 1115 Demonstration" Approval and Special Terms and Conditions, n.d., p. 111.) These efforts were pursued because lead is a known neurotoxin and lead poisoning may result in growth, developmental, and educational difficulties (*Case Studies in Environmental Medicine (CSEM) Lead Toxicity*, n.d.) Young children (under 6 years) and children exposed *in utero* were most at risk (*Case Studies in Environmental Medicine (CSEM) Lead Toxicity*, n.d.) Access to health care and support services was necessary to ensure appropriate screening and monitoring to identify and manage the impacts associated with lead exposure.

MDHHS applied for the waiver because they identified that access to health care services was a concern in the affected region. Access was compromised among this resource poor community due to individuals lacking health insurance. Approximately 10% of the city's population were uninsured around the time of the crisis (*Flint, MI*, n.d.). In addition, some individuals with health insurance lacked sufficient resources to absorb cost-sharing requirements associated with seeking healthcare. According to 2017 United States Census data, Flint had the highest poverty rate compared to other cities of its size in the United States. Nearly 60% of children were living below the federal poverty level and the area ranked 82nd out of 83 counties in the state for general health outcomes and 71st out of 83 counties specifically for child health outcomes (*Flint & Genesee County, Michigan - Community Health Needs Assessment*, 2019). MDHHS estimated that approximately 47,000 individuals were covered by Medicaid in the City of Flint in 2016. The 2019 Community Health Needs Assessment provided additional information that, despite having access to Medicaid, these children experienced higher rates of inpatient hospitalization and longer lengths of stay (*Flint & Genesee County, Michigan - Community Health Needs*

Assessment, 2019). Thus, the demonstration's intent to expand eligibility to higher federal poverty levels, eliminate cost-sharing, and add a targeted case management (TCM) benefit focused on coordinating care was expected to partially address these health care barriers.

Lead pipe replacement was a major factor in reducing the ongoing risk of lead exposure. As of the renewal submission in April 2020, 90% of lead pipes had been replaced, but individuals were still eligible to sign up for free removal. While the lead content in the water is currently below federal standards, the water has not yet been deemed safe. MDHHS applied for, and was granted, a 5-year renewal of the original Flint Michigan 1115 Demonstration, 11-W00302/5 to run 9/15/21 - 9/30/26 reflecting Demonstration Years (DYs) 6-10 because of the ongoing exposure to the community and the need to continue supporting the health and well-being of exposed individuals.

2) The name of the demonstration to be evaluated is the Flint Michigan Section 1115 Demonstration, which was renewed effective September 15, 2021, and will run through September 30, 2026, with a matching evaluation period. The summative final report is due March 31, 2027. The demonstration will be referred to as the *Flint Medicaid Expansion Demonstration* (FME Demonstration) in this proposal.

3) *Description and History of the Demonstration*
This FME demonstration was intended to address potential health issues for individuals exposed to the contaminated water in Flint from April 2014 until a date where the water is deemed safe. Work continues to mitigate ongoing exposure to lead in the water supply through proper treatments and lead pipe replacement. While the concentration of lead contaminants has been reduced below federal thresholds, no amount of lead exposure is acceptable. As of December 2021, the water has not been deemed safe since lead pipe replacement is not finished.

The Flint Michigan Section 1115 Demonstration was originally approved for the period 3/3/16-2/28/21, with an extension through 9/14/21. The years 2016-2021 reflected DYs 1-5. The overarching goals of the FME Demonstration were to "improve access to services, expand Medicaid eligibility, and create better health outcomes." These were addressed through the expansion of eligibility by increasing income thresholds, adding a TCM benefit, and eliminating cost-sharing. The review of the FME Demonstration's influence during DYs 1-5 suggests the activities associated with the FME Demonstration supported the state's goals, although some mixed findings were observed as described in the Summative Evaluation Report.

MDHHS submitted a renewal for the FME Demonstration with no program changes in April 2020. The renewal application was designed with the belief that health care coverage for lead exposed individuals needed to continue and the expectation that additional health care needs would become more apparent over time. The request resulted in the 5-year renewal authorization for DYs 6-10 of the Flint, Michigan, 1115 Demonstration, defined as 9/15/21 - 9/30/26.

4) *Description of changes to the demonstration during the approval period, how Evaluation Design altered/augmented to address changes*
The renewal application was submitted with no program changes. However, lessons learned from DYs 1-5 along with review of other FME Demonstration metrics and public comments provided opportunities to augment the evaluation design. Particularly, the hypotheses associated with FME Demonstration required revision, in consideration of data availability and appropriate comparison group(s) selection. The key goals of the renewal application emphasized access to care, expanded eligibility and improved health outcomes. These goals required slight modifications of the original FME Demonstration's reporting. One modification was the recategorization of specific hypotheses. An example of this was moving the lead assessment measure under the Access to Care Domain. We further incorporated the stand-alone TCM Domain from the original FME Demonstration evaluation as part of the renewal's Access to Care Domain. Another modification was to establish a domain to specifically focus on the Expanded Eligibility goal. The renewal evaluation will be further augmented by increasing enrollee input through surveys, inviting additional partners with education subject matter expertise to the team, and increasing focus on operational aspects of FME that may influence the enrollee experience.

5) *Describe the population groups impacted by the demonstration.*
The FME Demonstration is intended to support individuals who were exposed to the contaminated water from April 2016 through a date when the water is deemed safe. The groups targeted in the original FME Demonstration were children up to age 21 and pregnant women. Lead is known to affect brain development, particularly for fetuses and children. Adults would be less likely to experience adverse neurological impacts. Pregnant women were included due to concerns for the developing fetus. Residence in the City of Flint or Genesee County was not a requirement for eligibility. Individuals could have been exposed through child-care, school, or employer locations. In addition to documented water exposure, eligibility criteria included:
    - Increased income threshold to offer coverage to any pregnant woman or child up to age 21 in households with incomes from 212% federal poverty level (FPL) up to and including 400% FPL during the approved timeframe.
    - Any children born to a pregnant woman during the approved timeframe.

B. Evaluation Questions and Hypotheses

1) MDHHS' stated goals for the renewal FME Demonstration were to:
    - improve access to services,
    - expand Medicaid eligibility, and
    - create better health outcomes.

These goals would be addressed through the specific authorizations including expanding eligibility for pregnant women and children up to age 21 having incomes up to 400% FPL. The expanded income threshold would allow individuals who would not normally qualify for Medicaid coverage to do so. The addition of the TCM benefit would support access to services by offering coordination and linkages to needed medical, social, educational, and other types of services. The ability to obtain health care and other services would in turn result in improved health outcomes.

The following domains are offered to translate the FME Demonstration goals into measurable targets. The domains are briefly described with more detail provided in subsequent sections.

Domain 1: Access to services

Hypothesis 1.1: "Enrollees will access services to identify and address physical or behavioral health issues associated with lead exposure at a rate higher than non-enrollees with similar individual and neighborhood characteristics over the duration of the FME Demonstration." The approved FME Demonstration is expected to continue to provide Medicaid coverage and access to health care services to individuals exposed to the contaminated water. The expanded eligibility will provide health care services to individuals who might otherwise be uninsured. Existing Medicaid enrollees would benefit from the additional TCM benefit and the elimination of existing cost-sharing requirements. Further included in approved expenditures is coverage for evaluation of potential lead exposures in homes of eligible enrollees without documentation of elevated blood lead levels. Hypothesis 1.1 will be broken into sub-hypotheses, each focusing on specific preventive care services recommended for children up to age 21 and pregnant women.

Hypothesis 1.2: *"*Enrollees who participate with TCM services will access medical, social, educational, and other services at a rate higher than enrollees with similar individual and neighborhood characteristics who do not participate with TCM services over the duration of the FME Demonstration.*"* The FME Demonstration provides an additional benefit, specifically TCM, to facilitate enrollee access to needed medical, social, educational, and other services. Required elements of TCM have been described in MDHHS policy and include assessments, planning, linkage, advocacy, coordination, referral, monitoring, and

follow-up activities. The rationale for this hypothesis is that TCM participants will have additional help navigating the health care system and securing resources to assist with the consequences of lead exposure. Conversely, those who do not participate with TCM navigate the system independently and may not know about additional supports or services that could be available to them. This hypothesis will also be further subdivided to measure the impact of TCM on enrollees' adherence to recommended health services.

Domain 2: Expand Medicaid Eligibility

Hypothesis 2: "The proportion of new enrollees between 212-400% FPL will increase over the duration of the FME Demonstration representing an increase in the proportion of individuals having health care coverage." MDHHS received authorization to offer Medicaid coverage to individuals at higher income levels and the uptake of this coverage depends on several factors. Potentially eligible individuals and human service organizations responsible for enrollment would need to be aware of the revised qualifications. Also, enrollment processes need to be understood and easily implemented. In addition to standardized quantitative metrics, such as enrollment and disenrollment counts, enrollee and community organization qualitative inputs will inform evaluation of the processes required to participate with the FME Demonstration as well as the degree to which the expanded eligibility represented a new opportunity to obtain health insurance or was used as replacement coverage for other existing forms of health insurance.

Domain 3: Improved Health Outcomes

Hypothesis 3: "Enrollees will have improved health outcomes compared to non-enrollees with similar individual and neighborhood characteristics over the duration of the FME Demonstration." The approved demonstration will provide opportunities for access to health care and additional support leading to improved overall health status and health outcomes for enrollees. Measures such as complete childhood immunization and birth weight will serve as proxies for overall health outcomes. Individualized feedback will be sought through qualitative processes for self-reported health status measures.

2) **Table 1. Domains as the drivers of the FME demonstration, including primary and secondary drivers of the domain**.

Aim (Goal or Objective of the Work)	**Primary Drivers** (Key Drivers: System components or factors contributing directly to achieving aim)	**Secondary Drivers** (Actions, interventions, or lower-level components necessary to achieve the primary driver)
FME Demonstration enrollees will have increased access to selected health care services compared to non-enrollees having similar individual and neighborhood characteristics by 9/30/2026.	Individual having health care insurance	History of lead exposure from contaminated water
		Household income level (FPL%)
	Individual level of cost-sharing for health care services	Household income level (FPL%)
	Ability to navigate health care system	Eligible population knowledgeable about demonstration eligibility and benefits
		TCM and community service organization staff knowledge about FME demonstration eligibility and benefits
	Health literacy	Enrollees seek care in primary care settings rather than urgent or emergent care settings
		Enrollee knowledgeable about recommended preventive care services
The number and proportion of FME demonstration enrollees at 212-400% FPL will increase by 9/30/26 representing an increase in the proportion of individuals having health care coverage.	Eligible population knowledgeable about demonstration eligibility and benefits	FME Demonstration communications and dissemination to potentially affected community
		Community partner(s) knowledgeable about demonstration eligibility and benefits
		Efficient FME demonstration enrollment processes
	Eligible population willing to choose Medicaid	FME Demonstration provides continuity and Stability of coverage
FME Demonstration enrollees will have improved selected health outcomes compared to non-enrollees having similar individual and neighborhood characteristics by 9/30/2026.	Receipt of age-appropriate recommended preventive care services	Enrollee has reduced financial strain associated with having to pay for health care services
	Receipt of care coordination	Enrollee participation with TCM services.
		Enrollee is more confident in managing chronic conditions
	Healthy living environments	Enrollee awareness of the state's redesigned Elevated Blood Lead-Nurse Case Management (EBL-NCM) program and the Lead Safe Home Program (LSHP)

Outlined here are the FME demonstration Domains and the corresponding sub-hypotheses for each.

Domain 1: Access to Services:

Hypothesis 1.1: "FME Demonstration enrollees will access services to identify and address physical or behavioral health issues associated with lead exposure at a rate higher than ***non-enrollees with similar individual and neighborhood characteristics*** over the duration of the demonstration." This hypothesis will focus on comparing rates of selected services among enrollees to rates among selected comparison group(s). The specific services are identified below.

H1.1.1: FME Demonstration enrollees will access age-appropriate well-child exams at a rate higher than non-enrollees with similar individual and neighborhood characteristics over the duration of the demonstration.
H1.1.2: FME Demonstration enrollees will access age-appropriate developmental screening at a rate higher than non-enrollees with similar individual and neighborhood characteristics over the duration of the demonstration.
H1.1.3: FME Demonstration enrollees will access age-appropriate lead testing and follow-up/retesting as indicated at a rate higher than non-enrollees with similar individual and neighborhood characteristics over the duration of the demonstration.
H1.1.4: Pregnant FME Demonstration enrollees will access timely prenatal and postpartum care at a rate higher than non-enrollees with similar individual and neighborhood characteristics over the duration of the demonstration.
H1.1.5: Pregnant FME Demonstration enrollees will access recommended lead testing at a rate higher than non-enrollees with similar individual and neighborhood characteristics over the duration of the demonstration.
H1.1.6: Pregnant FME Demonstration enrollees will participate in the state's Maternal Infant Health Program (MIHP) at a rate higher than non-enrollees with similar individual and neighborhood characteristics over the duration of the demonstration.
H1.1.7: FME Demonstration enrollees will attest to improved health care access as a result of waiver participation.
H1.1.8: FME Demonstration enrollees will attest to satisfaction with their ability to access health care services as a result of waiver participation.
H1.1.9: FME Demonstration enrollees will attest to having evaluation of potential lead exposure in their home if their pipes have not been replaced.

Hypothesis 1.2: *"*FME Demonstration enrollees who participate with TCM services will access medical, social, educational, and other services at a rate higher than FME ***demonstration enrollees with similar individual and neighborhood characteristics who do not participate with TCM*** services over the duration of the demonstration.*"* This hypothesis will focus on

comparing rates of selected services among enrollees who have TCM involvement to rates among enrollees lacking evidence of TCM involvement. The same services in Hypothesis 1.1 will be targeted.

H1.2.1: FME Demonstration enrollees who participate with TCM will access age-appropriate well-child exams at a rate higher than enrollees who do not participate with TCM over the duration of the demonstration.

H1.2.2: FME Demonstration enrollees who participate with TCM will access age-appropriate developmental screening at a rate higher than enrollees who do not participate with TCM over the duration of the demonstration.

H1.2.3: FME Demonstration enrollees who participate with TCM will access age-appropriate lead testing and follow-up/retesting at a rate higher than enrollees who do not participate with TCM over the duration of the demonstration.

H1.2.4: Pregnant FME Demonstration enrollees who participate with TCM will access timely prenatal and postpartum care at a rate higher than pregnant enrollees who do not participate with TCM over the duration of the demonstration.

H1.2.5: Pregnant FME Demonstration enrollees who participate with TCM will access recommended lead testing at a rate higher than pregnant enrollees who do not participate with TCM over the duration of the demonstration.

H1.2.6: Pregnant FME Demonstration enrollees who participate with TCM will participate with MIHP at a rate higher than pregnant enrollees who do not participate with TCM over the duration of the demonstration.

H1.2.7: FME Demonstration enrollees who participate with TCM will attest to improved health care access as a result of waiver participation at a rate higher than enrollees who do not participate with TCM.

H1.2.8: FME Demonstration enrollees who participate with TCM will attest to satisfaction with their ability to access services as a result of TCM participation.

H1.2.9: FME Demonstration enrollees who participate with TCM will attest to having evaluation of potential lead exposure in their home if their pipes have not been replaced as a result of TCM participation.

Domain 2: Expand Medicaid Eligibility

Hypothesis 2: The proportion of new FME Demonstration enrollees between 212-400% FPL will increase over the duration of the demonstration representing an increase in the proportion of individuals having health care coverage.

H2.1: FME Demonstration enrollees between 212-400% FPL will attest to having information regarding expanded Medicaid eligibility resulting in waiver participation.

H2.2: Community partners involved with Medicaid enrollment will attest to awareness of FME Demonstration eligibility and enrollment processes.

H2.3: Community partners involved with Medicaid enrollment will attest to satisfaction with FME Demonstration enrollment processes.
H2.4: FME Demonstration enrollees between 212-400% FPL will attest that the demonstration authorized expanded Medicaid eligibility offered a new opportunity to obtain health care coverage versus serving as a replacement for existing health care coverage.

Domain 3: Improved Health Outcomes

Hypothesis 3: FME Demonstration enrollees will have improved health outcomes compared to non-enrollees with similar individual and neighborhood characteristics over the duration of the demonstration.

Health Outcomes:

H3.1: FME Demonstration enrollees will have improved age-appropriate completed immunization status compared to non-enrollees with similar individual and neighborhood characteristics over the duration of the demonstration. This outcome is included in Domain 3 as opposed to Domain 1 because a driver of health outcomes is the receipt of recommended preventive care services (Table 1).
H3.2: Pregnant FME Demonstration enrollees will have higher birth weights compared to non-enrollees with similar individual and neighborhood characteristics over the duration of the demonstration.
H3.3: FME Demonstration enrollees will report improved health status as a result of the waiver participation.
H3.4: FME Demonstration enrollees will report improved confidence in chronic condition self-management as a result of the waiver participation.

Educational outcomes:

H3.5: FME Demonstration enrollees will have an increased rate of referrals to specialized programs intended to mitigate potential educational and/or behavioral disabilities during childhood (ages 0-21) as a result of waiver participation.

3) Alignments of the hypotheses with overarching goals of the demonstration are described here.

The hypotheses identified in **Domain 1** evaluate the use of specified services including: well-child visits, developmental screening assessments, testing of blood lead levels in pregnant women and children, prenatal and postpartum care, MIHP participation, improved access to care, satisfaction with access to care, and evaluation of potential lead exposure. The majority of these hypotheses reflect services that are endorsed by the US Preventive Services Task Force to promote overall health. The FME Demonstration's goal to improve access to services may be met through a variety of mechanisms as suggested in the driver diagram. Access to health care is influenced by the availability of health insurance to cover the costs associated with obtaining

these services. Costs may be incurred through paying out of pocket in the absence of health insurance as well as having cost-sharing requirements for each instance of service use even with health insurance coverage. In addition to financial aspects of the health care transaction, the cost-benefit analysis in terms of non-financial costs (i.e., time to receive the service, difficulty navigating to an appointment, stress, and mental health) at the individual level influences adherence to these recommendations. The availability of the TCM benefit is expected to assist enrollees in overcoming barriers to seeking care as well as providing information and education on the importance of these services. The evaluation questions for Domain 1 will inform whether the goal to improve access to care was met by measuring enrollee adherence to having the recommended services as evidenced by claims/encounter data. Qualitative data obtained from surveys from enrollees and TCM professionals will provide context to the types of barriers that may impede access and the types of strategies to overcome these barriers.

The hypothesis identified for **Domain 2** is related to the FME Demonstration's goal to increase enrollment by expanding Medicaid eligibility. Authorization to offer Medicaid coverage to individuals at higher income levels was granted along with the elimination of cost-sharing measures. The intention was to eliminate these financial impediments to health care so exposed individuals could seek needed services. However, expanding eligibility criteria is just the first step to increase enrollment. Potentially eligible individuals need to know they may qualify for coverage under the expanded criteria which would require communication and dissemination of this information in a consumable format. Additionally, community partners who support Medicaid enrollment would need to be informed about changes so that they did not assume ineligibility based on prior criteria and/or have the necessary information to operationally enroll individuals. Even with health insurance, there must then be sufficient healthcare providers willing to accept new Medicaid patients. Administrative data along with survey data will be used to address this hypothesis.

The hypothesis identified for **Domain 3** establishes that individuals participating with the FME Demonstration should experience better health outcomes than similar individuals who do not participate. The specific health outcomes represent proxy measures that might reasonably be susceptible to lead exposure among individuals who would be identified as high-risk for lead exposure and represent the target population for the FME Demonstration application. They represent measures of optimum care which presumably would be facilitated through the increased access to health care coverage and the involvement of TCM. While some of these sub-hypotheses may be more accurately described as process measures, the association of each with optimized health status is well documented. The evaluation question associated with improved health outcomes relates to the belief the FME Demonstration addressed barriers to health care so enrollees could seek services as recommended. The financial constraints are

believed to be reduced through the eligibility expansion and elimination of cost-sharing. The availability of TCM professionals to work with enrollees to provide education, secure referrals for care, identify and provide solutions to barriers to care (i.e., transportation, difficulty making appointments) also supports the ability to obtain recommended services to the fullest extent possible. The evaluation team will reach out to enrollees and TCM professionals to obtain qualitative reports on the factors associated with health outcome status and the degree to which the FME demonstration impacted these factors.

Flint City schools are unique because they are composed of both public schools and charter schools totaling 21 distinct districts (Green, 2019). To further elucidate this, Flint has 68 schools within these districts and many of which have very small enrollment counts. Due to administrative circumstances including the water crisis, the State of Michigan and the Genesee Intermediate School District act as the intermediary for all special education for the 21 school districts. For this reason, not all Flint school data, that are necessary to make accurate reports for the progress of school-age children, are publicly available and housed collectively. Additionally, some schools are so small that valuable data on special education and services are often limited. Although individual level education metrics are unavailable due to the Family Educational Rights and Privacy Act (FERPA), it is the intent of the evaluation team to work with several sources such as Michigan State University College of Education, Flint Registry, and Genesee Health System Neurodevelopmental Center of Excellence. The evaluation team plans to aggregate data from the sources listed with administrative data of families enrolled in the waiver. For instance, each entity may have different levels and types of developmental data such IEPs and special services for behavior and educational delays that can inform reasonably accurate benchmarks and trends. Further, enrollee surveys will be designed to capture qualitative child behavioral and educational data to explore the relation to administrative data and the progression of children in Flint. The primary focus of this methodology is to depict close approximations of developmental milestones observed in Flint children exposed to lead in the tap water.

4) The objective of Title XIX was to provide medical and health related services for individuals with low income. The FME Demonstration includes specific authorizations intended to promote the availability of medical and health related services to more individuals at low-income levels through expanded eligibility and elimination of cost-sharing. The evaluation of the FME Demonstration will document the degree to which newly eligible individuals based on expanded criteria are able to seek health care and the degree to which the FME Demonstration resulted in greater health insurance coverage for the affected community. Another benefit of the FME Demonstration is that it offers case management professionals to assist with navigating the health care system. The evaluation will measure whether enrollees received services to a greater degree

with the involvement of these professionals.

C. Methodology

1) *Evaluation Design*

Depending on the types of outcomes, the renewal evaluation will use different designs. For changes of outcomes over time a pre- and post-period with two-group comparison design will be used; and for cross-sectional outcomes, a two-group comparison design will be used. To avoid selection bias, we will not use beneficiaries in Flint who were potentially eligible as the comparison group. Potentially eligible individuals are those residing in the same allowable areas impacted by the water crisis, having the same income levels, and in the same age group(s) but did not choose to enroll in the FME demonstration. This design choice is based on the concern over self-selection bias; there is no reason to believe that we can use statistical methods to control for all systematic differences between FME Demonstration enrollees and non-enrollees. In addition, some statistical methods (e.g., Heckman's selection model, instrumental variables) require researchers to observe factors that are meaningfully related to decisions to participate but are not related to the outcomes to correct the selection bias. Thus, the comparison groups will be selected using a two-step procedure which will first focus on some geographic areas with the larger policy environment like that of Genesee County and then selection of individuals within those areas.

Specifically, in the *first* step, we will use 1) the K-means clustering method to select up to 3 or 4 counties in the Lower Peninsula that are like Genesee County in socioeconomic, demographic, and health characteristics; or 2) a synthetic control (SC) method to construct weighted combinations of counties that had similar trends as that of Flint in the percentage of children under age 6 with elevated blood lead level (EBLL) in the period prior to the expansion.

In the *second* step, we will obtain the administrative claims data and residential census block group or census tract information for Medicaid children up to age 21 and pregnant women in the selected counties together with the data from the target population to estimate a propensity score (PS) for the likelihood of enrolling in the FME demonstration. The estimated PS will be combined with outcome regressions to estimate the average treatment effect on the treated using doubly robust estimation methods (Schuler & Rose, 2017; Zhong et al., 2021).

Details of the two-step procedure and the covariates for estimations are discussed in subsection (ii) of section 6) Analytic Methods.

2) *Target and Comparison Populations*

The FME Demonstration is intended to support individuals who were exposed to contaminated water from April 2016 through a date when the water is deemed safe. The groups targeted in the original FME demonstration were children up to age 21 and pregnant women. Thus, in the renewal evaluation, the same groups of beneficiaries will still be the *target population*.

We may further distinguish existing versus newly enrolled *individuals* in the renewal FME Demonstration. During the first waiver period, the evaluation team considered those at the higher income thresholds of 212-400% FPL would have been considered "newly eligible". These persons did not qualify for existing Medicaid coverage based on current restrictions. The FME demonstration was specifically designed to expand coverage to this group. However, when analyzing the available eligibility data, information regarding income levels was incomplete which compromised the ability to compare the "newly eligible" group to those that would have qualified at the non-FME demonstration levels. Discussions with MDHHS are in process to identify opportunities to obtain complete data to support these comparisons with sufficient rigor.

Additional patterns were noted in the FME demonstration enrollment data suggesting that some individuals could have voluntarily disenrolled from the FME demonstration benefit package but retained other Medicaid coverage. This anomaly is being reviewed with MDHHS representatives to determine if these observations represent errors in the data or potential operational edits. Examples of these patterns are noted in Table 2. Table 2 shows the beneficiaries' enrollment status in Medicaid (where "elig" and "no elig" indicate the person being in Medicaid or not, irrespective of specific FME demonstration enrollment) versus also in the initial FME demonstration (where "fme" indicates the person having at least one enrollment flag). For example, the first row represents individuals who enrolled in the demonstration ("fme") for at least one month in each period from 5/2016 to 4/2020; and among them, 20,307 (subgroup 1) were in Medicaid prior to 5/2016 and 2,619 (subgroup 2) were new to Medicaid starting sometime in 5/2016-4/2017 (e.g, no prior evidence of being a Medicaid beneficiary before 5/2016). The second row of the table represents individuals enrolled in the FME demonstration from 5/2016 to 4/2019, but did not enroll in 5/2019 to 4/2020; and among them, 368 were in Medicaid prior to 5/2016 and 31 were new to Medicaid starting from 5/2016. The rest of the rows of the table read similarly. In total, we found 31,494 existing (before 5/2016) beneficiaries (subgroup 1) and 11,028 new beneficiaries (subgroup 2) who had at least one month enrollment in the FME demonstration in the initial FME demonstration period (2016-2020). Depending on the potential sample sizes of the renewal demonstration, we may target the subgroup of new enrollees.

Ideally, we will assess the impact of the demonstration for those who became eligible through the higher income eligibility criteria as well as individuals already enrolled in Medicaid prior to

the demonstration. The feasibility will depend on the potential new information we may receive from MDHHS on program participation. Using current data, we found 916 (~4%) out of 22,765 enrolled children had income level greater than 212% FPL. The current sample sizes may not allow separate analyses of the impact of the FME Demonstration.

The general criteria for selecting the *comparison populations* will include: 1) children up to age 21 or pregnant women, 2) residing in one of the selected comparison counties using either K-means or synthetic controls method, 3) with estimated propensity scores that overlap with the propensity scores of the target population, and 4) in the appropriate subgroup of the target population defined by the outcome domain metric. Additional criteria for specific outcomes and the justification and limitation of these comparison groups are discussed in subsection (i) of the section 6) Analytic Methods.

Table 2. History of Flint Medicaid expansion (FME) enrollment among existing and new members who were children up to age 21 and pregnant women with a Flint ZIP code or at least one month enrollment in the demonstration after 5/2016.

5/2016 - 4/2017	5/2017 - 4/2018	5/2018 - 4/2019	5/2019 - 4/2020	Subgroup 1 (N=31,494)	Subgroup 2 (N=11,028)
fme	fme	fme	fme	20307	2619
fme	fme	fme	elig	368	31
fme	fme	fme	no elig	1722	282
fme	fme	elig	fme	100	14
fme	fme	elig	elig	351	47
fme	fme	elig	no elig	147	22
fme	fme	no elig	fme	248	64
fme	fme	no elig	elig	42	12
fme	fme	no elig	no elig	1906	615
fme	elig	fme	fme	67	8
fme	elig	fme	elig	16	2
fme	elig	fme	no elig	16	3
fme	elig	elig	fme	48	3
fme	elig	elig	elig	360	46
fme	elig	elig	no elig	87	17
fme	elig	no elig	fme	9	3

fme	elig	no elig	elig	12	5	
fme	elig	no elig	no elig	144	67	
fme	no elig	fme	fme	163	33	
fme	no elig	fme	elig	17	1	
fme	no elig	fme	no elig	64	19	
fme	no elig	elig	fme	6	8	
fme	no elig	elig	elig	24	0	
fme	no elig	elig	no elig	26	5	
fme	no elig	no elig	fme	112	39	**Subtotal**
fme	no elig	no elig	elig	35	9	**Subgroup 2a =**
fme	no elig	no elig	no elig	1977	773	**4747**
elig	fme	fme	fme	654	490	
elig	fme	fme	elig	78	17	
elig	fme	fme	no elig	116	54	
elig	fme	elig	fme	14	2	
elig	fme	elig	elig	88	19	
elig	fme	elig	no elig	21	7	
elig	fme	no elig	fme	11	10	
elig	fme	no elig	elig	6	4	
elig	fme	no elig	no elig	161	142	
elig	elig	fme	fme	226	39	
elig	elig	fme	elig	94	11	
elig	elig	fme	no elig	39	5	
elig	elig	elig	fme	251	46	
elig	elig	no elig	fme	27	9	
elig	no elig	fme	fme	42	4	
elig	no elig	fme	elig	3	0	
elig	no elig	fme	no elig	15	2	**Subtotal**
elig	no elig	elig	fme	8	1	**Subgroup 2b =**

elig	no elig	no elig	fme	61	3	**865**
no elig	fme	fme	fme	211	1027	
no elig	fme	fme	elig	24	60	
no elig	fme	fme	no elig	108	304	
no elig	fme	elig	fme	7	11	
no elig	fme	elig	elig	34	62	
no elig	fme	elig	no elig	17	35	
no elig	fme	no elig	fme	14	30	**Subtotal**
no elig	fme	no elig	elig	7	7	**Subgroup 2c =**
no elig	fme	no elig	no elig	147	461	**1997**
no elig	elig	fme	fme	23	151	
no elig	elig	fme	elig	7	25	
no elig	elig	fme	no elig	4	24	**Subtotal**
no elig	elig	elig	fme	33	147	**Subgroup 2d =**
no elig	elig	no elig	fme	3	13	**360**
no elig	no elig	fme	fme	181	1028	**Subtotal**
no elig	no elig	fme	elig	29	106	**Subgroup 2e =**
no elig	no elig	fme	no elig	71	354	**1488**
						Subtotal
no elig	no elig	elig	fme	26	152	**Subgroup 2f=**
no elig	no elig	no elig	fme	259	1419	**1571**

Footnote: “fme” means the beneficiary had at least one month enrollment in the demonstration program. “elig” means the beneficiary was in the Medicaid program. “no elig” means the beneficiary did not have any enrollment month in Medicaid.

3) *Evaluation Period*

In the initial evaluation, the critical time periods were May 1, 2013 – April 30, 2014, as ‘pre’ water switch period (T1), May 1, 2014 – April 30, 2016, as the ‘pre’ demonstration implementation period (T2), and all subsequent years since the demonstration began in May 2016 as the ‘post’ implementation period (T3). For the renewal evaluation, we will continue with the strategy using each 12-month period, starting from May 2016, as one study period and will include activity from 9/15/21 - 9/30/26.

Timeframe Code	Timeframe Description
T1	Baseline year prior to the water switch (May 1, 2013 – April 30, 2014).

T2	Post water switch, FME not implemented (May 1, 2014 – April 30, 2016).
T3	Post water switch, FME implemented (May 1, 2016 – present).

4) *Evaluation Measures*

As described in the Evaluation Questions and Hypotheses section, the evaluation measures fall in three domains: 1) Access to Services, 2) Eligibility Expansion, and 3) Improved Health Outcomes. We will provide the definitions of each outcome measure here. Summary tables of all measures by domain are available in Appendix A-1.

Domain 1 measures

Age-appropriate well-child exam: the Healthcare Effectiveness Data and Information Set (HEDIS) algorithms will be used to define the following measures.

- The HEDIS well child visits in the first 15 months of life measures "the percentage of children who had between one and six or more well-child visits by the time they turned 15 months of age." The corresponding procedure codes and principal diagnosis in the HEDIS value set will be used to construct the variables.
- The HEDIS well child visits in the third, fourth, fifth and sixth years of life measures "The percentage of members 3-6 years of age who had one or more well-child visits with a PCP (primary care practitioner) during the measurement year." The corresponding procedure codes and principal diagnosis in the HEDIS value set will be used to construct the variables.
- The HEDIS adolescent well-care visits measures "the percentage of enrolled members 12-21 years of age who had at least one comprehensive well-care visit with a PCP or an OB/GYN practitioner during the measurement year." The corresponding procedure codes and principal diagnosis in the HEDIS value set will be used to construct the variables.

Age-appropriate developmental screening: the HEDIS value set procedure codes will be used to construct the following variables.

- The percentage of children screened for risk of developmental, behavioral, and social delays using a standardized screening tool (CPT 96110) in the first three years of life.
- The percentage of children/adolescents 4-17 years of age who had at least one socio-emotional/behavioral screen (CPT 96127) with a primary care practitioner or an OB/GYN practitioner during the measurement year.

Age-appropriate lead testing and follow-up/retesting:

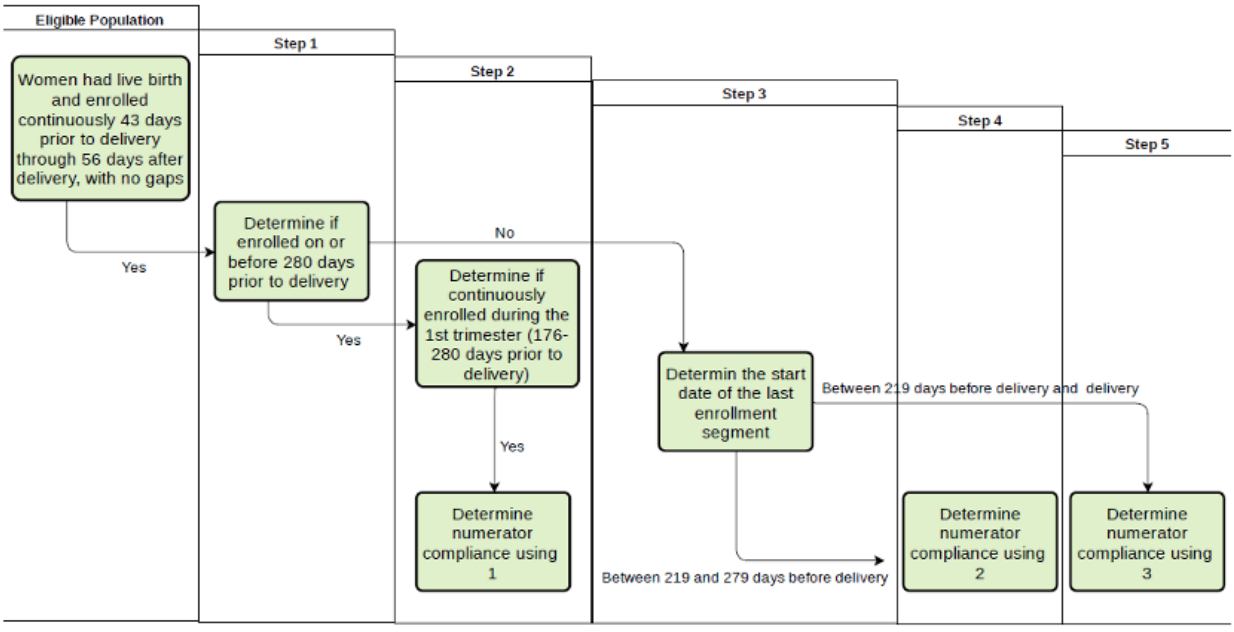
- The modified HEDIS lead screening in children measures "the percentage of children 6 years of age who had 1 or more capillary or venous lead blood test for lead poisoning by

their second birthday." We will use both claims coding and lab data to identify who had a lead test. We will use the Centers for Disease Control and Prevention guideline for the recommended timing for appropriate follow-up as of the evaluation period.

Pregnant enrollees with timely prenatal and postpartum care as defined in HEDIS specifications:

- The percentage of deliveries that received a prenatal care visit in the first trimester, on the enrollment start date or within 42 days of enrollment. Figure A shows the steps to identify the denominator and numerator for this measure.
- The percentage of deliveries that had a postpartum visit on or between 21 and 56 days after delivery. A postpartum visit for a pelvic exam or postpartum care on or between 21 and 56 days after delivery is identified using any of the following criteria: a postpartum visit (Postpartum Visits Value Set); a cervical cytology (Cervical Cytology Value Set); or a bundled service (Postpartum Bundled Services Value Set).

Figure A. The HEDIS procedure defines the percentage of deliveries that received a prenatal care visit in the first trimester, on the enrollment start date or within 42 days of enrollment.



1, Identifying Prenatal Care for Women Continuously Enrolled During the First Trimester
2, Identifying Prenatal Care for Women Not Continuously Enrolled During the First Trimester and find a visit on or between the last enrollment start date and 176 days before delivery
3, Identifying Prenatal Care for Women Not Continuously Enrolled During the First Trimester and find a visit on the enrollment start date or within 42 days after enrollment.

Pregnant enrollees with recommended lead testing:

- We will use the same claim codes and lab data identified for the child lead testing, but the time frame will be specific for pregnant women.

Pregnant enrollees participating in the Maternal and Infant Health Program (MIHP):

- Specific procedure codes for the MIHP in Michigan will be used to identify participants.

Enrollee attestation to improved health care access:

- Survey data questionnaire, for example: Since {Reference date}, the FME demonstration has made it easier to get the health care that I need.
    a. Strongly Agree
    b. Agree
    c. Neutral
    d. Disagree
    e. Strongly Disagree

Enrollee satisfaction with ability to access health care services:

- Survey data questionnaire, for example: Since {Reference date}, how satisfied have you been with your Supports Coordinator?
    a. Very Satisfied
    b. Somewhat Satisfied
    c. Somewhat Dissatisfied
    d. Very Dissatisfied

Evaluation of potential lead exposure at home:

- Environmental Reports from the community
- Survey data questionnaire, for example: Since {Reference date}, did you know that you could have your home evaluated for potential lead exposures? Did you have your home evaluated for potential lead exposures?
- Utilize a variety of analyses to map waterline replacement and associated neighborhood characteristics. We will geocode enrollee addresses and link their survey data with these characteristics, which include but are not limited to water age, previous lead levels in water, area socioeconomic characteristics, vacancy rates, physical disorder. From these connections, we will assess statistical relationships between enrollee health data and their neighborhood context. Subsequent maps will assist in visualizing patterns among these variables.

<u>Domain 2 measures</u>

Enrollee attestation to demonstration information leading to enrollment:

- Data from enrollee survey

Community partner awareness of demonstration enrollment processes:

- Data from community partner survey

Community partner attestation to enrollment processes:

- Data from community partner survey

Enrollee attestation to waiver providing new vs. replacement insurance coverage

- Data from enrollee survey

Domain 3 measures

Age-appropriate immunization status:

- The percentage of children 2 years of age who were fully immunized per the Advisory Committee on Immunization Practices: had four diphtheria, tetanus, and acellular pertussis (DTaP); three polios; one measles, mumps, and rubella (MMR); three haemophilus influenza type B (HiB); three hepatitis B (HepB), one chicken pox (VZV); four pneumococcal conjugates (PCV); one hepatitis A (HepA); two or three rotavirus (RV); and two influenza (flu) vaccines by their second birthday. The children with all 10 immunization records will be counted as part of the numerator.
- The percentage of adolescents 13 years of age who had one dose of meningococcal conjugate vaccine, one tetanus, diphtheria toxoids and acellular pertussis (Tdap) vaccine, and have completed the human papillomavirus (HPV) vaccine series by their 13th birthday. The measure calculates a rate for each vaccine and two combination rates. The adolescents with all 3 vaccines will be counted as part of the numerator.

Birth weights:

- Linked vital records data will be used to find the birth weights.
- Live births with birth weight < 2500 grams will be defined as low birth weight.

Increase in self-reported health status:

- Survey data questionnaire, for example: Since {Reference date}, how would you rate your overall health (both physical and behavioral/emotional)?
    a. Excellent
    b. Very Good
    c. Good
    d. Fair
    e. Poor

Chronic condition self-management confidence:

- Survey data questionnaire, for example: Since {Reference date}, I have access to more resources that help with self-management of my chronic condition(s)
    a. Strongly Agree
    b. Agree
    c. Neutral
    d. Disagree
    e. Strongly Disagree
- Since {Reference date}, I am more confident that I can manage my chronic condition(s) (such as asthma or diabetes).
    a. Strongly Agree
    b. Agree
    c. Neutral
    d. Disagree
    e. Strongly Disagree

Educational Delays:

- Since {Reference date}, have you been told by a doctor or nurse that your child has a behavioral or emotional problem?
    a. Yes
    b. No
    c. Don't know
- Since {Reference date}, has a daycare or school teacher or school nurse told you that your child has a behavioral or emotional problem?
    a. Yes
    b. No
    c. Don't know
    d. Child not school aged/not in school
- Has a daycare or school teacher or school nurse told you that your child has an educational delay requiring special support through an IEP?
    a. Yes
    b. No
    c. Don't know
    d. Child not school aged/not in school

Additional data for educational outcomes will be pursued through a potential partnership with the Genesee Intermediate School District.

5) *Data Sources*

The major data sources for the renewal evaluation will include:

(i) MDHHS (Michigan Department of Health and Human Services) Medicaid enrollment, utilization (claims/encounter) data, Lead Poisoning Prevention Program Data
(ii) TCM program information (administrative data and surveys)
(iii) Michigan Care Improvement Registry (MCIR)
(iv) Enrollee, non-enrollees, and community partner Surveys
(v) Publicly available data (Robert Wood Johnson Foundation County Health Rankings, and census block group and census tract data in American Community Survey)

Each data source and quality control measures are briefly described below.

(i) MDHHS Health Services Data Warehouse – Enrollment and Utilization

MDHHS maintains a data warehouse containing information at an individual level regarding a variety of health-related services and data points. IHP employs staff with the necessary permissions and expertise to access the MDHHS Health Services Data Warehouse (HSDW) and acquire the elements needed to support analyses through an honest broker arrangement. However, despite the storage of a variety of health-related program data in the HSDW, access to these data is controlled by each program.

Specific information contained within the data warehouse includes Medicaid eligibility/enrollment records, final paid Medicaid claims/encounter data, and blood lead program data. While much of the Medicaid claims/encounter data lack clinical care values, the Michigan Childhood Lead Poisoning Prevention Program (MCLPPP) does collect this information.

Reviews of routinely reported information are conducted by MDHHS program and warehouse staff to identify potential issues with data loading or when changes to warehouse tables are made. The evaluation team will not validate the data extracted from the warehouse with primary sources such as medical record reviews. Instead, periodically scheduled conversations between the IHP staff responsible for pulling data and state program and warehouse staff will ensure that relevant fields are captured, and coded variables are correctly interpreted. Data review will be an ongoing, iterative process and continue throughout the duration of the evaluation. Independent review and validation of code used to process data and conduct statistical analyses will be performed by evaluation team statisticians.

(ii) Targeted Case Management Program Information

The supplementary TCM benefit approved in the waiver necessitates additional data sources to support the evaluation beyond the claims/encounter information contained in the HDSW. While the provision of TCM services can be identified through specific procedure codes entered onto billing data, the data elements required to discriminate between specific services is not

available via this administrative data. Although in the initial evaluation, the evaluation team established a Business Associates' Agreement (BAA) with Genesee Health System (GHS) to access their records for purposes of this evaluation, the level of detail needed to support the evaluation was insufficient. The hope was that additional details regarding specific service delivery would be available from this source. Unfortunately, the existing documentation did not permit evaluators to discriminate between referrals to address needs associated with the water exposure versus referrals to address other pre-existing or concomitant social, physical, or behavioral needs. Thus, in the renewal evaluation we will not assess TCM referrals. Instead, enrollee surveys will provide additional data regarding the TCM benefit in Domain 1. More descriptions of the survey are in (iv), and details of the sampling design and analysis are in section 6) Analytic Methods.

(iii) Michigan Care Improvement Registry (MCIR)

In the renewal evaluation we will use MCIR data to complement the HSDW data to evaluate the participants' immunization status. A recent report showed that vaccine coverage declines among most children at milestone ages in May 2020 compared to previous May estimates (Bramer et al., 2020). We will use future MCIR publications as benchmarks to assess the coverage in enrollees.

(iv) Enrollee, non-enrollees, dis-enrollees, and community partner surveys

In the initial evaluation, we found that Flint community members preferred a web-based survey to the paper- or telephone-based survey. Initially, we adopted a longitudinal survey strategy and followed a random sample of enrollees over a 3-year period. However, the low response rates made longitudinal analyses difficult. In addition, the beneficiaries get in and out of Medicaid frequently (Table 2) and the COVID-19 pandemic will also affect the sampling frames. Thus, in the renewal evaluation, we will conduct repeated cross-sectional surveys each year.

MSU is working with MDHHS to clarify apparent voluntary disenrollment that was identified during the first evaluation cycle. If these patterns are confirmed, the following options will be pursued. To address these potential issues of non-enrollment and disenrollment, we will explore the potential of using Medicaid eligibility data to identify two additional groups for surveys. First, children up to age 21 in Medicaid who have at least one residential ZIP code in the list of Flint water service qualified ZIP codes, but no FME demonstration enrollment will be the basis for non- FME demonstration enrollees. Second, children up to age 21 who had at least one FME demonstration benefit flag in the year prior but do not have the benefit flag in the current evaluation year (e.g., the second row in Table 2 showing individuals who were enrolled for three years but not in year 4) will serve as the basis for FME demonstration disenrollees.

For details for the sampling frame, sampling procedure and analysis plan, see the subsection (iv) in section 6) Analytic Methods.

The focus on operational aspects of the FME demonstration for the impact on enrollment requires input from community partners who are involved with Medicaid eligibility verification and enrollment processes. These community partners will provide information through surveys and key informant interviews on topics such as awareness of revised eligibility for the demonstration and ease of processing enrollments.

(v) Publicly Available Data

American Community Surveys (ACS)

Recent literature on social determinants of health in general and the environmental correlates to elevated blood levels in Flint specifically suggests that social and built environments are important predictors for health outcomes (Sadler et al., 2017). Lacking individual-level data on these factors, we will link enrollees' addresses geocoded to the census tract or census block group level with the ACS to find proxies to the neighborhood socioeconomic backgrounds.

Childhood Opportunity Index (COI)

COI is a multidimensional depiction of the neighborhood beyond the population composition and socioeconomic conditions at the census tract level for 2010 and 2015. It captures "neighborhood resources and conditions that matter for children's healthy development" in a single metric. The index focuses on contemporary features of neighborhoods that are affecting children. It is based on 29 indicators spanning 3 domains: education, health and environment, and social and economic." (*Child Opportunity Index 2.0 Database*, n.d.)

Social Vulnerability Index (SVI)

SVI ranks census tracts on 15 social factors and groups them into four related themes: socioeconomic (income, poverty, employment, education), household composition and disability (age, single parenting, disability), minority status and language (race, ethnicity, English-language proficiency), and housing and transportation (housing structure, crowding, vehicle access). Each census tract receives a ranking for each theme, and an overall ranking within the state (*CDC/ATSDR Social Vulnerability Index 2018 Database Michigan.*, 2021).

County Health Rankings & Roadmaps (CHR&R)

CHR&R "provides data, evidence, guidance, and examples to build awareness of the multiple factors that influence health and support community leaders working to improve health and increase health equity". The Rankings are unique in their ability to measure the health of nearly every county in all 50 states, and are complemented by guidance, tools, and resources designed to accelerate community learning and action" (*How Healthy Is Your County?*, n.d.). The data elements will be used primarily in the first step of the comparison county selection procedure and listed under the "Covariates" section in 6).

Area Deprivation Index (ADI)

Researchers at the University of Wisconsin-Madison created the ADI using ACS (5-year data) at the block group level. It is "composed of 17 education, employment, housing-quality, and poverty measures originally drawn from long-form Census data … updated to incorporate more recent ACS data" (Kind & Buckingham, 2018).

Michigan Medicaid Statewide Weighted HEDIS Measures

Although the Michigan Medicaid summary HEDIS statewide report reflects statewide estimates rather than county level information, these reports will be reviewed to provide additional context to the results obtained through the renewal evaluation. However, the evaluation team is cognizant of the fact that several of the targeted measures reported by the statewide summary are based on hybrid (administrative and medical record review) reporting method by health plans. Hybrid rates are known to exceed administrative rates.

6) *Analytic Methods*

This section describes the identification strategies for the causal effects of interest in Domains 1-3 in the renewal evaluation plan. The analytic strategies depend on the period of comparisons (one year or longitudinal), the type of outcomes (continuous or discrete), the data source (administrative or survey), and the availability of a comparison group. The general hypothesis is driven by the intent of the FME Demonstration and services provided by the TCM. We will focus on the average treatment effect on the treated (ATT), which asks the question: "what would the difference in outcomes be had the FME Demonstration enrollees not participated in the program?" This section is divided into five subsections: subsection (i) describes whether there will be a potential comparison group for each outcome measure, subsection (ii) describes the two-step procedure to select potential comparison groups, subsection (iii) clearly lays out the assumptions and statistical methods that will be employed to identify and estimate the effects of interest, subsection (iv) presents the enrollee, non-enrollee, dis-enrollee survey sampling designs and analysis plans, and subsection (v) discusses potential sensitivity and robustness analyses.

Throughout this section we will refer to the renewal FME Demonstration as the program (first level intervention), the FME Demonstration enrollees as enrollees, the TCM services as the treatment (second level intervention), and the TCM recipients as the participants. Enrollees who do not use the TCM services will be called non-participants and the term non-enrollees will be reserved for beneficiaries who are potentially eligible for the FME Demonstration but do not enroll. The term comparison may refer to either comparison with enrollees or comparison with participants, depending on the context. The comparison group(s) for enrollees will be selected from other counties; and the comparison group(s) for participants will be selected from non-participants.

(i) Availability of Potential Comparison Groups

The causal inference problem is a missing data problem because the outcomes of the enrollees/participants if they had not enrolled in the program or received the treatment are never observable. To estimate the causal effect of any intervention, we must rely on the outcomes of an appropriate comparison group or multiple comparison groups as the counterfactual outcomes of the treated group.

The ideal comparison group should be comprised of individuals who are not exposed to the intervention, are like the enrollees in confounding factors (i.e., determinants of both enrollment and the outcome of interest), observed or unobserved, and "exposed to the same policy environment." (Contreary et al., 2018) However, the environment in Flint is unique due to the water crisis and the FME demonstration is only designed for individuals exposed to the crisis. All other Medicaid programs for children and pregnant women in Michigan have lower income limits (217% for children and 200% for pregnant women), thus the enrollees with income higher than these levels (approximately 5% of all enrollees in the initial FME demonstration period) will not have a natural comparison group.

Other Medicaid children and pregnant women with income higher than that allowed by non-FME demonstration programs may also have access to health care when their medical expenses equal or exceed their deductible (formerly known as spend-down) amount. The spend-down population may be closest to the high income (over 217%) enrollees in the FME demonstration. For the spend-down population we also may be missing some of their healthcare services through other insurance, which could also be true for enrollees. In addition, the initial FME demonstration enrollees whose income was higher than 200% federal poverty level (FPL) accounted for only approximately 5% of the total number of enrollees, and most of the initial FME demonstration enrollees had income levels similar to that of the selected comparison group in the initial evaluation.

Thus, the best strategy to approximate a 'same policy environment' is to first focus on some geographic areas with a larger policy environment like that of Genesee County (whose county seat and largest city is Flint). Genesee County is the 5th most populous county in Michigan, with approximately one-quarter enrolled in Medicaid each year. We chose a two-step procedure to select comparison groups when possible (see below).

Table 3 displays the outcomes of each domain by the availability of potential comparison groups. In general, outcomes measured using claims/encounter data may have a potential comparison group and outcomes assessed through surveys will not have a comparison group. When possible, the overarching criteria for a comparison group include: 1) children up to age 21 or pregnant women, 2) residing in one of the selected comparison counties using the two-step procedure, 3) with estimated propensity scores that overlap with the propensity scores of

the target population, and 4) in the appropriate subgroup of the target population defined by the outcome domain metric.

Table 3. Evaluation outcomes with or without a potential comparison group

Presence of comparison	Domain	Outcomes	Data sources
Yes	1	Age-Appropriate well-child exam	Enrollment and claims
		Age-appropriate developmental screening	Enrollment and claims
		Age-appropriate lead testing and follow-up/retesting	Enrollment, claims and lab tests
		Timely prenatal and postpartum care	Enrollment and claims
		Lead testing during pregnancy	Enrollment, claims and lab tests
		Participation in MIHP	Enrollment and claims
	3	Age-appropriate immunization status	Enrollment, claims and lab tests
		Birth weight	Enrollment, claims and vital records
No	1	Enrollee attestation of access	Survey
		Enrollee satisfaction	Survey
	2	Enrollee attestation of dissemination	Survey
		Community partner awareness	Survey
		Community partner attestation	Survey
	3	Self-reported health status	Survey
		Confidence in chronic disease management	Survey
		Education/behavior outcomes	Survey

(ii) Two-step Procedure for Selecting Comparison Groups

In the renewal evaluation we will continue the use of a pre- and post-period with two-group comparison design for changes of outcomes over time, and a two-group comparison design for cross-sectional outcomes, but the comparison populations in both designs will be refined. Previously, we used all pregnant women and children up to age 21 in Saginaw County as the comparison group. Saginaw County was selected using the K-means method. However, our experience revealed some limitations of this approach (detailed in the publication of an unrelated project) (Strutz et al., 2021). Thus, in the renewal evaluation for outcomes in Table 3 with enrollees as the target population, we will select up to 3 or 4 comparison counties from the Lower Peninsula and use individual- and census tract- or census block group-level data in the selected counties and the enrollees together to estimate propensity scores for enrolling in the FME demonstration. When the target population is the treated population (i.e., utilizing the TCM services) for outcomes in Table 3, we will compare the participants with non-participants estimating another propensity score.

As we outlined in the Evaluation Design section, the two-step procedure is as follows. In the *first* step, we will use 1) the K-means clustering method to select up to 3 or 4 counties in the Lower Peninsula that are like Genesee County in socioeconomic, demographic, and health characteristics; or 2) a synthetic control (SC) method to construct weighted combinations of counties that had similar trends as that of Flint in the outcomes in the period prior to the expansion. In the *second* step, we will obtain the administrative claims data and residential census block group or census tract information for Medicaid children up to age 21 and pregnant women in the selected counties together with the data from the target population to estimate a propensity score (PS) for the likelihood of enrolling in the FME demonstration. The estimated PS will be combined with outcome regressions to estimate the average treatment effect on the treated using double robust estimation methods. For different evaluation hypotheses we will consider different potential covariates (e.g., for age-appropriate immunization outcomes in children we may consider exact matching on age and sex; and for prenatal care measures we may consider matching on previous pregnancy history which can be identified through linked vital records). Below we provide some details of these steps.

The K-means clustering method

This is a common unsupervised learning method that we exploit to find other counties in Michigan like Genesee County in important socioeconomic, demographic, educational, physical environment, and health indicators. Traditionally, the K-means method aims at segregating a population into subgroups (clusters) such that the within cluster variation is minimized. The K-means solution is sensitive to the initial centroids of clusters and the final number of clusters, thus, we take advantage of these properties and use different initial centroids and different number of clusters many times (1,000 in each scenario) and find 3 or 4 counties in the Lower Peninsula that are most often clustered in the same subgroup as Genesee County.

The variables used in the K-means method are the key for success in this selection strategy. Table 4 shows health outcomes, health behavior, clinical care, social economic environment, and physical environment used by the CHR&R to rank counties in the US. We will choose relevant confounding characteristics that may influence the outcome of interest and the presence of potential programs (a total of 48 variables, but subject to change and selection in the renewal evaluation with updated years of data) under the assumption that counties similar in these characteristics as Genesee County will have a similar policy environment (Bradley et al., 2020).

Table 4. County Health Ranking measures and source data used in the initial evaluation*

Health Outcomes		
Measure	**Description**	**Source**
Poor or fair health	Percentage of adults reporting fair or poor health (age-adjusted)	Behavioral Risk Factor Surveillance System
Poor physical health days	Average number of physically unhealthy days reported in past 30 days (age-adjusted)	Behavioral Risk Factor Surveillance System
Poor mental health days	Average number of mentally unhealthy days reported in past 30 days (age-adjusted)	Behavioral Risk Factor Surveillance System
Low birthweight	Percentage of live births with low birthweight (< 2500 grams)	National Center for Health Statistics - Natality files
Infant mortality	Average infant death per 10,000 live births	Health Indicators Warehouse
Frequent physical distress	Percent population experiencing frequent physical distress	Behavioral Risk Factor Surveillance System
Frequent mental distress	Percent population experiencing frequent mental distress	Behavioral Risk Factor Surveillance System
Health Behaviors		
Measure	**Description**	**Source**
Food environment index	Index of factors that contribute to a healthy food environment, 0 (worst) to 10 (best)	USDA Food Environment Atlas, Map the Meal Gap
Teen births	Teen birth rate per 1,000 female population, ages 15-19	National Center for Health Statistics - Natality files
Food insecurity	Percent population with food insecurity	Map the Meal Gap
Access to healthy foods	Percent population with limited access to healthy foods	USDA Food Environment Atlas
Drug induced deaths	Number of deaths induced by drug overdose	Michigan Health Statistics
Insufficient sleep	Percent population with reported insufficient sleep	Behavioral Risk Factor Surveillance System
Clinical Care		
Measure	**Description**	**Source**
Uninsured	Percentage of population under age 65 without health insurance	Small Area Health Insurance Estimates
Primary care physicians	Ratio of population to primary care physicians	Area Health Resource File/American Medical Association

Dentists	Ratio of population to dentists	Area Health Resource File/National Provider Identification file
Uninsured adults	Percentage of population age 18 and above without health insurance	Small Area Health Insurance Estimates
Uninsured children	Percentage of population under age 18 without health insurance	Small Area Health Insurance Estimates
Health care costs	Average health care costs	Dartmouth Atlas of Health Care
Other primary care providers	Ratio of primary care physicians to per 10,000 population	CMS, National Provider Identification file
Social and Economic Environment		
Measure	**Description**	**Source**
High school graduation	Percentage of ninth-grade cohort that graduates in four years	EDFacts
Some college	Percentage of adults ages 25-44 years with some post-secondary education	American Community Survey
Unemployment	Percentage of population ages 16 and older unemployed but seeking work	Bureau of Labor Statistics
Children in poverty	Percentage of children under age 18 in poverty	Small Area Income and Poverty Estimates
Income	Median household income	Small Area Income and Poverty Estimates
Income inequality	Ratio of household income at the 80th percentile to income at the 20th percentile	American Community Survey
Children in single-parent households	Percentage of children that live in a household headed by single parent	American Community Survey
Children eligible for free lunch	Percent of children that are eligible for free lunch or lunch at the reduced price	National Center for Education Statistics
Violent crime	Number of reported violent crime offenses per 100,000 population	Uniform Crime Reporting – FBI and Michigan State Police
Homicide	Number of reported homicides per 100,000 population	CDC (Centers for Disease Control) WONDER mortality data
Property crime	Number of reported property-related crimes per 100,000 population	Uniform Crime Reporting – FBI and Michigan State Police
Physical Environment		
Measure	**Description**	**Source**

Air pollution - particulate matter	Average daily density of fine particulate matter in micrograms per cubic meter (PM2.5)	Environmental Public Health Tracking Network
Drinking water violations	Indicator of the presence of health-related drinking water violations. 1 - indicates the presence of a violation, 0 - indicates no violation	Safe Drinking Water Information System
Severe housing problems	Percentage of households with at least 1 of 4 housing problems: overcrowding, high housing costs, or lack of kitchen or plumbing facilities	Comprehensive Housing Affordability Strategy (CHAS) data
Demographics		
Measure	**Description**	**Source**
Population	Population Sizes	Census Population Estimates
Children	Percent population below 18 years of age	Census Population Estimates
Elderly	Percent population 65 and older	Census Population Estimates
Race-ethnicity	Percent population Non-Hispanic African American	Census Population Estimates
Race-ethnicity	Percent population American Indian and Alaskan Native	Census Population Estimates
Race-ethnicity	Percent population Asian	Census Population Estimates
Race-ethnicity	Percent population Native Hawaiian/Other Pacific Islander	Census Population Estimates
Race-ethnicity	Percent population Hispanic	Census Population Estimates
Race-ethnicity	Percent population non-Hispanic white	Census Population Estimates
Proficient in English	Percent population not proficient in English	American Community Survey
Female	Percent population females	Census Population Estimates
Rural	Percent population in rural areas	Census Population Estimates

* Information taken from County Health Ranking Reports https://www.countyhealthrankings.org

The K-means algorithm is as follows: 1) Randomly assign a number from 1 to K to each county where K is the assumed number of clusters; 2) compute the cluster centroid (defined by the feature means in each cluster) and reassign each county to the cluster whose centroid is closest using, say, the Euclidean distance to itself; and 3) iterate until the cluster assignments stop changing.

One issue of the K-means clustering method is that the resulting assignments depend on the random starting point. The K-means algorithm does not guarantee to lead to global minimum, so the starting points should be varied to examine the end partitioning. The second issue of the K-means algorithm is that sometimes a variable with high variability would dominate the cluster analysis. A common solution is to standardize variables, but there are multiple ways of standardizing variables and standardization could also hide the true groupings in the data (Schaffer & Green, 1996; Steinley, 2006). This is a case-by-case decision depending on the type of data and the nature of the groups. Finally, the optimal choice of the final number of clusters, K, is not always clear.

We will test solutions for 3 to 10 clusters for S iterations (say S=5,000) with randomly selected starting centroid values. We will use scree plots to visualize the curve of the within sum of squares (WSS) or its logarithm for all cluster solutions and a kink in the curve, if present, will be the number K. We will use the GAP statistics to estimate and confirm the optimal number of clusters (Tibshirani et al., 2001). If the scree plot does not produce any obvious kink point, or if the kink point suggested by the scree plot does not agree with the optimal solution based on the Gap statistic, we will use the number of clusters K* that passes the Gap statistic test. We will then generate S random starting values to run the K-means algorithm for K* clusters. Next, we count how many times a county is assigned to the same cluster as Genesee County out of the S iterations. The 3 or 4 counties most often clustered together with Genesee County will be chosen as the comparison counties. We will use the five standardization methods in addition to the z-score to calculate the distances between the selected and Genesee County using the Euclidean, L1, Canberra and 1-correlation distance measures based on the subset of relevant covariates from Table 3. If the majority of the distance measures suggest that the selected counties are closer to Genesee County than unselected counties, then the K means selection will be accepted (Schaffer & Green, 1996).

As an illustration, in the initial evaluation, the Gap statistic based on the z-score standardized features in Table 4 indicated the 68 Lower Peninsula counties were best grouped in 9 clusters. Using the 9-cluster solution, we ran the K-means algorithm with 5,000 random starting values and Saginaw County was clustered within the same group as Genesee County 4,405 times, followed by Muskegon and Calhoun with 4,183 and 4,124 times, respectively. Thus, Saginaw County was the chosen county in the initial evaluation.

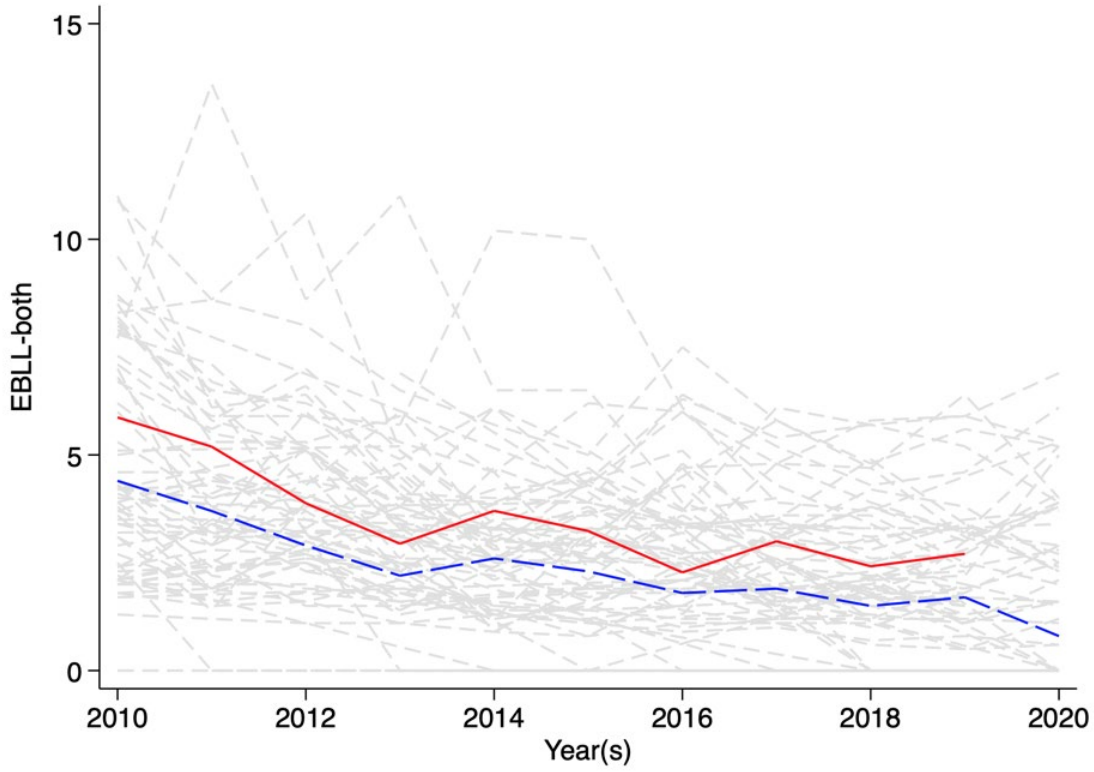
In K-means analyses, if all variables are standardized then clustering based on correlation (similarity) is equivalent to that based on squared distance (dissimilarity). Therefore, as a robustness check, we will run the K-means twice, with and without z score standardization of all features.

The Synthetic Control Method

The second approach the renewal evaluation will consider is the synthetic control (SC) method (Abadie et al., 2010). Since no single county is like Genesee County in all characteristics under consideration, we will explore using a weighted combination of counties as controls. The SC idea is to impute a counterfactual outcome of Genesee as a weighted average of other counties (not including the upper peninsula counties). The weights are computed by minimizing a vector distance between Genesee and other counties over a set of pre-treatment covariates that are predictive of the outcome.

The evaluation has numerous outcomes and the SC method, unlike the K-means method, needs to be conducted separately for each outcome to estimate the weights specific to that outcome. Here we use elevated blood lead levels (EBLL) for illustration. Even though this is not an outcome for the renewal evaluation, it may be informative as to what this approach can and cannot achieve and the required data elements and assumptions for the method to be valid. First, we extracted county-level and ZIP code-level data for the proportion of children < 6 years of age who were tested and had EBLL from 2010 to 2020, using the Michigan Childhood Lead Poisoning Prevention Program (MCLPPP) annual reports and data portal. Figure B shows the EBLL of children in the 11 ZIP code approved by the Flint waiver demonstration (red solid line), Genesee County (blue dashed line), and the rest of the 67 counties in the Lower Peninsula (light gray dashed lines, excluding the city of Detroit). We can see a more pronounced uptick of the trend in Flint than that in Genesee County in 2014.

Figure B. Percent children under age 6 with elevated blood lead level (EBLL) using either capillary or venous test. The red line is for children in Flint and blue dashed line is for children in Genesee County. (Note: The City of Detroit is excluded from the Wayne County data.)



We then use the 2010-2019 variables in Table 4 of the 68 counties in the Lower Peninsula of Michigan to construct an SC county for Flint (Genesee County is removed in this analysis and the county covariates are used for the 11 ZIP codes) using parametric and non-parametric SC methods (Cerulli, 2020). Table 5 shows that in 12 of the specifications of predictors and models, Saginaw was selected 10 times as one of the top 4 counties with the largest weights in the synthetic controls, followed by Wayne (6 times), Jackson (5 times) and St. Clair (5 times), Muskegon (4 times) and Monroe (4 times). Overall, the unstandardized predictors and non-parametric models had smaller biases and smaller root mean-squared prediction error (RMSPE).

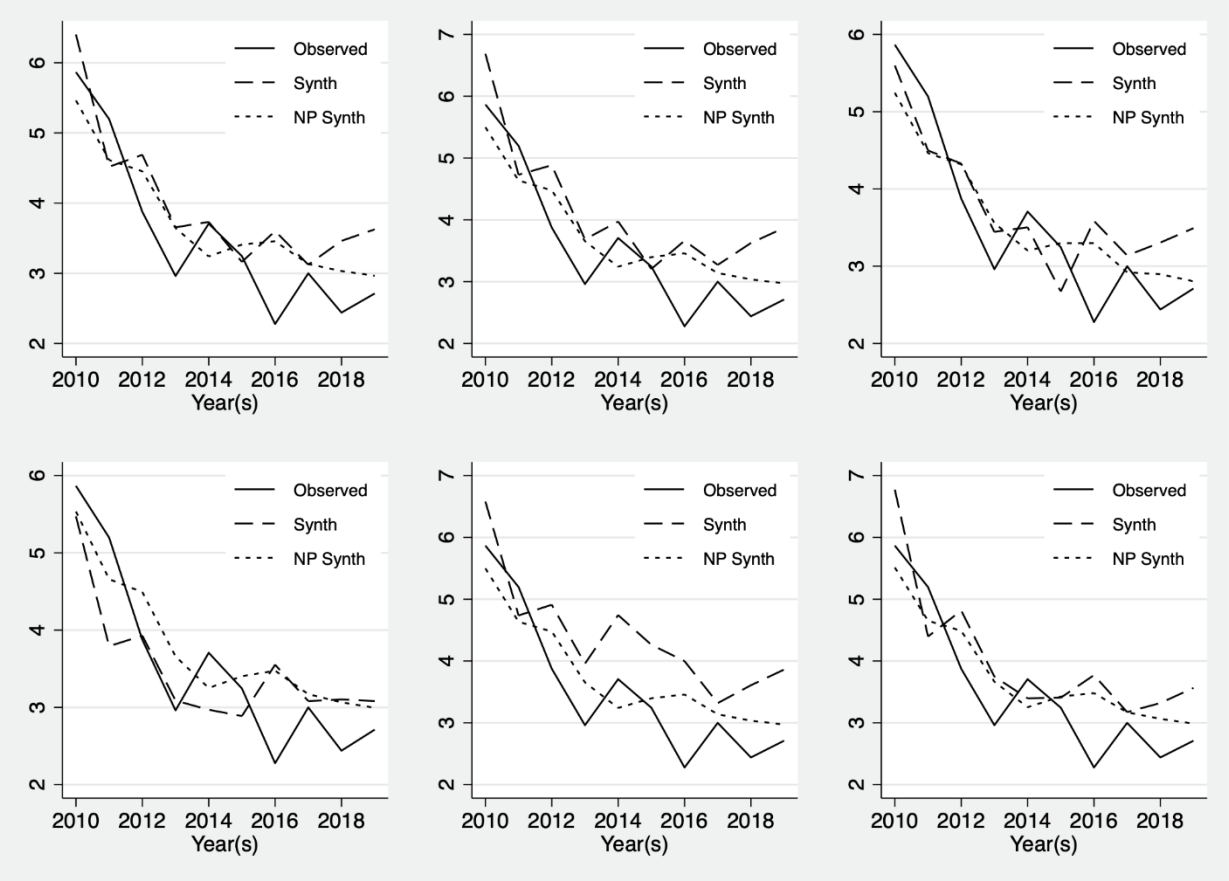
Figure C shows that the specifications in the top row and first column (unstandardized covariates and non-parametric model) tracks the Flint data the best prior to 2016; and all other specifications fall short in some aspect. The selected top counties in the best case are St. Clair, Saginaw, Jackson, and Monroe (row 2 of Table 5).

Table 5. The parametric and non-parametric* synthetic control models' root mean-squared prediction error (RMSPE), 4 counties with the highest weights, and average bias in the pre-treatment period. (Note: Wayne does not include the City of Detroit.)

Predictors***	Model	RMSPE	4 Highest weight counties	Bias in years prior to 2016
$unstd	Parametric	0.745	Saginaw, Wayne, Muskegon, St. Clair	-0.218
	Non-parametric	0.581	St. Clair, Saginaw, Jackson, Monroe	0.005
$std	Parametric	0.851	Saginaw, Muskegon, Wayne**	-0.385
	Non-parametric	0.581	Jackson, Monroe, St. Clair, Saginaw	-0.009
$unstd+$std	Parametric	0.668	Wayne, Muskegon, Cass**	0.135
	Non-parametric	0.548	Ottawa, Livingston, Oakland, Washtenaw	0.128
$unstd-pc10	Parametric	0.709	Wayne, Saginaw, Calhoun, St. Joseph	0.452
	Non-parametric	0.586	Saginaw, Monroe, Calhoun, Jackson	-0.025
$std-pc10	Parametric	1.032	Saginaw, Muskegon, Wayne, Lenawee	-0.724
	Non-parametric	0.581	Jackson, Monroe, St. Clair, Saginaw	-0.009
$unstd+$std-pc10	Parametric	0.828	Saginaw, St. Clair, St. Joseph, Wayne	-0.281
	Non-parametric	0.590	Jackson, Saginaw, Bay, Calhoun	-0.024

*Almost all counties have equal weights.
** Only 3 counties have non-zero weights
*** The list of variables in the unstandardized and standardized covariates are not the same.

Figure C. The parametric and non-parametric synthetic controls compared with the observed trends in Flint. The 6 panels from left to right and top to bottom are based on the following formats: 1) unstandardized variables, 2) standardized variables, 3) all variables, 4) first 10 principal components (PCs) of the unstandardized variables, 5) first 10 PCs of the standardized variables, 6) first 10 PCs of all variables.



The above illustration shows some disadvantages of the SC method. First, the method requires re-calibration of weights for each outcome because different counterfactual weights may be required to construct an SC that is similar in the respective hypothesis to be tested. Summary measures of the outcomes and time-varying covariates that are predictive of each outcome at the county level (and Flint) for each hypothesis many years prior to the FME demonstration expansion will be required. Extracting all the required data from the HSDW will be time-consuming and the predictive power of the covariates in the CHR&R may be weak. This is the main reason for which we prefer to use the K-means method or the nearest-neighbors method to find comparison counties in the first step of the evaluation.

Second, the SC method works best if the outcomes of interest have clear trends over time before and after the intervention. However, many of the outcomes in the renewal evaluation

have stable distribution and there is no compelling evidence of the change in the slope of the trends after the intervention.

Given these limitations, we will consider the SC method only as the secondary approach in selecting comparison counties.

Propensity Score (PS) Estimation Protocol

PS for Enrollment:

Once the comparison counties are selected, we will find children up to age 21 and pregnant women in these counties who meet the criteria in the appropriate subgroups of the target enrollees defined by the outcome measures. Their data will be combined with the data of the target enrollees to estimate a PS for the probability of enrollment in the FME demonstration.

We will use a logistic regression to estimate the PS when the number of covariates is not large as the literature shows that in this case a logistic regression performs as well as some machine learning algorithms (P. Austin et al., 2013). The covariates in the estimation of the PS have been traditionally selected using some statistical variable selection methods that are significant predictors of the intervention. However, more recent literature has shown that doing so may compromise causal effect estimation and inference. In addition, confounding variables should be the ones that can block the biasing pathways (e.g., the backdoor path from the intervention to the outcome), not just predictors of the intervention. Thus, we will not follow the traditional variable selection approach to estimate the PS. Instead, we will focus on examining covariate balance using the weighted standardized differences between enrollees and comparison persons using the inverse probability weighting (IPW) by the PS (P. C. Austin & Stuart, 2015). Note: because we are not using the PS matching estimators, we will not use the usual paired standardized differences to examine balance in covariates. It will be an iterative process until all weighted standardized differences are smaller than 0.1. If for some covariates this cannot be achieved, we will use them in outcome regression adjustment (ORA) to control residual confounding.

PS for Participation:

For hypotheses involving comparing FME demonstration enrollees who used the TCM services (i.e., participants) and FME demonstration enrollees who did not use the service (i.e., non-participants), we will estimate the PS for the probability of utilizing the TCM services with a logistic regression using data from all FME demonstration enrollees in the subpopulations relevant to the hypotheses. The protocol will be the same as the one above.

Covariates for PS and ORA models:

We will use individual-level and census tract- or census block group-level variables relevant to each hypothesis as covariates for the PS and ORA models for the double-robust estimation methods. For example, for age-appropriate well-child exam, we will use children's age, sex, race/ethnicity, and the COI at the census tract level as covariates; and for timely prenatal care, we will use women's age, race/ethnicity, pregnancy history in vital records, comorbidity index constructed using claims data, and the SVI at the census tract level or the ADI at the census block group level as covariates.

(iii) Identification Assumptions and Statistical Methods

The double-robust methods, incorporating both outcome and treatment mechanisms, can minimize the influence of model misspecification and outperform g-formula and IPW methods in both point and confidence interval estimation (Díaz, 2020; Le Borgne et al., 2021; Luque-Fernandez et al., 2018; Schuler & Rose, 2017; Zhong et al., 2021). With the assistance of machine learning techniques, these methods can further mitigate the influence of model misspecification (Kreif & DiazOrdaz, 2019). However, it is important to understand the underlying causal and statistical assumptions needed for these methods. All assumptions (conditional exchangeability for emulating randomization, sequential exchangeability for censoring and compliance, consistency, positivity, and stable unit of treatment value) are inherently untestable (Hernán & Robins, 2020). We will provide potential steps we may take to guard against violations of assumptions.

For the Pre-Post Two-Group (PPTG) Comparison Design:

This design will be used when the effect of interest is the change in outcomes over time. It is essentially the difference in differences (DID) design, which can be implemented using repeated cross sections or panel data, i.e., different individuals over time or the same individuals (Stuart et al., 2014). As Medicaid beneficiaries tend to go in and out of enrollment (churning), we will use repeated cross sections. In the initial evaluation, the critical time periods were May 1, 2013 – April 30, 2014, as 'pre' water switch period (T1), May 1, 2014 – April 30, 2016, as the 'pre' demonstration implementation period (T2), and all subsequent years since the demonstration began in May 2016 as the 'post' implementation period (T3). The two pre-periods, T1 and T2, will be used separately when feasible and the post-period will be the evaluation years. The "treated" population in the pre-periods will include individuals in the Flint area designated by the 11 zip codes and meeting the age restriction or pregnancy condition. We have extracted data from 2013 to 2021 for the initial evaluation. Very recent literature on DID methodologies suggests that having multiple pre-treatment periods may help satisfy the parallel trend assumption crucial to the analysis (Callaway & Sant'Anna, n.d.; Wooldridge, 2021). However, if we use T1 or T2 as the pre-period, we will not be able to take advantage of the multi-year data before the FME Demonstration.

For the Two-Group (TG) Comparison Design:

This design will be used when the effect of interest is the difference between the target population and the comparison population. This design is especially vulnerable to unmeasured confounding. We will perform sensitivity analysis and provide the E-values of the estimates and the confidence limits (VanderWeele & Ding, 2017). For both the PPTG and TG designs, we will appropriately consider the nesting of observations within individuals if present, and the nesting of individuals within clusters (census tract or census block group).

(iv) Enrollee, Non-enrollee, Dis-enrollee Survey Sampling Design and Analysis Plan

For each evaluation period, we will use the first 6-month FME demonstration enrollment data from MDHHS to identify FME demonstration enrollees who had at least one TCM benefit flag to form the sampling frame for the FME demonstration enrollee survey. Previously we used a longitudinal survey design but had poor response rates. In addition, the FME demonstration enrollees displayed the 'churning' phenomenon as in the general Medicaid population (as seen in Table 2). Thus, in the renewal evaluation, we will conduct repeated cross-sectional surveys and each sample will be representative of the FME demonstration enrollees of that year who had at least one month of enrollment (assuming the second 6 months enrollees are similar in characteristics). We will use a stratified (age, race, geography) unequal probability sample and the sample size will be based on 5% margin of error for the key question related to enrollment attestation and satisfaction.

We are interested in exploring the feasibility of surveying FME demonstration non-enrollees. For the non-enrollee survey, we would use the same first 6-month enrollment data from MDHHS to find the "potentially" eligible beneficiaries who 1) were up to age 21, 2) had one residential ZIP code in the list of 11 ZIP codes used by MDHHS to determine eligibility, 3) had no prior enrollment history, and 4) had income level >212%. These individuals would form the sampling frame of the FME demonstration non-enrollee survey. Since we have the age, race/ethnicity, and geographic information for these beneficiaries, we would use the same stratified unequal probability of sampling to select the survey samples and the sample size consideration will be based on the key question related to non-FME demonstration enrollment (e.g., main reason). However, we remain concerned about the traditional Medicaid income limits compromising the ability to identify sufficient individuals.

For the FME Demonstration dis-enrollee survey, we will use the previous year's enrollment data from MDHHS to identify individuals who had enrolled for at least 6 months in that year but had not enrolled in the first 6 months of the current evaluation year, and these individuals will form the sampling frame of the FME demonstration dis-enrollee survey. The sampling design and sample size consideration will be the same as in the two cases above.

For all three surveys, we will use Stata's svy prefixed commands for generalized linear models with proper sampling design features to estimate the parameters of interest.

(v) Potential Sensitivity and Robustness Analyses

Because we will employ double-robust estimation methodologies and not use statistical significance as a criterion to select covariates, we expect some degree of robustness of our statistical estimation. However, as we mentioned above, all observational studies suffer the potential bias for unmeasured confounding and endogenous selection, and we will perform quantitative bias analysis, i.e., sensitivity analysis, in these two categories. First, for binary outcomes, the E-value mentioned above is defined as "the minimum strength of association that an unmeasured confounder would need to have with both the treatment and the outcome to fully explain away a specific treatment-outcome association, conditional on the measured covariates" (VanderWeele & Ding, 2017). A large E-value implies that considerable unmeasured confounding would be needed to explain away an effect estimate. A small E-value implies little unmeasured confounding would be needed to explain away an effect estimate. Second, assessing selection bias is more difficult. We will use a negative-control idea to gauge the potential severity of the selection bias. We will use an outcome measure that is unlikely or assumed to have no reason to be affected by the program or the TCM services, e.g., say, accidental injury, and use the models for the analysis on this outcome. If our modeling strategy is sound and if the negative control outcome is not influenced by the program or the TCM services, then we should see zero treatment effects. On the contrary, if we found significant treatment effect on a negative control outcome, then we may suspect model misspecifications in some stage of our analysis, from selection of comparison sample to propensity score estimation, and to outcome regression modeling. If we find zero effect on the negative outcomes, then we will be more reassured of the evaluation results.

D. Limitations

Limitations associated with the planned evaluation include difficulty identifying individuals who would be eligible for the program at the higher income levels but have not come through the enrollment process. The FME Demonstration enrolled cohort further presents challenges due to missing data after enrollment if the FME demonstration enrollment is secondary coverage. We will attempt to document these participants who have other forms of health care coverage through documentation collected by the state for coordination of benefit processing which may give us additional strata for comparison. To better understand the participation process, we plan to use the survey mechanism and key-informant interviews.

The impacts of the COVID pandemic will continue to be felt during this renewal cycle as a full return to 'normalcy' has not yet been achieved. Nationally, ambulatory care visits dropped approximately 60% in 2020, according to some reports, although visits appeared to have rebounded in 2021 (Mehrotra et al., 2020). Care delivery shifted from an in-person model to one using telemedicine and virtual visits to a much greater degree. However, the key component of the demonstration, i.e., TCM, was not authorized for telemedicine delivery.

Evaluating changes in health care visits is a ripe topic for investigation. We will compare trends observed in our data against state and national estimates as those data become available through literature.

E. Attachments

1) *Independent Evaluator.*
The Michigan State University Institute for Health Policy (MSU-IHP) has been involved with health care quality improvement, program evaluation, and health services research for over two decades. MSU's College of Human Medicine maintains a community campus in Flint, Michigan, with associated clinical practices and faculty who may interact with MDHHS regarding Medicaid policies or reimbursement. The evaluation team at MSU-IHP, however, operates independently of the clinical practices and has no business interest in the expansion of Medicaid and the provision of services to the affected population. Thus, we believe no conflict of interest exists to conducting the evaluation and are willing to provide a "No Conflict of Interest" statement.

With specific regards to the FME demonstration, MSU-IHP was involved with the evaluation conducted on DYs 1-5. We are prepared to leverage the processes and tools that were successful in the first round and have identified lessons learned that will serve to augment the evaluation for the renewal period (DYs 6-10). The evaluation team includes expertise in Medicaid operations and Data Warehouse, Program Evaluation, Biostatistics and Epidemiology, Health Economics, Health Disparities, Nursing, Women and Children's Health, and Geospatial Epidemiology. Current members of the team include:

- Sabrina Ford, PhD, Institute for Health Policy & Department of Obstetrics and Gynecology, College of Human Medicine, MSU
- Nicole Jones, PhD, Division of Public Health, College of Human Medicine, MSU
- Joan Ilardo, PhD, LMSW; Office of Research, College of Human Medicine, MSU
- Zongqiang Liao, PhD, Institute for Health Policy, College of Human Medicine, MSU
- Zhehui Luo, PhD; Department of Epidemiology and Biostatistics, College of Human Medicine, MSU
- Kathleen Oberst, PhD, RN; Institute for Health Policy, College of Human Medicine, MSU
- Richard Sadler, PhD, MPH; Division of Public Health, College of Human Medicine, MSU

2) *Evaluation Budget.*
Budget submitted follows MDHHS fiscal year master agreement timelines. Start date of 01/01/22 reflects project start date in FY23 master agreement amendment. A budget for implementing the evaluation shall be provided with the draft Evaluation Design. It will

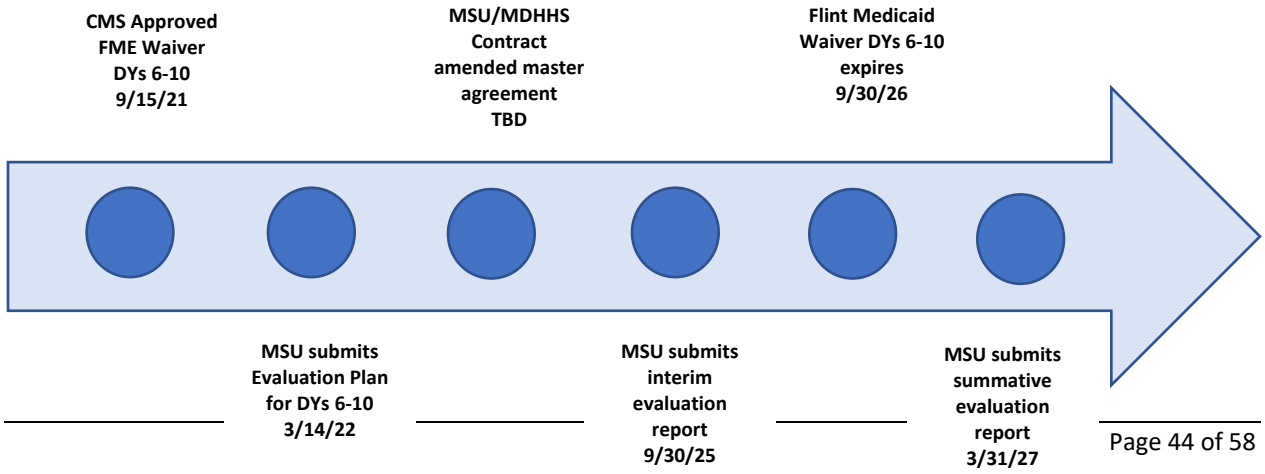
include the total estimated cost, as well as a breakdown of estimated staff, administrative, and other costs for all aspects of the evaluation. Examples include but are not limited to: the development of all survey and measurement instruments; quantitative and qualitative data collection; data cleaning 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 or if CMS finds that the draft Evaluation Design is not sufficiently developed. Refer to Table 6 below.

Table 6. Evaluation Budget

MSU Institute for Health Policy						
Flint Lead Waiver Renewal						
01/01/22-09/30/26						
	Year One	Year Two	Year Three	Year Four	Year Five	TOTAL
Salaries	180,309	245,221	250,125	255,128	260,338	1,191,121
Fringe Benefits	46,149	63,090	66,233	67,838	69,517	312,827
Supplies/Materials	6,200	6,200	6,200	6,200	6,200	31,000
Survey Expense	55,000	60,000	61,736	62,049	75,864	314,649
Graduate Assistant Tuition	24,000	24,720	25,462	26,226	26,226	126,634
Indirect Expense @ 20%	62,332	79,846	81,951	83,488	87,629	395,246
Total Expenses	373,990	479,077	491,707	500,929	525,774	2,371,477

Timeline and Major Milestones

- 9/15/21 - CMS approved Flint Medicaid Waiver DYs 6-10
- 3/14/22 - MSU submits Evaluation Plan for 9/15/21 - 9/30/26 to CMS
- TBD - MSU contract amended to MSU/MDHHS master agreement
- 9/30/25 - MSU submits Interim Evaluation Report
- 9/30/26 - Flint Medicaid Waiver DYs 6-10 expires
- 3/31/27 - MSU submits summative evaluation report



References

Abadie, A., Diamond, A., & Hainmueller, J. (2010). Synthetic Control Methods for Comparative Case Studies: Estimating the Effect of California's Tobacco Control Program. *Journal of the American Statistical Association*, *105*(490), 493–505. https://doi.org/10.1198/jasa.2009.ap08746

Austin, P. C., & Stuart, E. A. (2015). Moving towards best practice when using inverse probability of treatment weighting (IPTW) using the propensity score to estimate causal treatment effects in observational studies. *Statistics in Medicine*, *34*(28), 3661–3679. https://doi.org/10.1002/sim.6607

Austin, P., Tu, J., Ho, J., Levy, D., & Lee, D. (2013). Using methods from the data-mining and machine-learning literature for disease classification and prediction: A case study examining classification of heart failure subtypes. *Journal of Clinical Epidemiology*, *66*(4).

Bradley, K., Heeringa, J., Pohl, R. V., Reschovsky, J. D., & Samra, M. (2020, October). Selecting the Best Comparison Group and Evaluation Design: A Guidance Document for State Section 1115 Demonstration Evaluations. Centers for Medicare and Medicaid Services. Retrieved September 2022, from https://www.medicaid.gov/medicaid/section-1115-demo/downloads/evaluation-reports/comparison-grp-eval-dsgn.pdf

Bramer, C. A., Kimmins, L. M., Swanson, R., Kuo, J., & Vranesich, P. (2020). Decline in Child Vaccination Coverage During the COVID-19 Pandemic—Michigan Care Improvement Registry, May 2016–May 2020. *MMWR. Morbidity and Mortality Weekly Report*, *69*. https://doi.org/10.15585/mmwr.mm6920e1

Callaway, B., & Sant'Anna, P. (n.d.). Difference-in-Differences with multiple time periods. *Journal of Economics*, *225*(2), 200–230. https://doi.org/10.1016/j.jeconom.2020.12.001.

Case Studies in Environmental Medicine (CSEM) Lead Toxicity. (n.d.). Agency for Toxic Substances and Disease Registry.

CDC/ATSDR Social Vulnerability Index 2018 Database Michigan. (2021, August 27). Centers for Disease Control and Prevention/ Agency for Toxic Substances and Disease Registry/ Geospatial Research, Analysis, and Services Program. https://www.atsdr.cdc.gov/placeandhealth/svi/data_documentation_download.html

Cerulli, G. (2020). Nonparametric synthetic control using the npsynth command. *The Stata Journal: Promoting Communications on Statistics and Stata*, *20*(4), 844–865. https://doi.org/10.1177/1536867X20976315

Child Opportunity Index 2.0 database. (n.d.). Diversitydatakids.Org. 2021. Waltham, MA: Institute for Child, Youth and Family Policy, Heller School for Social Policy and Management, Brandeis University. Retrieved February 7, 2022, from https://data.diversitydatakids.org/dataset/coi20-child-opportunity-index-2-0-database

Contreary, K., Bradley, K., & Chao, S. (2018). *Best Practices in Causal Inference for Evaluations of Section 1115 Eligibility and Coverage Demonstrations*. Mathematica Policy Research White Paper.

Díaz, I. (2020). Machine learning in the estimation of causal effects: Targeted minimum loss-based estimation and double/debiased machine learning. *Biostatistics (Oxford, England)*, *21*(2), 353–358. https://doi.org/10.1093/biostatistics/kxz042

Flint & Genesee County, Michigan—Community Health Needs Assessment. (2019). Hurley Medical Center.

Flint, MI. (n.d.). Data USA. Retrieved February 1, 2022, from https://datausa.io/profile/geo/flint-mi/#:~:text=Between%202018%20and%202019%2C%20the,various%20types%20of%20health%20insurance.

"Flint Michigan Section 1115 Demonstration" approval and Special Terms and Conditions. (n.d.). US Department of Health & Human Services, Centers for Medicare & Medicaid Services and Michigan Department of Health and Human Services.

Green, Erica L. (2019). *Flint's Children Suffer in Class After Years of Drinking the Lead-Poisoned Water.* The New York Times. https://www.nytimes.com/2019/11/06/us/politics/flint-michigan-schools.html

Hernán, M., & Robins, J. (2020). Causal Interference: What If. *Boca Raton: Chapman &Hall/CRC*.

How Healthy is your County? (n.d.). County Health Rankings & Roadmaps. Retrieved February 7, 2022, from https://www.countyhealthrankings.org/county-health-rankings-roadmaps

Kind, A., & Buckingham, W. (2018). *Making Neighborhood-Disadvantage Metrics Accessible—The Neighborhood Atlas*. New England Journal of Medicine.

Kreif, N., & DiazOrdaz, K. (2019). Machine Learning in Policy Evaluation: New Tools for Causal Inference. *Oxford Research Encyclopedia of Economics and Finance*. https://doi.org/10.1093/acrefore/9780190625979.013.256

Le Borgne, F., Chatton, A., Léger, M., Lenain, R., & Foucher, Y. (2021). G-computation and machine learning for estimating the causal effects of binary exposure statuses on binary outcomes. *Scientific Reports*, *11*(1), 1435. https://doi.org/10.1038/s41598-021-81110-0

Luque-Fernandez, M., Schomaker, M., Rachet, B., & Schnitzer, M. (2018). Targeted maximum likelihood estimation for a binary treatment: A tutorial—PubMed. *Statistics in Medicine*. https://pubmed.ncbi.nlm.nih.gov/29687470/

Mehrotra, A., Chernew, M., Linetsky, D., Hatch, H., & Cutler, D. (2020, May 19). *The Impact of the COVID-19 Pandemic on Outpatient Visits: A Rebound Emerges | Commonwealth Fund*. The Commonwealth Fund. https://www.commonwealthfund.org/publications/2020/apr/impact-covid-19-outpatient-visits

Sadler, R. C., LaChance, J., & Hanna-Attisha, M. (2017). Social and Built Environmental Correlates of Predicted Blood Lead Levels in the Flint Water Crisis. *American Journal of Public Health*, *107*(5), 763–769. https://doi.org/10.2105/AJPH.2017.303692

Schaffer, C. M., & Green, P. E. (1996). An Empirical Comparison of Variable Standardization Methods in Cluster Analysis. *Multivariate Behavioral Research*, *31*(2), 149–167. https://doi.org/10.1207/s15327906mbr3102_1

Schuler, M. S., & Rose, S. (2017). Targeted Maximum Likelihood Estimation for Causal Inference in Observational Studies. *American Journal of Epidemiology*, *185*(1), 65–73. https://doi.org/10.1093/aje/kww165

Section 1115 Demonstrations. (n.d.). US Department of Health & Human Services, Centers for Medicare and Medicaid Services. Retrieved April 14, 2016, from https://www.medicaid.gov/Medicaid-CHIP-Program-Information/By-Topics/Waivers/1115/Section-1115-Demonstrations.html

Steinley, D. (2006). K-means clustering: A half-century synthesis. *The British Journal of Mathematical and Statistical Psychology*, *59*(Pt 1), 1–34. https://doi.org/10.1348/000711005X48266

Strutz, K., Luo, Z., Raffo, J., Meghea, C., Vander Meulen, P., & Roman, L. (2021). *Determining County-Level Counterfactuals for Evaluation of Population Health Interventions: A Novel Application of K-Means Cluster Analysis Kelly L. Strutz, Zhehui Luo, Jennifer E. Raffo, Cristian I. Meghea, Peggy Vander Meulen, Lee Anne Roman, 2021*. https://doi.org/10.1177/00333549211030507

Stuart, E. A., Huskamp, H. A., Duckworth, K., Simmons, J., Song, Z., Chernew, M., & Barry, C. L. (2014). Using propensity scores in difference-in-differences models to estimate the effects of a policy change. *Health Services & Outcomes Research Methodology*, *14*(4), 166–182. https://doi.org/10.1007/s10742-014-0123-z

Tibshirani, R., Walther, G., & Hastie, T. (2001). Estimating the Number of Clusters in a Data Set via the Gap Statistic. *Journal of the Royal Statistical Society. Series B (Statistical Methodology)*, *63*(2), 411–423.

VanderWeele, T. J., & Ding, P. (2017). Sensitivity Analysis in Observational Research: Introducing the E-Value. *Annals of Internal Medicine*, *167*(4), 268–274. https://doi.org/10.7326/M16-2607

Wooldridge, J. M. (2021). *Two-Way Fixed Effects, the Two-Way Mundlak Regression, and Difference-in-Differences Estimators* (SSRN Scholarly Paper ID 3906345). Social Science Research Network. https://doi.org/10.2139/ssrn.3906345

Zhong, Y., Kennedy, E. H., Bodnar, L. M., & Naimi, A. I. (2021). AIPW: An R Package for Augmented Inverse Probability-Weighted Estimation of Average Causal Effects. *American Journal of Epidemiology*, *190*(12), 2690–2699. https://doi.org/10.1093/aje/kwab207

Sub-hypotheses details for each Domain

Hypothesis 1 is made of 2 subgroups.

- H1.1 focuses on comparing enrollee services to non-enrollees (i.e. comparison group)
- H1.2 focuses on the impact of TCM services on enrollees adhering to recommended care, thus comparing TCM participants to non-participants *among those who are enrolled in the waiver*. The belief is that participants who take advantage of these services are better educated both as to the importance of preventive care and offered direct assistance and support in navigating the health care system. Thus, we repeat the targeted measures from H1.1 with further sub-categorization among all enrollees comparing TCM participants to non-participants. If sufficient data is available, we intend to explore whether a dose-response effect of TCM visits can be identified. Qualitative data from enrollees and TCM professionals will provide context to the findings.

Domain 1: Access to Services							
Hypothesis 1.1: "Enrollees will access services to identify and address physical or behavioral health issues associated with lead exposure at a rate higher non-enrollees with similar individual and neighborhood characteristics over the duration of the demonstration." *Hypothesis 1.2: "Enrollees who participate with TCM services will access medical, social, educational, and other services at a rate higher than enrollees with similar individual and neighborhood characteristics who do not participate with TCM services over the duration of the demonstration."*							
Characteristic	**Detail Description**	**Detail Description**	**Detail Description**	**Detail Description**	**Detail Description**	**Detail Description**	**Detail Description**
Measure Title	Well Child Visits in the First 15 months of Life	Well Child visits in the Third, Fourth, Fifth and Sixth Years of Life	Adolescent Well-Care Visits	Developmental Screening in the First Three Years of Life	Socio-emotional/Behavioral Screening for Children 4-17 years of age	Lead Screening in Children	Follow-up of elevated blood lead level
Measure Description	The percentage of children 15 months old who had the recommended number of well-child visits with a PCP during their first 15 months of life.	The percentage of children 3-6 years of age who had one or more well-child visits with a primary care provider during the measurement year.	The percentage of children/ adolescents 12-21 years of age who had at least one comprehensive well-care visit with a primary care provider or an OB/GYN practitioner during the measurement year.	The percentage of children screened for risk of developmental, behavioral, and social delays using a standardized screening tool in the first three years of life.	The percentage of children/ adolescents 4-17 years of age who had at least one socio-emotional/behavioral screen (CPT 96127) with a primary care provider or an OB/GYN practitioner during the measurement year.	The percentage of children 2 years of age who had 1 or more capillary or venous lead blood test for lead poisoning by their second birthday.	The percentage of children with elevated blood lead levels having retests according to recommended timeframes established by MDHHS Lead Policy.
NQF Number	1392	1516	n/a	1448	n/a	n/a	n/a
Measure Steward	National Committee for Quality Assurance	National Committee for Quality Assurance (Child Core Set)	National Committee for Quality Assurance (Child Core Set)	Oregon Health & Science University	n/a	National Committee for Quality Assurance	Early and Periodic Screening, Diagnostic, and Treatment (EPSDT)-

							CMS/American Academy of Pediatrics
Numerator	This measure has 7 discrete numerators: • # Children who received 0 well-child visits • # Children who received 1 well-child visit • # Children who received 2 well-child visits • # Children who received 3 well-child visits • # Children who received 4 well-child visits • # Children who received 5 well-child visits • # Children who received 6 or more well-child visits	This measure has 1 discrete numerator: • At least one well-child visit with a primary care provider	This measure has 1 discrete numerator: • At least one comprehensive well-care visit with a PCP or an OB/GYN practitioner during the measurement year.	This measure has 4 discrete numerators: • # Children who had screening for risk of development, behavioral and social delays using a standardized screening tool that was documented by their first birthday. • # Children who had screening for risk of development, behavioral and social delays using a standardized screening tool that was documented by their second birthday. • # Children who had screening for risk of development, behavioral and social delays using a standardized screening tool that was documented by their third birthday. • # Children who had screening for risk of development, behavioral and social delays using a standardized screening tool that	This measure has 1 discrete numerator: • At least one socio-emotional/behavioral screen with a PCP or an OB/GYN practitioner during the measurement year.	# of children with at least one lead capillary or venous blood test on or before the child's second birthday.	# of children with elevated blood lead levels having re-testing with specified timeframes.

				was documented by their first, second, or third birthday. (Combination estimate)			
Denominator	Children 15 months old during the measurement period.	This measure has 1 discrete denominator: • Children 3-6 years of age during the measurement period.	This measure has 1 discrete denominator: • Children/ adolescents 12-21 years of age during the measurement period.	This measure has 4 discrete denominators (respectively): • # Children who turn 1 by the end of the measurement period. • # Children who turn 2 by the end of the measurement period. • # Children who turn 3 by the end of the measurement period. • # Children who turn 1 or 2 or 3 by the end of the measurement period.	This measure has 1 discrete denominator: • Children/adolescents 4-17 years of age during the measurement period.	# of children who turn 2 years old during the measurement period.	# of children with elevated blood lead levels during the measurement period.
Baseline Value(s)	DY 1-5 results	DY 1-5 results	DY 1-5 results	DY 1-5 results	DY 1-5 results	DY 1-5 results	DY 1-5 results
Sampling Methodology	No sampling – plan to use 100% available claims/encounter data	No sampling – plan to use 100% available claims/encounter data	No sampling – plan to use 100% available claims/encounter data	No sampling – plan to use 100% available claims/encounter data	No sampling – plan to use 100% available claims/encounter data	No sampling – plan to use 100% available claims/encounter data matched with MCIR and Childhood Lead Prevention Program	No sampling – plan to use 100% available claims/encounter data
Anticipated Data Source	Administrative claims/encounters in the MDHHS data warehouse	Administrative claims/encounters in the MDHHS data warehouse	Administrative claims/encounters in the MDHHS data warehouse	Administrative claims/encounters in the MDHHS data warehouse	Administrative claims/encounters in the MDHHS data warehouse	Administrative claims/encounters, MCIR, and Childhood Lead Screening Data in the MDHHS data warehouse	Administrative claims/encounters in the MDHHS data warehouse linked to state lead

							screening and TCM monitoring data

Domain 1: Access to Services (continued)							
Hypothesis 1.1: "Enrollees will access services to identify and address physical or behavioral health issues associated with lead exposure at a rate higher non-enrollees with similar individual and neighborhood characteristics over the duration of the demonstration." *Hypothesis 1.2: "Enrollees who participate with TCM services will access medical, social, educational, and other services at a rate higher than enrollees with similar individual and neighborhood characteristics who do not participate with TCM services over the duration of the demonstration."*							
Characteristic	**Detail Description**	**Detail Description**	**Detail Description**	**Detail Description**	**Detail Description**	**Detail Description**	**Detail Description**
Measure Title	Timeliness of Prenatal Care	Postpartum Care	Lead screening in pregnancy	MIHP Participation	Enrollee Attestation for Improved Access to Care	Enrollee satisfaction with Medicaid expansion coverage	Evaluation of potential lead exposure in home
Measure Description	Percentage of Medicaid live birth deliveries between February 4 of the year prior to the measurement period and February 3 of the measurement period	The percentage of deliveries that had a postpartum visit on or between 21 and 56 days after delivery.	The percentage of pregnant women screened for elevated blood lead levels during pregnancy.	The percentage of deliveries participating with the Maternal Infant Health Program.	Surveyed enrollees will agree or strongly agree with a statement acknowledging the Medicaid program as one method for improving access to health care.	Surveyed enrollees ranking of their health care coverage using 0-10 scale (0=worst health care possible, 10=best health care possible)	Surveyed enrollees reporting accessing lead evaluation service offered through TCM
NQF Number	1517	1517	n/a	n/a	n/a	--	n/a
Measure Steward	National Committee for Quality Assurance	National Committee for Quality Assurance	American Congress of Obstetricians and Gynecologists	n/a	Agency for Healthcare Research and Quality – Consumer Assessment of Healthcare Providers and Systems (AHRQ-CAHPS) Question Modification	AHRQ CAHPS Question Modification	n/a
Numerator	Percentage of deliveries that received a prenatal care visit as a patient in the first trimester or within 42 days of enrollment.	Percentage of deliveries that had a postpartum visit on or between 21 and 56 days after delivery.	Percentage of deliveries that received 1 or more capillary or venous lead blood test during pregnancy.	Percentage of deliveries receiving 1 or more visit with MIHP during pregnancy or after birth.	Number of respondents who report they "agree "or "strongly agree" with a statement about Medicaid improving health care access. *Sample questions:* "In the last 6 months, how often was it easy to get the care, tests, or	Mean of health care scores provided by survey enrollees. *Sample question:* "Using any number from 0 to 10, where 0 is the	Proportion of households evaluated for potential lead exposure provided by survey enrollees.

					treatment you needed?" (never/sometimes/usually/always) "Overall, enrolling in the Medicaid expansion made it easier to get the health care that I needed" (strongly agree to strongly disagree)	worst health care possible and 10 is the best health care possible, what number would you use to rate all your health care"	
Denominator	Medicaid deliveries of live births between February 4 of the year prior to the measurement period and February 3 of the measurement period.	Medicaid live birth deliveries between February 4 of the year prior to the measurement period and February 3 of the measurement period.	Medicaid live birth deliveries between February 4 of the year prior to the measurement period and February 3 of the measurement period.	Medicaid deliveries of live births between February 4 of the year prior to the measurement period and February 3 of the measurement period.	Number of survey participants.	Number of survey participants.	Number of survey participants.
Baseline Value(s)	DY 1-5 results	DY 1-5 results	DY 1-5 results	DY 1-5 results	DY 1-5 results	DY 1-5 results	n/a
Sampling Methodology	No sampling – plan to use 100% available claims/encounter data	No sampling – plan to use 100% available claims/encounter data	No sampling – plan to use 100% available claims/encounter data	No sampling – plan to use 100% available claims/encounter data	Random/weighted sampling	Random/weighted sampling	Random/weighted sampling
Anticipated Data Source	Administrative claims/encounters in the MDHHS data warehouse linked to Vital Records	Administrative claims/encounters in the MDHHS data warehouse linked to Vital Records	Administrative claims/encounters in the MDHHS data warehouse linked to Vital Records data	Administrative claims/encounters in the MDHHS data warehouse linked to MIHP visit and TCM Monitoring data	Enrollee survey	Enrollee survey	Enrollee survey

Domain 2: Expand Medicaid Eligibility			
Hypothesis 2: "The proportion of new enrollees between 212-400% FPL will increase over the duration of the demonstration representing an increase in the proportion of individuals having health care coverage."			
Characteristic	**Detail Description**	**Detail Description**	**Detail Description**
Measure Title	Enrollee attestation to demonstration leading to enrollment	Community partner awareness	Community partner attestation
Measure Description	Surveyed enrollees will agree or strongly agree with a statement acknowledging the waiver implementation provided information leading to enrollment.	Interviewed community partners … will agree or strongly agree with a statement acknowledging waiver eligibility, increased income limits, elimination of cost-sharing.	Interviewed community partners … will agree or strongly agree with a statement acknowledging that process to enroll individuals in the Flint Waiver is easy, they have contacts available if there are questions, the process is sufficiently automated for timely enrollment
NQF Number	n/a	n/a	n/a
Measure Steward	Agency for Healthcare Research and Quality – Consumer Assessment of Healthcare Providers and Systems (AHRQ-CAHPS) Question Modification	n/a	n/a
Numerator	Number of respondents who report they "agree" or "strongly agree" with the statement. *Sample questions:* I received information about the Flint Medicaid Waiver that told me how to find out if I qualify. The information I received about the Flint Medicaid Waiver was helpful to let me know that I could qualify for Medicaid. The information I received about the Flint Medicaid Waiver told me about special benefits only available to people enrolled in the waiver. The information I received about the Flint Medicaid Waiver told me about extra help that was available to help me get needed services.	Number of partners knowledgeable Sample questions: I/my agency received information about the Flint Medicaid Waiver eligibility guidelines. I/my agency received information about cost-sharing elimination so that I could inform potential enrollees.	Number of partners reporting positive experience with enrollment process Sample question: The information I/my agency received about the Flint Medicaid Waiver was helpful to understand the mechanisms to check eligibility and enroll new members, I am able to use existing systems with helpful prompts to check potential eligibility and enroll new individuals,
Denominator	Number of survey participants.	Number of partners interviewed	Number of partners interviewed
Baseline Value(s)	--	--	--
Sampling Methodology	Random/weighted sampling	n/a	n/a

Anticipated Data Source	Enrollee survey		Key informant interviews/surveys with Targeted partners		Key informant interviews/surveys with Targeted partners		
Domain 3: Improved Health Outcomes							
Hypothesis 3: "Enrollees will have improved health outcomes compared to non-enrollees with similar individual and neighborhood characteristics over the duration of the demonstration."							
Characteristic	**Detail Description**	**Detail Description**	**Detail Description**	**Detail Description**	**Detail Description**	**Detail Description**	**Detail Description**
Measure Title	Childhood Immunization Status	Immunizations for Adolescents	Low Birth Weight Rate	Enrollee Self-Reported Health Status	Enrollee Self-Reported Confidence of Chronic Condition Management	Enrollee Self-Report Cognitive and Education Status	Childhood Independent Educational Plan (IEP)
Measure Description	Percentage of children 2 years of age who had 4 diphtheria, tetanus and acellular pertussis (Tdap), polio (IPV); one measles, mumps and rubella (MMR); three H influenza type B (HiB): three hepatitis B (HepB); one chicken pox (VZV); four pneumococcal conjugate (PCV); one hepatitis A (HepA); two or three rotavirus (RV); and two influenza (flu) vaccines by their second birthday.	Percentage of adolescents 13 years of age who had the recommended immunizations (meningococcal vaccine and one tetanus, diphtheria toxoids and acellular pertussis vaccine (Tdap) or one tetanus, diphtheria toxoids vaccine (Td))) by their 13th birthday.	Low birth weight (<2500 gram) infants per 1,000 newborns (excluding transfers)	Surveyed enrollees' self-evaluation for overall health status.	Surveyed enrollees' self-evaluation for managing chronic conditions	Surveyed enrollees' self-evaluation of childhood educational delays.	MI Schools Dashboard school counts of IEP
NQF Number	0038	1407	0278	--	--	--	--
Measure Steward	National Committee for Quality Assurance	National Committee for Quality Assurance	Agency for Healthcare Research & Quality	AHRQ CAHPS/BRFSS Question Modification	--	--	State of Michigan Department of Education
Numerator	# children who received the recommended vaccines by their second birthday. Separate rates calculated for each	# adolescents 13 years of age who had one dose of meningococcal vaccine and one tetanus, diphtheria	# of newborns, among cases meeting inclusion/exclusion rules for the denominator, with	Number of respondents participating with at least 2 survey waves who have an increase in the level of self-	Number of respondents participating with at least 2 survey waves who report increase in confidence in managing chronic conditions.	Number of respondents participating in at least 2 survey waves who report childhood	Number of students who have official IEP for each age group.

	vaccine as well as 9 separate combination rates.	toxoids and acellular pertussis vaccine (Tdap) or one tetanus, diphtheria toxoids vaccine (Td) by their 13th birthday.	any-listed ICD-9-CM (ICD-10) diagnosis codes for birth weight less than 2,500 grams.	reported health status. *Sample questions:* "In general, how would you rate your overall health?" (excellent/very good/good/fair/poor) "In general, how would you rate your overall mental or emotional health?" (excellent/very good/good/fair/poor)	*Sample Tools:* Adult/Pediatric Asthma Control Test	cognitive and educational delays.	
Denominator	# children who turn 2 years of age during the measurement period.	# adolescents who turn 13 years of age during the measurement period.	# of newborns in region	Number of survey participants.	Number of survey participants.	Number of survey participants	Number of student counts for Flint City Schools
Baseline Value(s)	DY 1-5	DY 1-5	DY 1-5	DY 1-5	DY 1-5	DY 1-5	DY 1-5
Sampling Methodology	No sampling – plan to use 100% available claims/encounter data	No sampling – plan to use 100% available claims/encounter data	No sampling – plan to use 100% available claims/encounter data	Random/weighted sampling	Random/weighted sampling	Random/weighted sampling	No sampling - plan to use 100% available student counts for Flint City Schools
Anticipated Data Source	Administrative claims/encounters in the MDHHS data warehouse	Administrative claims/encounters in the MDHHS data warehouse	Administrative claims/encounters in the MDHHS data warehouse linked to Vital Records	Enrollee survey responses	Enrollee survey responses	Enrollee survey responses	MI Schools Dashboard