State of Alaska Department of Health, Division of Behavioral Health

Alaska Section 1115 Demonstration Waiver – Behavioral Health Reform

Renewal Evaluation Design

January 2025

HSAG HEALTH SERVICES ADVISORY GROUP

Table of Contents

1. Background

The Centers for Medicare & Medicaid Services (CMS) and federal law establish standards for the minimum care that states must provide to Medicaid-eligible populations, while also allowing states to design and test their own strategies for funding and providing healthcare services. Under Section 1115 of the Social Security Act, states can implement innovative demonstration projects and evaluate state-specific policy changes to improve efficiency and reduce costs. On March 26, 2024, CMS approved Alaska's request to extend its Section 1115 Demonstration Waiver, now titled Behavioral Health Reform.[1-1] This extension is approved for five years, from March 26, 2024, through December 31, 2028.

Alaska's Substance Use Disorder Landscape

In line with national trends, opioid use and overdose in Alaska became significantly more prevalent over the last decade. Since 2008, deaths involving opioids reached historical highs. Although there were small improvements at the turn of the last decade, the most recent available data showed that Alaskan opioid death counts continued to rise from 2013 to 2020.[1-2] By 2022, opioid-related overdose deaths nearly quadrupled from 2010, averaging 24.9 deaths per 100,000.[1-3] From 2018 to 2022, 633 of Alaska's 886 drug overdose deaths involved opioids, slightly over 70 percent.[1-4] Opioid misuse was not exclusive to the State of Alaska; rates of self-reported opioid misuse were similar or higher in Alaska compared to national trends, with 3.0 percent of Alaskans reporting misuse of any opioids and 22.8 percent of Alaskans reporting illicit drug use, compared to national rates of 3.3 percent and 15.5 percent, respectively, in 2022.[1-5] According to the 2021–2022 National Survey on Drug Use and Health (NSDUH), 21.1 percent of Alaskan adults reported binge alcohol use in the past month, compared to a national rate of 21.7 percent. Additionally, 23.0 percent of Alaskans had a substance use disorder (SUD), compared to a national rate of 17.0 percent,[1-6] and 6.7 percent of Alaskans reported needing but not receiving treatment for illicit drug use in the past year, compared with a national rate of 4.8 percent.[1-7] Notably, alcohol misuse was prominent in Alaska, which ranked 15th in the nation for highest prevalence rate of adult binge drinking in 2022.[1-8]

1-1 From January 1, 2019, to March 25, 2024, Behavioral Health Reform was called the Substance Use Disorder-Behavioral Health Program.

1-2 Centers for Disease Control and Prevention, National Center for Health Statistics. Multiple Cause of Death 1999-2020 on CDC WONDER Online Database released in 2021. Data are from the Multiple Cause of Death Files, 1999-2020, as compiled from data provided by the 57 vital statistics jurisdictions through the Vital Statistics Cooperative Program. 2020. Available at: http://wonder.cdc.gov/mcd-icd10.html. Accessed on: Aug 1, 2024.

1-3 Alaska Department of Health. 2022 Drug Overdose Mortality Update. Available at: https://health.alaska.gov/dph/VitalStats/Documents/PDFs/DrugOverdoseMortalityUpdate_2022.pdf. Accessed on: Aug 2, 2024.

1-4 Ibid.

1-5 Substance Abuse and Mental Health Services Administration. Behavioral Health Barometer, Alaska, Volume 6. Available at: https://www.samhsa.gov/data/sites/default/files/reports/rpt45304/2022-nsduh-barometer-region-10.pdf. Accessed on: Aug 2, 2024.

1-6 Ibid.

1-7 Kaiser Family Foundation. State Health Facts: Mental Health & Substance Use. Available at: https://www.kff.org/state-category/mental-health/alcohol-drug-dependence-and-abuse/. Accessed on: Aug 2, 2024.

1-8 Statista. Binge drinking prevalence among adults in the United States as of 2022, by state. Available at: https://www.statista.com/statistics/378966/us-binge-drinking-rate-adults-by-state/#:~:text=U.S.%20binge%20drinking%20among%20adults%20by%20state%202022&text=As%20of%202022%2C%20the%20U.S.,last%2030%20days%20in%202022. Accessed on: Aug 2, 2024.

The need for behavioral health (BH) services, which often coincided with the need for SUD treatment, was more prominent among Alaskans than the rest of the nation. Data from the 2022 Behavioral Risk Factor Surveillance System (BRFSS) showed that 15.2 percent of Alaskans and 17.2 percent of Alaska Natives reported frequent mental distress, defined as experiencing poor mental health (MH) 14 or more days per month.[1-9] In addition, Alaska's 2022 suicide rate of 26.7 per 100,000 Alaskans was more than twice the 2015 national rate of 12.32 per 100,000 Alaskans, and the Alaska Native population was over two times as likely to die by suicide compared to non-Alaska Natives.[1-10] With rates of mental illness, suicide, illicit and opioid drug use, overdose deaths, and binge drinking stable or increasing, and often in line with or surpassing national trends, Alaskans continue to need SUD and BH services, as well as intervention to address downstream effects that further perpetuate the need for these services.

For example, between 2017 and 2021, the number of Medicaid-covered infants diagnosed with neonatal abstinence syndrome (NAS) increased by over 35 percent, from 85.3 per 1,000 births to 116.4 per 1,000 births.[1-11] Additionally, children living with adults with SUD and other BH ailments often experienced adverse childhood experiences (ACEs), placing them at a significantly higher likelihood of risky behaviors such as substance misuse, alcoholism, smoking, and unsafe sex practices and subsequent sexually transmitted infections (STIs). Children with a high prevalence of ACEs are more likely to experience physical and mental morbidities including certain cancers, obesity, depression, or premature mortality including suicide, in adulthood.[1-12] In 2021, 16.96 percent of Alaskan children had experienced two or more adverse events,[1-13] compared to 14.0 percent nationally.[1-14] The higher rates of ACEs in Alaska not only coincided with higher rates of adult SUD and BH ailments, they also perpetuated a cycle of high rates of SUD and BH ailments as ACE-affected children aged into adulthood with an increased aptitude to partake in risky behaviors. As a result, there was a clear need for intervention across all age groups in Alaska.

Further exacerbating the challenges of providing SUD and BH interventions in Alaska is the State's unique infrastructure. While Alaska is the largest state by land mass, its population density is significantly lower than that

---

1-9 Centers for Disease Control and Prevention. BRFSS Prevalence & Trends Data. Available at: https://nccd.cdc.gov/BRFSSPrevalence/rdPage.aspx?rdReport=DPH_BRFSS.ExploreByLocation&rdProcessAction=&SaveFileGenerated=1&irbLocationType=States&islLocation=02&islState=&islCounty=&islClass=CLASS20&islTopic=TOPIC71&islYear=2022&hidLocationType=States&hidLocation=02&hidClass=CLASS20&hidTopic=TOPIC71&hidTopicName=Healthy+Days&hidYear=2022&irbShowFootnotes=Show&rdICL-iclIndicators=_PHYS14D%2c_MENT14D&iclIndicators_rdExpandedCollapsedHistory=&iclIndicators=_PHYS14D%2c_MENT14D&hidPreviouslySelectedIndicators=&DashboardColumnCount=2&rdShowElementHistory=divTopicUpdating%3dHide%2cislTopic%3dShow%2cdivYearUpdating%3dHide%2cislYear%3dShow%2c&rdScrollX=0&rdScrollY=200&rdRnd=27721. Accessed on: Aug 2, 2024.

1-10 Alaska Department of Health Division of Public Health. Alaska Vital Statistics 2022 Annual Report. Available at: https://health.alaska.gov/dph/VitalStats/Documents/PDFs/VitalStatistics_Annualreport_2022.pdf. Accessed on: Aug 2, 2024.

[11]Centers for Medicare & Medicaid Services. Number and rate of NAS per 1,000 births in newborns whose deliveries were covered by Medicaid or CHIP, 2017–2021. Available at: https://data.medicaid.gov/dataset/9c9ad0d1-c59b-4a25-9314-8e7e44e7f281?conditions[0][property]=state&conditions[0][value]=Alaska&conditions[0][operator]=%3D. Accessed on: Aug 2, 2024.

1-12 Felitti VJ, Anda RF, Nordenberg D, et al. Relationship of Childhood Abuse and Household Dysfunction to Many of the Leading Causes of Death in Adults. *Am. J Prev Med* 1998;14(4). Available at: https://www.ajpmonline.org/action/showPdf?pii=S0749-3797%2898%2900017-8. Accessed on: Aug 2, 2024.

1-13 An adverse event includes parent divorce or separation; living with someone who had an alcohol or drug problem; neighborhood violence victim or witness; living with someone who was mentally ill, suicidal, or severely depressed; domestic violence witness; parent served jail time; being treated or judged unfairly due to race/ethnicity; or death of a parent.

1-14 United Health Foundation. America's Health Rankings, Adverse Childhood Experiences, Alaska. 2021. Available at: https://www.americashealthrankings.org/explore/measures/ACEs_8. Accessed on: Aug 2, 2024.

of cities in the contiguous United States. For example, Alaska's largest city, Anchorage, had an estimated population of 286,075 in July 2023, much smaller than many cities in the lower 48 states, which have populations exceeding one million.[1-15] In addition, Alaskan communities are widely distanced and inaccessible by road, resulting in them being medically underserved. The large geographic size and small population of Alaska make SUD and BH support less accessible, and there are fewer healthcare professionals compared to communities in the contiguous United States. Moreover, Alaska's northern and unforgiving climate constantly poses a challenge for accessibility and delivery of healthcare services.

Lastly, Alaska has a diverse population that includes 229 federally recognized tribes, 20 different native languages, and a growing immigrant population throughout the State. Alaska is home to 37 tribal health organizations that serve the tribal population, many of which receive grants from the Division of Behavioral Health (DBH). This diversity presents a challenge for providing culturally and regionally appropriate care.

Waiver Background

On March 1, 2023, the Alaska Department of Health (DOH) submitted an application to CMS to renew its Medicaid Section 1115 Demonstration Waiver. The renewal application proposed continuing the SUD-BH Program under a new name, Behavioral Health Reform. The name change reflects the same MH and SUD services as under the previous waiver under the broader term of BH.[1-16] CMS approved Behavioral Health Reform on March 26, 2024, with a demonstration period running from March 26, 2024, to December 31, 2028. Behavioral Health Reform will continue to increase access to SUD and BH services for Alaskans, aiming to anticipate or eliminate crises and strengthen the continuum of care, including early intervention services and community support. The primary objectives of Behavioral Health Reform include:

- Rebalancing the current BH system of care to reduce Alaska's over-reliance on acute, institutional care and shift to more community or regionally based care.
- Intervening as early as possible in the lives of Alaskans to address BH symptoms before symptoms cascade into functional impairments.
- Improving the overall BH system accountability by reforming the existing system of care.

Additionally, the State identified long-term goals for the Behavioral Health Reform:[1-17]

- Increased rates of identification, initiation, and engagement in treatment for substance use and BH issues.
- Increased adherence to and retention in treatment for substance use and BH issues.
- Reduced overdose deaths, particularly those due to opioids.

1-15 United States Census Bureau. Quick Facts. Available at: https://www.census.gov/quickfacts/fact/table/anchoragemunicipalityalaska,US/PST045221. Accessed on: Aug 2, 2024.

1-16 In this evaluation, behavioral health (BH) refers to both substance use disorder (SUD) and mental health (MH) services, while 'SUD' and 'MH' refer specifically to their respective services.

1-17 Centers for Medicare & Medicaid Services. CMS Approval Behavioral Health Reform. Available at: https://www.medicaid.gov/medicaid/section-1115-demonstrations/downloads/ak-behavioral-health-refm-aprvl-03262024.pdf. Accessed on: Aug 2, 2024.

- Reduced utilization of emergency departments (EDs) and inpatient (IP) hospital settings for substance use and BH treatment where the utilization is preventable or medically inappropriate through improved access to other more appropriate and focused services.
- Fewer readmissions to the same or higher level of care where readmission is preventable or medically inappropriate.
- Improved access to care for physical health (PH) conditions among beneficiaries.

Program Population

Behavioral Health Reform impacts three Alaskan Medicaid beneficiary population groups. Behavioral Health Reform did not impact Medicaid eligibility standards.

Group 1: Children, adolescents, and their parents or caretakers with or at risk of MH disorders and SUD	Group 1 beneficiaries include those under the supervision or in the custody of the Alaska DOH Office of Children's Services, the Division of Juvenile Justice, or in tribal custody; formerly in kinship care, foster care, or residential care; or at risk of an out-of-home placement. Behavioral Health Reform services for this population include home-based family treatment, intensive case management (ICM), partial hospitalization program (PHP) services, intensive outpatient (IOP) services, children's residential treatment (CRT) level 1, and therapeutic treatment homes.
Group 2: Transitional age youth and adults with acute MH needs	Group 2 beneficiaries include transitional age youth and adults between the ages of 18 and 21 as well as between the ages 18 and 25 who experience MH disorders and have comorbidities or dual diagnoses, such as intellectual, developmental, or sensory disabilities, making their care needs more complex. Two different transitional age groups will be analyzed due to variations in the definition of transitional ages in the literature. Individuals in this population are at higher risk of emerging MH issues and SUD.[1-18] Behavioral Health Reform services for Group 2 include assertive community treatment (ACT) services, ICM, PHP services, adult MH residential services, and peer-based crisis services.
Group 3: Adults, adolescents, and children with SUD	Group 3 consists of adults, adolescents, and children between the ages of 12 and 64 years who have at least one diagnosis for substance-related and addictive disorders from the Diagnostic and Statistical Manual of Mental Disorders, 5th Edition (DSM-5), or the most current version.[1-19]

Services for this group aim to enhance the availability of care and provide a more comprehensive continuum of treatment for SUD, including:

- Opioid treatment services
- IOP services
- PHP services
- Residential treatment
- Medically monitored intensive IP services

[1-18] Martel A, Fuchs C. Transitional age youth and mental illness – influences on young adult outcomes. *Child and Adolescent Psychiatric Clinics*. 2017; 26(2): XIII–XVII.

[1-19] American Psychiatric Association. Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition, Text Revision (DSM-5-TR®). 2022.

- Medically managed intensive IP services
- Ambulatory withdrawal management
- Clinically managed residential withdrawal management
- Medically monitored IP withdrawal management
- Medically managed intensive IP withdrawal management

Select Behavioral Health Reform services replaced existing State plan services, with additional services introduced. Behavioral Health Reform services adhere to the American Society of Addiction Medicine (ASAM) level-of-care criteria to place patients in the right setting at the right time.[1-20]

Previous Report Findings

The March 2023 Interim Evaluation Report, approved by CMS on September 27, 2023,[1-21] provides evidence of Alaska's progress toward its goals. The report suggested that SUD beneficiaries were transitioning away from ED utilization to outpatient (OP) care with the SUD-BH Program implementation. Additionally, SUD and BH provider capacity expanded, and new services became more widely available. Timely initiation of SUD treatment improved, and the average length of stay (LOS) in institutions for mental diseases (IMDs) significantly decreased. However, the coronavirus disease 2019 (COVID-19) public health emergency (PHE) likely impacted utilization, and workforce shortages impacted expansion of services in the State.

The independent evaluator will incorporate a synthesis of results from the prior demonstration period Summative Evaluation Report in the Behavioral Health Reform's Interim Evaluation Report, due to CMS by December 31, 2027.

Additional research questions and measures have been added to this Evaluation Design since the approval of the prior demonstration period's Interim Evaluation Report in September 2023. Table 1-1 lists the research questions that are new to Behavioral Health Reform.

Table 1-1—New Research Questions for Behavioral Health Reform

New Research Questions
1.2: Does Behavioral Health Reform mitigate barriers to maintaining and providing SUD and BH resources?
2.2: Does Behavioral Health Reform reduce the number of beneficiaries who experience or are exposed to adverse events?

1-20 Alaska Department of Health and Social Services. Medicaid Section 1115 Behavioral Health Demonstration Application. Available at: https://www.medicaid.gov/Medicaid-CHIP-Program-Information/By-Topics/Waivers/1115/downloads/ak/behavioral-health/ak-behavioral-health-demo-pa.pdf. Accessed on: Aug 2, 2024.

1-21 Centers for Medicare & Medicare Services. SUD Interim Evaluation Report. Available at: https://www.medicaid.gov/medicaid/section-1115-demonstrations/downloads/ak-behavioral-health-sud-interim-evaluton-rpt-09272023.pdf. Accessed on: Aug 2, 2024.

2. Evaluation Questions and Hypotheses

This section provides the logic model, hypotheses, research questions, and measures which focus on evaluating the impact of the Behavioral Health Reform Section 1115 Demonstration Waiver.

Logic Model

Figure 2-1 displays the logic model for Behavioral Health Reform. The model depicts the relationship between the demonstration's goal/purpose/aim, the primary drivers that contribute to realizing that purpose, and the secondary drivers that are necessary to achieve the primary drivers. Both the secondary and primary drivers will influence each Centers for Medicare & Medicaid Services (CMS Goal), though the relationships between the secondary and primary drivers may be multidirectional.[2-1]

Figure 2-1—Driver Diagram

Primary Drivers[1]	Secondary Drivers[2]
Behavioral Health Reform Goal 1: Increased rates of identification, initiation, and engagement in treatment for substance use and BH issues (Hypotheses 1 and 2)	
Access to Critical Levels of Care for SUD/MH Treatment[2-2]	• Universally screen all beneficiaries, regardless of setting, using industry-recognized, evidence-based SUD and MH screening instruments to identify symptoms for preventive measures and intervene as early as possible • Improve care coordination through the integration of BH into primary care and specialty medical settings • Increase SUD and BH treatment options for youth (ages 12–17) and adult (ages 18 and older) beneficiaries, particularly non-residential, step-up, and step-down treatment options • Increase the number of Medicaid qualified professionals who provide SUD or BH services
Use of Evidence-Based, SUD and MH Patient Placement Criteria	• Implement ASAM Criteria to match beneficiaries with SUD with the services and tools necessary for recovery • Improve beneficiary knowledge of available treatment services

2-1 Centers for Medicare & Medicaid Services. Defining and Using Aims and Drivers for Improvement. January 1, 2013. Available at: https://www.cms.gov/priorities/innovation/files/x/hciatwoaimsdrvrs.pdf. Accessed on Aug 24, 2024.

2-2 In this evaluation, behavioral health (BH) refers to both substance use disorder (SUD) and mental health (MH) services, while 'SUD' and 'MH' refer specifically to their respective services.

Primary Drivers[1]	Secondary Drivers[2]
Behavioral Health Reform Goal 2: Increased adherence to and retention in treatment for substance use and BH issues (Hypotheses 1 and 2)	
Use of Evidence-Based, SUD and MH Patient Placement Criteria	• Improve care coordination through the integration of BH into primary care and specialty medical settings • Increase SUD and MH treatment options for youth (ages 12–17) and adult (ages 18 and older) beneficiaries, particularly non-residential, step-up, and step-down treatment options • Increase access to OP, IOP, and residential treatment for SUD • Ensure patients are satisfied with services • Increase the number of Medicaid qualified professionals who provide SUD or MH services
Use of Evidence-Based, SUD-Specific Patient Placement Criteria	• Implement ASAM Criteria to match beneficiaries with SUD with the services and tools necessary for recovery
Behavioral Health Reform Goal 3: Reduced overdose deaths, particularly those due to opioids (Hypotheses 1 and 2)	
Implementation of Comprehensive Treatment and Prevention Strategies to Address Opioids	• Universally screen all beneficiaries, regardless of setting, using industry-recognized, evidence-based SUD screening instruments to identify symptoms for preventive measures and intervene as early as possible before use becomes dependence • Improve SUD provider infrastructures and capacity utilizing industry-recognized standards for certification and ongoing accountability (with emphasis on residential providers, but across-the-board) • Implement ASAM Criteria to match beneficiaries with SUD with the services and tools necessary for recovery
Improve adherence to treatment for OUD and other SUDs	• Improve care coordination and transitions between levels of care, ensuring residential and IP facilities link beneficiaries with community-based services and supports • Increase SUD and MH treatment options for youth (ages 12–17) and adult (ages 18 and older) beneficiaries, particularly non-residential, step-up, and step-down treatment options

Primary Drivers[1]	Secondary Drivers[2]
Behavioral Health Reform Goal 4: Reduced utilization of EDs and IP hospital settings for substance use and MH treatment where the utilization is preventable or medically inappropriate through improved access to other more appropriate and focused services (Hypotheses 1 and 2)	
Increase Access to Continuum of Care	• Improve SUD provider infrastructures and capacity utilizing industry-recognized standards for certification and ongoing accountability (with emphasis on residential providers, but across-the-board) • Improve care coordination through the integration of BH into primary care and specialty medical settings • Promote appropriate service utilization through improved access to care for PH conditions for beneficiaries
Use of Evidence-Based, SUD-Specific Patient Placement Criteria	• Implement ASAM Criteria to match beneficiaries with SUD with the services and tools necessary for recovery • Increase SUD and BH treatment options for youth (ages 12–17) and adult (ages 18 and older) beneficiaries, particularly non-residential, step-up, and step-down treatment options
Behavioral Health Reform Goal 5: Fewer readmissions to the same or higher level of care where readmission is preventable or medically inappropriate (Hypotheses 1 and 2)	
Increase Access to Continuum of Care	• Improve SUD provider infrastructures and capacity utilizing industry-recognized standards for certification and ongoing accountability (with emphasis on residential providers, but across-the-board) • Improve care coordination and transitions between levels of care, ensuring residential and IP facilities link beneficiaries with community-based services and supports • Promote appropriate service utilization through improved access to care for PH conditions for beneficiaries • Ensure patients are satisfied with services
Use of Evidence-Based, SUD-Specific Patient Placement Criteria	• Implement ASAM Criteria to match beneficiaries with SUD with the services and tools necessary for recovery • Increase SUD and MH treatment options for youth (ages 12–17) and adult (ages 18 and older) beneficiaries, particularly non-residential, step-up, and step-down treatment options

Primary Drivers[1]	Secondary Drivers[2]
Behavioral Health Reform Goal 6: Improved access to care for PH conditions among beneficiaries (Hypothesis 2)	
Integrate PH and BH care	• Improve care coordination through the integration of BH into primary care and specialty medical settings • Increase access to preventive health services for beneficiaries overall, and those with comorbid conditions

[1] Primary drivers are major domains through which Alaska may accomplish the six goals adapted from CMS' special terms and conditions (STCs).
[2] Secondary drivers are from Alaska's implementation plan, utilizing key milestones identified by CMS.
Note: ASAM: American Society of Addiction Medicine; BH: behavioral health; IOP: intensive outpatient; IP: inpatient; MH: mental health; OP: outpatient; PH: physical health; SUD: substance use disorder.

Hypotheses and Research Questions

To comprehensively evaluate Behavioral Health Reform, three hypotheses, listed in Table 2-1, will be assessed using eight research questions.

Table 2-1—Hypotheses

	Hypotheses
1	Behavioral Health Reform will maintain or increase the identification of and access to SUD and MH disorder treatment services by increasing access to community-based care.
2	Beneficiaries receiving BH services will experience maintained or improved health and well-being outcomes.
3	Behavioral Health Reform will maintain or reduce the cost of Medicaid for Alaska and the federal government.

Hypothesis 1 is designed to identify whether Behavioral Health Reform increased community-based care and if the identification of substance use disorder (SUD) and mental health (MH) treatment disorder services increased in response. The measures and research questions associated with Hypothesis 1 are presented in Table 2-2.

Table 2-2—Hypothesis 1 Research Questions and Measures

Hypothesis 1: Behavioral Health Reform will maintain or increase the identification of and access to SUD and MH disorder treatment services by increasing access to community-based care.	
Research Question 1.1: Does Behavioral Health Reform increase the number of beneficiaries in the waiver population who are screened, referred to, and engaged in treatment for SUD and MH disorders?	
1-1	Number of beneficiaries screened for symptoms of SUD using industry-recognized, evidence-based screening instruments
1-2	Number of beneficiaries screened for symptoms of MH disorders using industry recognized, evidence-based screening instruments
1-3	Number of beneficiaries in the waiver population with a SUD or MH diagnosis, by setting

Hypothesis 1: Behavioral Health Reform will maintain or increase the identification of and access to SUD and MH disorder treatment services by increasing access to community-based care.	
1-4	Number of child beneficiaries who are taking medication for ADD/ADHD, ASD, difficulties with emotions, concentration, or behavior
1-5	Percentage of beneficiaries who had initiation of SUD treatment
1-6	Percentage of beneficiaries who had engagement of SUD treatment
1-7	Follow-up after discharge within 7-days and 30-days from ED visits for SUD, and specifically for OUD
1-8	Follow-up after discharge within 7-days and 30-days from ED visits for a MH disorder
1-9	Number of Medicaid qualified SUD providers (identified by provider ID numbers) who bill for SUD services
1-10	Number of Medicaid qualified professionals licensed in the State to provide MH who bill for MH disorder services
1-11	Number of respondents who received substance use treatment
1-12	Number of respondents who were classified as needing substance use treatment
1-13	Number of respondents who received MH treatment
Research Question 1.2: Does Behavioral Health Reform mitigate barriers to maintaining and providing SUD and MH resources?	
1-14	Providers' experience providing SUD and MH services
1-15	Providers' reported barriers maintaining the integration of SUD and MH services
1-16	Provider's reported successes maintaining the integration of SUD and MH services
1-17	Providers' reported changes in the ability to provide care after DBH procurement of a BH organization
1-18	Administrators' reported barriers to maintaining SUD and MH services
1-19	Administrators' reported successes in maintaining SUD and MH services
1-20	Administrators' plan for program sustainability and anticipated challenges
1-21	Alaska tribal entities' reported continued changes in quality of care and access to care in providing SUD and MH services
Research Question 1.3: Does Behavioral Health Reform decrease utilization of ED, IP, or institutional settings within the beneficiary population?	
1-22	IP admissions for SUD, and specifically for OUD
1-23	IP admissions for MH disorders
1-24	ED visits for SUD
1-25	ED visits for OUD
1-26	Number of ED visits among beneficiaries who had at least one ED admission for SUD
1-27	Number of ED visits for SUD among high utilizing beneficiaries
1-28	ED visits for MH disorders
1-29	Number of ED visits among beneficiaries who had at least one ED admission for MH
1-30	Number of ED visits for MH among high utilizing beneficiaries
1-31	Mean length of stay among IMDs measured from admission date to discharge date
1-32	30-day readmission rate to IP facilities following hospitalization for an SUD-related diagnosis
1-33	30-day readmission rate to IP facilities following hospitalization for a MH-related diagnosis

Hypothesis 1: Behavioral Health Reform will maintain or increase the identification of and access to SUD and MH disorder treatment services by increasing access to community-based care.	
Research Question 1.4: Does Behavioral Health Reform increase the percentage of beneficiaries who adhere to treatment for SUD and MH disorders?	
1-34	Number of beneficiaries with a SUD diagnosis including those with OUD who used services in the last month or year, by service or benefit type
1-35	Number of beneficiaries with a MH diagnosis who used services in the last month or year, by service or benefit type
1-36	Time to treatment among beneficiaries who access SUD treatment

Note: ADD/ADHD: attention deficit disorder/attention deficit hyperactivity disorder; ASD: autism spectrum disorder; BH: behavioral health; DBH: Department of Behavioral Health; ED: emergency department; IMD: Institutions for Mental Disease; IP: inpatient; MH: mental health; OUD: opioid use disorder; SUD: substance use disorder

Hypothesis 2 will evaluate whether Behavioral Health Reform improved health outcomes for beneficiaries. The measures and research questions associated with Hypothesis 2 are presented in Table 2-3.

Table 2-3—Hypothesis 2 Research Questions and Measures

Hypothesis 2: Beneficiaries receiving BH services will experience maintained or improve health and well-being outcomes.	
Research Question 2.1: Does Behavioral Health Reform increase the percentage of beneficiaries with an SUD or a MH disorder who experience care for comorbid conditions?	
2-1	Percentage of adults who accessed preventive/ambulatory health services
2-2	Percentage of beneficiaries 3–21 years of age with a well-care visit with a PCP or OB/GYN
2-3	Screening for chronic conditions relevant to State Medicaid population
2-4	Screening for co-morbidity of BH disorders and SUDs
2-5	Prevention Quality Chronic Composite
2-6	Pediatric Quality Chronic Composite
2-7	Percentage of beneficiaries who have a high rating of their healthcare quality (8, 9, or 10 on a scale of 0–10)
2-8	Percentage of beneficiaries who rate their overall mental or emotional health as “very good” or “excellent”
2-9	Percentage of beneficiaries who demonstrate “very good” or “excellent” knowledge of available SUD/MH treatment and services
2-10	Percentage of beneficiaries who are knowledgeable of the number of SUD and MH services available
2-11	Percentage of mothers who often or always felt depressed since their new baby was born
2-12	Percentage of beneficiaries who indicated poor MH in the last 30 days
2-13	Percentage of mothers who indicate that they have someone who would help them while sick
2-14	Desire to obtain SUD/MH treatment and obtainment of SUD treatment in the past three months
Research Question 2.2: Does Behavioral Health Reform reduce the number of beneficiaries who experience or are exposed to adverse events?	
2-15	Percentage of mothers who reported that during the past 12 months, their husband or partner pushed, hit, slapped, kicked, choked, or physically hurt them in any other way
2-16	Percentage of mothers who reported that in the past 12 months, their husband or partner threatened them, limited their activities against their will, or made them feel unsafe in any other way
2-17	Percentage of respondents whose child lived with someone who had a problem with alcohol or drugs
2-18	Percentage of respondents whose child lived with someone who was mentally ill, suicidal, or severely depressed
2-19	Percentage of respondents whose child witnessed violence or physical abuse between household beneficiaries

Hypothesis 2: Beneficiaries receiving BH services will experience maintained or improve health and well-being outcomes.	
2-20	Percentage of respondents who reported that their child saw or heard parents or adults slap, hit, kick, or punch one another in the home
2-21	Maltreatment types among victims
2-22	Caregiver risk factors among child victims
2-23	Maternal use of marijuana or cannabis in any form
2-24	Frequency of maternal marijuana or cannabis use in the past 30 days
Research Question 2.3: Does Behavioral Health Reform decrease the rate of drug overdoses and overdose deaths due to opioids?	
2-25	Rate of overdose deaths, specifically overdose deaths due to any opioid
2-26	Non-fatal overdoses (all cause)
2-27	Use of opioids at high dosage in persons without cancer
2-28	Use of opioids from multiple providers
2-29	Risk of continued opioid use

Note: BH: behavioral health; OB/GYN: obstetrician/gynecologist; PCP: primary care physician; MH: mental health; SUD: substance use disorder

Hypothesis 3 seeks to measure the cost-effectiveness of Behavioral Health Reform. The measures and research questions associated with Hypothesis 3 are presented in Table 2-4.

Table 2-4—Hypothesis 3 Research Questions and Measures

Hypothesis 3: Behavioral Health Reform will maintain or reduce the cost of Medicaid for Alaska and the federal government.	
Research Question 3.1: Does Behavioral Health Reform maintain or reduce Alaska's per capita Medicaid BH costs?	
3-1	Total costs of healthcare (sum of parts below), by State and federal share
3-2	Total cost of SUD, SUD-IMD, and SUD-Other, and Non-SUD, by setting, including claims data (IP, OP, Rx, LTC)
3-3	Total cost of MH diagnosis by IMD and Other, by setting, including claims data (IP, OP, Rx, LTC)

Note: IMD: Institutions for Mental Diseases; IP: inpatient; LTC: long term care; MH: mental health; OP: outpatient; Rx: prescription medicine; SUD: substance use disorder

3. Methodology

To assess the impact of Alaska's Section 1115 Demonstration Waiver, Behavioral Health Reform, a comparison of outcomes between the intervention group and a valid counterfactual—the intervention group had they not been exposed to the intervention—must be made. The gold standard for experimental design is a randomized controlled trial which would be implemented by first identifying an intervention population, and then randomly assigning individuals to the intervention and the rest to a comparison group, which would serve as the counterfactual.[3-1] The use of random assignment is limited in healthcare evaluations because it is often infeasible and unethical to provide treatment to an intervention group and deny treatment to those assigned to the comparison group.

As such, a variety of quasi-experimental or observational methodologies have been developed for evaluating the effect of policies on outcomes. The research questions presented in the previous section will be addressed using at least one of these methodologies. The selected methodology depends on data availability factors relating to: (1) data to measure the outcomes, (2) data for a valid comparison group, and (3) data during the time periods of interest—typically defined as the year prior to implementation and annually thereafter. Table 3-1 illustrates a sampling of standard analytic approaches and whether the approach requires data gathered at the baseline (i.e., pre-implementation); requires a comparison group; or allows for causal inference to be drawn. It also notes key requirements unique to a particular approach.

Table 3-1—Sampling of Analytic Approaches

Analytic Approach	Baseline Data	Comparison Group	Allows Causal Inference	Notes
Difference-in-Differences	✓	✓	✓	Trends in outcomes should be similar between comparison and intervention groups at baseline.
Interrupted Time Series	✓		✓	Requires sufficient data points prior to and following implementation.
Pre-Test/Post-Test	✓			

Evaluation Design Summary

Summary of Approach

The SUD-BH Program targeted three populations, beginning in January 2019, and did not undergo substantive changes upon renewal, known as Behavioral Health Reform, in March 2024.[3-2, 3-3] A comprehensive evaluation of

3-1 Contreary K, Bradley K, Chao S. Best Practices in Causal Inference for Evaluations of Section 1115 Eligibility and Coverage Demonstrations. *Mathematica Policy Research*. June 2018. Available at: Best Practices in Causal Inference for Evaluations of Section 1115 Eligibility and Coverage Demonstrations (hhs.gov). Accessed on: Aug 7, 2024.

3-2 Centers for Medicare & Medicaid Services. CMS Approval Substance Use Disorder Treatment and Alaska Behavioral Health Program. Available at: https://www.medicaid.gov/Medicaid-CHIP-Program-Information/By-Topics/Waivers/1115/downloads/ak/behavioral-health/ak-behavioral-health-demo-no-imp-plan-20190321.pdf. Accessed on: Aug 21, 2024.

3-3 Centers for Medicare & Medicaid Services. CMS Approval Behavioral Health Reform. Available at: https://www.medicaid.gov/medicaid/section-1115-demonstrations/downloads/ak-behavioral-health-refm-aprvl-03262024.pdf. Accessed on: Aug 21, 2024.

the SUD-BH Program and its impacts among substance use disorder (SUD) and behavioral health (BH) care and outcomes was conducted in the Interim Evaluation Report and will be conducted in the forthcoming Summative Evaluation Report for the January 2019 through March 2024 time period. As such, the evaluation of the renewal will primarily seek to determine whether Behavioral Health Reform, previously known as the SUD-BH Program, goals continue to show promising results during the renewal period. The independent evaluator has made adjustments to further refine the intervention populations and to provide national comparisons when data are available. Quantitative methods will include a pre-test and post-test approach, showing change over time in both counts and rates for specific metrics, or interrupted time series (ITS) analysis to assess whether the Behavioral Health Reform affected changes across specific outcome measures. Beneficiary survey results will be analyzed using a pre-test/post-test approach to assess changes in beneficiaries' rating of their personal doctor and overall healthcare as well as knowledge of available services before and after the renewal of the demonstration. The independent evaluator will use national surveys to construct a difference-in-differences (DiD) or synthetic controls analysis between Alaska and national respondents to compare changes in BH and exposure to adverse events. The independent evaluator may use additional data models to properly identify and control for heterogeneity in outcomes. Non-inferiority testing will determine if Behavioral Health Reform is performing as well or better than the counterfactual group for measures where the hypothesis is framed as maintaining, improving, or reducing. Where data is available, a health equity analyses will compare the relative performance of subgroups to a reference group. The results of these analyses may better inform the ability of members to access care and care outcomes across demographic groups. Finally, the independent evaluator will complete a qualitative component of Behavioral Health Reform to obtain the perspectives of providers, tribal health organizations' (THOs), and State administrators' perspectives regarding the renewal of the Behavioral Health Reform demonstration.

Intervention and Comparison Populations

Behavioral Health Reform targets three groups of Medicaid recipients:

- **Group 1**: Children, adolescents, and their parents or caretakers with or at risk of mental health (MH)[3-4] disorders and SUDs
- **Group 2**: Transitional age youth and adults with acute MH needs
- **Group 3**: Adults, adolescents, and children with SUDs

Since, Behavioral Health Reform covers a majority of Medicaid beneficiaries with SUD or MH diagnoses, the viability of an in-state comparison group is limited. For example, beneficiaries enrolled in Behavioral Health Reform may be fundamentally different from those in the general Medicaid population. The theoretical in-state comparison group may consist of beneficiaries who do not have an SUD or MH diagnosis, which would make these beneficiaries have vastly different health needs and outcomes than the intervention population. It is possible that these groups could serve as a comparison group with a risk-adjustment algorithm applied; however, this approach is unlikely to sufficiently adjust for the substantial differences across subpopulations to produce accurate and reliable results. Since Alaska does not have an all-payer claims database, it is not possible to identify and use an in-state low-income non-Medicaid population as a comparison group.

[3-4] In this evaluation, BH refers to both SUD and MH services, while 'SUD' and 'MH' refer specifically to their respective services.

To evaluate of survey measures in hypotheses 1 and 2, the independent evaluator will compare survey data from Alaskan residents with those nationally. This comparison group will help identify potential program impacts related to SUD, MH, or exposure to adverse events. However, one caveat to the potential survey data sources is that only Behavioral Risk Factor Surveillance System (BRFSS) data includes a mechanism to identify Medicaid recipients. As such, the intervention populations of these surveys will reflect Alaska as a whole, rather than Behavioral Health Reform alone.

Evaluation Periods

Table 3-2 presents the baseline and evaluation periods for Behavioral Health Reform. The Mid-Point Assessment (MPA) is distinct from the Interim and Summative Evaluation Reports. For the MPA, the independent evaluator will work with the Department of Behavioral Health (DBH) to present monitoring metrics data and feedback from key stakeholders, rather than completing a comprehensive evaluation, as described in this Evaluation Design.[3-5] Although Behavioral Health Reform is a renewal of the SUD-BH Program, the baseline period will only cover two years prior to the implementation of the renewal demonstration to allow for an isolated evaluation of the renewal's impact on SUD/MH outcomes in the State. The baseline and evaluation periods will be utilized for all measures where data is available. When appropriate, rates calculated for the SUD-BH Program evaluation reports will be leveraged to provide additional context.

Table 3-2—Evaluation Periods

Deliverable	Baseline Period	Evaluation Period
Mid-Point Assessment	March 26, 2022–March 25, 2024	March 26, 2024–April 30, 2026
Interim Evaluation Report	March 26, 2022–March 25, 2024	March 26, 2024–December 31, 2026
Summative Evaluation Report	March 26, 2022–March 25, 2024	March 26, 2024–December 31, 2028

Evaluation Measures

Table 3-3 presents the evaluation measures along with the respective comparison groups, data sources, and analytic approaches. Please see Appendix A for full measure specifications.

[3-5] Centers for Medicare and Medicaid Services. *Mid-Point Assessment Technical Assistance*. Available at: https://www.medicaid.gov/medicaid/section-1115-demo/downloads/evaluation-reports/1115-sud-smised-mid-point-assessment-ta.pdf. Accessed on: Aug 20, 2024.

Table 3-3—Evaluation Measures

Measure(s)	Comparison Group(s)	Data Source(s)	Analytic Approach	Frequency	Stratification
Hypothesis 1: Behavioral Health Reform will maintain or increase the identification of and access to SUD and MH disorder treatment services by increasing access to community-based care.					
Research Question 1.1: Does Behavioral Health Reform increase the number of beneficiaries in the waiver population who are screened, referred to, and engaged in treatment for SUD and MH disorders?					
1-1: Number of beneficiaries screened for symptoms of SUD using industry-recognized, evidence-based screening instruments	• N/A	• State eligibility and enrollment data • Claims/encounter data	• Pre-test/post-test	• Annually	• Gender • Age • Race/ethnicity • Geography • Criminal justice involvement • High-cost high needs
1-2: Number of beneficiaries screened for symptoms of MH disorders using industry-recognized, evidence-based screening instruments	• N/A	• State eligibility and enrollment data • Claims/encounter data	• Pre-test/post-test	• Annually	• Gender • Age • Race/ethnicity • Geography • Criminal justice involvement • High-cost high needs
1-3: Number of beneficiaries in the waiver population with a SUD or MH diagnosis, by setting	• N/A	• State eligibility and enrollment data • Claims/encounter data	• Pre-test/post-test	• Annually	• Gender • Age • Race/ethnicity • Geography • Criminal justice involvement • High-cost high needs
1-4: Number of child beneficiaries who are taking medication for ADD/ADHD, ASD, difficulties with emotions, concentration, or behavior	• National benchmarks	• National Survey of Children's Health	• Pre-test/post-test • DiD	• Annually	• N/A

Measure(s)	Comparison Group(s)	Data Source(s)	Analytic Approach	Frequency	Stratification
1-5: Percentage of beneficiaries who had initiation of SUD treatment	• N/A	• State eligibility and enrollment data • Claims/encounter data	• Pre-test/post-test • ITS	• Annually/ Monthly	• Gender • Age • Race/ethnicity • Geography • Criminal justice involvement • High-cost high needs
1-6: Percentage of beneficiaries who had engagement of SUD treatment	• N/A	• State eligibility and enrollment data • Claims/encounter data	• Pre-test/post-test • ITS	• Annually/ Monthly	• Gender • Age • Race/ethnicity • Geography • Criminal justice involvement • High-cost high needs
1-7: Follow-up after discharge within 7-days and 30-days from ED visits for SUD, and specifically for OUD	• N/A	• State eligibility and enrollment data • Claims/encounter data	• Pre-test/post-test • ITS	• Annually/ Monthly	• Gender • Age • Race/ethnicity • Geography • Criminal justice involvement • High-cost high needs
1-8: Follow-up after discharge within 7-days and 30-days from ED visits for a MH disorder	• N/A	• State eligibility and enrollment data • Claims/encounter data	• Pre-test/post-test • ITS	• Annually/ Monthly	• Gender • Age • Race/ethnicity • Geography • Criminal justice involvement • High-cost high needs

Measure(s)	Comparison Group(s)	Data Source(s)	Analytic Approach	Frequency	Stratification
1-9: Number of Medicaid qualified SUD providers (identified by provider ID numbers) who bill for SUD services	• N/A	• Provider data • Claims/encounter data	• Pre-test/post-test	• Annually	• N/A
1-10: Number of Medicaid qualified professionals licensed in the State to provide MH who bill for MH disorder services	• N/A	• Provider data • Claims/encounter data	• Pre-test/post-test	• Annually	• N/A
1-11: Number of respondents who received substance use treatment	• Synthetic control group • State/National rates	• NSDUH	• Pre-test/post-test • DiD • Synthetic controls	• Annually	• N/A
1-12: Number of respondents who were classified as needing substance use treatment	• Synthetic control group • State/National rates	• NSDUH	• Pre-test/post-test • DiD • Synthetic controls	• Annually	• N/A
1-13: Number of respondents who received MH treatment	• Synthetic control group • State/National rates	• NSDUH	• Pre-test/post-test • DiD • Synthetic controls	• Annually	• N/A
Research Question 1.2: Does Behavioral Health Reform mitigate barriers to maintaining and providing SUD and MH resources?					
1-14: Providers' experience providing SUD and MH services	• N/A	• Key informant interviews	• Qualitative synthesis	• Two rounds (Interim and Summative)[1]	• N/A
1-15: Providers' reported barriers maintaining the integration of SUD and MH services	• N/A	• Key informant interviews	• Qualitative synthesis	• Two rounds (Interim and Summative)	• N/A
1-16: Provider's reported successes maintaining the integration of SUD and MH services	• N/A	• Key informant interviews	• Qualitative synthesis	• Two rounds (Interim and Summative)	• N/A
1-17: Providers' reported changes in the ability to provide care after DBH procurement of a BH organization	• N/A	• Key informant interviews	• Qualitative synthesis	• Two rounds (Interim and Summative)	• N/A

Measure(s)	Comparison Group(s)	Data Source(s)	Analytic Approach	Frequency	Stratification
1-18: Administrators' reported barriers to maintaining SUD and MH services	• N/A	• Key informant interviews	• Qualitative synthesis	• Two rounds (Interim and Summative)	• N/A
1-19: Administrators' reported successes in maintaining SUD and MH services	• N/A	• Key informant interviews	• Qualitative synthesis	• Two rounds (Interim and Summative)	• N/A
1-20: Administrators' plan for program sustainability and anticipated challenges	• N/A	• Key informant interviews	• Qualitative synthesis	• Two rounds (Interim and Summative)	• N/A
1-21: Alaska tribal entities' reported continued changes in quality of care and access to care in providing SUD and MH services	• N/A	• Key informant interviews	• Qualitative synthesis	• Two rounds (Interim and Summative)	• N/A
Research Question 1.3: Does Behavioral Health Reform decrease utilization of ED, IP, or institutional settings within the beneficiary population?					
1-22: IP admissions for SUD, and specifically for OUD	• N/A	• State eligibility and enrollment data • Claims/encounter data	• Pre-test/post-test • ITS	• Annually/ Monthly	• Gender • Age • Race/ethnicity • Geography • Criminal justice involvement • High-cost high needs
1-23: IP admissions for MH disorders	• N/A	• State eligibility and enrollment data • Claims/encounter data	• Pre-test/post-test • ITS	• Annually/ Monthly	• Gender • Age • Race/ethnicity • Geography • Criminal justice involvement • High-cost high needs

Measure(s)	Comparison Group(s)	Data Source(s)	Analytic Approach	Frequency	Stratification
1-24: ED visits for SUD	• N/A	• State eligibility and enrollment data • Claims/encounter data	• Pre-test/post-test • ITS	• Annually/ Monthly	• Gender • Age • Race/ethnicity • Geography • Criminal justice involvement • High-cost high needs
1-25: ED visits for OUD	• N/A	• State eligibility and enrollment data • Claims/encounter data	• Pre-test/post-test • ITS	• Annually/ Monthly	• Gender • Age • Race/ethnicity • Geography • Criminal justice involvement • High-cost high needs
1-26: Number of ED visits among beneficiaries who had at least one ED admission for SUD	• N/A	• State eligibility and enrollment data • Claims/encounter data	• Pre-test/post-test • ITS	• Annually/ Monthly	• Gender • Age • Race/ethnicity • Geography • Criminal justice involvement • High-cost high needs
1-27: Number of ED visits for SUD among high utilizing beneficiaries	• N/A	• State eligibility and enrollment data • Claims/encounter data	• Pre-test/post-test • ITS	• Annually/ Monthly	• Gender • Age • Race/ethnicity • Geography • Criminal justice involvement • High-cost high needs

Measure(s)	Comparison Group(s)	Data Source(s)	Analytic Approach	Frequency	Stratification
1-28: ED visits for MH disorders	• N/A	• State eligibility and enrollment data • Claims/encounter data	• Pre-test/post-test • ITS	• Annually/ Monthly	• Gender • Age • Race/ethnicity • Geography • Criminal justice involvement • High-cost high needs
1-29: Number of ED visits among beneficiaries who had at least one ED admission for MH	• N/A	• State eligibility and enrollment data • Claims/encounter data	• Pre-test/post-test • ITS	• Annually/ Monthly	• Gender • Age • Race/ethnicity • Geography • Criminal justice involvement • High-cost high needs
1-30: Number of ED visits for MH among high utilizing beneficiaries	• N/A	• State eligibility and enrollment data • Claims/encounter data	• Pre-test/post-test • ITS	• Annually/ Monthly	• Gender • Age • Race/ethnicity • Geography • Criminal justice involvement • High-cost high needs
1-31: Mean length of stay among IMDs measured from admission date to discharge date	• N/A	• State eligibility and enrollment data • Claims/encounter data	• Pre-test/post-test	• Annually/ Monthly	• Gender • Age • Race/ethnicity • Geography • Criminal justice involvement • High-cost high needs

Measure(s)	Comparison Group(s)	Data Source(s)	Analytic Approach	Frequency	Stratification
1-32: 30-day readmission rate to IP facilities following hospitalization for an SUD-related diagnosis	• N/A	• State eligibility and enrollment data • Claims/encounter data	• Pre-test/post-test	• Annually/ Monthly	• Gender • Age • Race/ethnicity • Geography • Criminal justice involvement • High-cost high needs
1-33: 30-day readmission rate to IP facilities following hospitalization for a MH-related diagnosis	• N/A	• State eligibility and enrollment data • Claims/encounter data	• Pre-test/post-test	• Annually/ Monthly	• Gender • Age • Race/ethnicity • Geography • Criminal justice involvement • High-cost high needs
Research Question 1.4: Does Behavioral Health Reform increase the percentage of beneficiaries who adhere to treatment for SUD and BH disorders?					
1-34: Number of beneficiaries with a SUD diagnosis including those with OUD who used services in the last month or year, by service or benefit type	• N/A	• State eligibility and enrollment data • Claims/encounter data	• Pre-test/post-test • ITS	• Annually/ Monthly	• Gender • Age • Race/ethnicity • Geography • Criminal justice involvement • High-cost high needs
1-35: Number of beneficiaries with a MH diagnosis who used services in the last month or year, by service or benefit type	• N/A	• State eligibility and enrollment data • Claims/encounter data	• Pre-test/post-test • ITS	• Annually/ Monthly	• Gender • Age • Race/ethnicity • Geography • Criminal justice involvement • High-cost high needs

Measure(s)	Comparison Group(s)	Data Source(s)	Analytic Approach	Frequency	Stratification
1-36: Time to treatment among beneficiaries who access SUD treatment	• N/A	• State eligibility and enrollment data • Claims/encounter data	• Pre-test/post-test • ITS	• Annually/ Monthly	• Gender • Age • Race/ethnicity • Geography • Criminal justice involvement • High-cost high needs
Hypothesis 2: Beneficiaries receiving BH services will experience maintained or improve health and well-being outcomes.					
Research Question 2.1: Does Behavioral Health Reform increase the percentage of beneficiaries with an SUD or a MH disorder who experience care for comorbid conditions?					
2-1: Percentage of adults who accessed preventive/ambulatory health services	• N/A	• State eligibility and enrollment data • Claims/encounter data	• Pre-test/post-test	• Annually	• Gender • Age • Race/ethnicity • Geography • Criminal justice involvement • High-cost high needs
2-2: Percentage of beneficiaries 3–21 years of age with a well-care visit with a PCP or OB/GYN	• N/A	• State eligibility and enrollment data • Claims/encounter data	• Pre-test/post-test	• Annually	• Gender • Age • Race/ethnicity • Geography • Criminal justice involvement • High-cost high needs

Measure(s)	Comparison Group(s)	Data Source(s)	Analytic Approach	Frequency	Stratification
2-3: Screening for chronic conditions relevant to State Medicaid population	• N/A	• State eligibility and enrollment data • Claims/encounter data	• Pre-test/post-test	• Annually	• Gender • Age • Race/ethnicity • Geography • Criminal justice involvement • High-cost high needs
2-4: Screening for co-morbidity of MH disorders and SUDs	• N/A	• State eligibility and enrollment data • Claims/encounter data	• Pre-test/post-test	• Annually	• Gender • Age • Race/ethnicity • Geography • Criminal justice involvement • High-cost high needs
2-5: Prevention Quality Chronic Composite	• N/A	• State eligibility and enrollment data • Claims/encounter data	• Pre-test/post-test	• Annually	• Gender • Age • Race/ethnicity • Geography • Criminal justice involvement • High-cost high needs
2-6: Pediatric Quality Chronic Composite	• N/A	• State eligibility and enrollment data • Claims/encounter data	• Pre-test/post-test	• Annually	• Gender • Age • Race/ethnicity • Geography • Criminal justice involvement • High-cost high needs

Measure(s)	Comparison Group(s)	Data Source(s)	Analytic Approach	Frequency	Stratification
2-7: Percentage of beneficiaries who have a high rating of their healthcare quality (8, 9, or 10 on a scale of 0–10)	• National benchmarks	• Beneficiary survey	• Pre-test/post-test	• N/A	• Gender • Age • Race/ethnicity
2-8: Percentage of beneficiaries who rate their overall mental or emotional health as "very good" or "excellent"	• National benchmarks	• Beneficiary survey	• Pre-test/post-test	• N/A	• Gender • Age • Race/ethnicity
2-9: Percentage of beneficiaries who demonstrate "very good" or "excellent" knowledge of available SUD/MH treatment and services	• N/A	• Beneficiary survey	• Pre-test/post-test	• N/A	• Gender • Age • Race/ethnicity
2-10: Percentage of beneficiaries who are knowledgeable of the number of SUD and MH services available	• N/A	• Beneficiary survey	• Pre-test/post-test	• N/A	• Gender • Age • Race/ethnicity
2-11: Percentage of mothers who often or always felt depressed since their new baby was born	• Synthetic control group • State/National rates	• PRAMS	• Pre-test/post-test • DiD • Synthetic controls	• Annually	• N/A
2-12: Percentage of beneficiaries who indicated poor MH in the last 30 days	• Synthetic control group • State/National rates	• BRFSS	• Pre-test/post-test • DiD • Synthetic controls	• Annually	• N/A
2-13: Percentage of mothers who indicate that they have someone who would help them while sick	• N/A	• CUBS	• Pre-test/post-test	• Annually	• N/A
2-14: Desire to obtain SUD/MH treatment and obtainment of SUD/MH treatment in the past three months	• N/A	• CUBS	• Pre-test/post-test	• Annually	• N/A
Research Question 2.2: Does Behavioral Health Reform reduce the number of beneficiaries who experience or are exposed to adverse events?					
2-15: Percentage of mothers who reported that during the past 12 months, their husband or partner pushed, hit, slapped, kicked, choked, or physically hurt them in any other way	• Synthetic control group • State/National rates	• CUBS • PRAMS	• Pre-test/post-test • DiD • Synthetic controls	• Annually	• N/A

Measure(s)	Comparison Group(s)	Data Source(s)	Analytic Approach	Frequency	Stratification
2-16: Percentage of mothers who reported that in the past 12 months, their husband or partner threatened them, limited their activities against their will, or made them feel unsafe in any other way	• Single state comparison: New York	• CUBS • PRAMS	• Pre-test/post-test • DiD	• Annually	• N/A
2-17: Percentage of respondents whose child lived with someone who had a problem with alcohol or drugs	• National benchmarks	• CUBS • National Survey of Children's Health	• Pre-test/post-test • DiD	• Annually	• N/A
2-18: Percentage of respondents whose child lived with someone who was mentally ill, suicidal, or severely depressed	• National benchmarks	• CUBS • National Survey of Children's Health	• Pre-test/post-test • DiD	• Annually	• N/A
2-19: Percentage of respondents whose child witnessed violence or physical abuse between household beneficiaries	• N/A	• CUBS	• Pre-test/post-test	• Annually	• N/A
2-20: Percentage of respondents who reported that their child saw or heard parents or adults slap, hit, kick, or punch one another in the home	• National benchmarks	• National Survey of Children's Health	• Pre-test/post-test • DiD	• Annually	• N/A
2-21: Maltreatment types among victims	• National benchmarks	• Children's Bureau	• Pre-test/post-test • DiD	• Annually	• N/A
2-22: Caregiver risk factors among child victims	• National benchmarks	• Children's Bureau	• Pre-test/post-test • DiD	• Annually	• N/A
2-23: Maternal use of marijuana or cannabis in any form	• Synthetic control group • State/National rates	• PRAMS	• Pre-test/post-test • DiD • Synthetic controls	• Annually	• N/A
2-24: Frequency of maternal marijuana or cannabis use in the past 30 days	• N/A	• CUBS	• Pre-test/post-test	• Annually	• N/A

Measure(s)	Comparison Group(s)	Data Source(s)	Analytic Approach	Frequency	Stratification
Research Question 2.3: Does Behavioral Health Reform decrease the rate of drug overdoses and overdose deaths due to opioids?					
2-25: Rate of overdose deaths, specifically overdose deaths due to any opioid	• Synthetic control group • State/National rates	• State eligibility and enrollment data • Claims/encounter data • Vital Records • CDC WONDER	• Pre-test/post-test • DiD • Synthetic Controls	• Annually	• Gender • Age • Race/ethnicity • Geography • Criminal justice involvement • High-cost high needs
2-26: Non-fatal overdoses (all cause)	• N/A	• State eligibility and enrollment data • Claims/encounter data	• Pre-test/post-test	• Annually	• Gender • Age • Race/ethnicity • Geography • Criminal justice involvement • High-cost high needs
2-27: Use of opioids at high dosage in persons without cancer	• N/A	• State eligibility and enrollment data • Claims/encounter data	• Pre-test/post-test	• Annually	• Gender • Age • Race/ethnicity • Geography • Criminal justice involvement • High-cost high needs
2-28: Use of opioids from multiple providers	• N/A	• State eligibility and enrollment data • Claims/encounter data	• Pre-test/post-test	• Annually	• Gender • Age • Race/ethnicity • Geography • Criminal justice involvement • High-cost high needs

Measure(s)	Comparison Group(s)	Data Source(s)	Analytic Approach	Frequency	Stratification
2-29: Risk of continued opioid use	• N/A	• State eligibility and enrollment data • Claims/encounter data	• Pre-test/post-test	• Annually	• Gender • Age • Race/ethnicity • Geography • Criminal justice involvement • High-cost high needs
Hypothesis 3: Behavioral Health Reform will maintain or reduce the cost of Medicaid for Alaska and the federal government.					
Research Question 3.1: Does Behavioral Health Reform maintain or reduce Alaska's per capita Medicaid BH costs?					
3-1: Total costs of healthcare (sum of parts below), by State and federal share	• N/A	• State eligibility and enrollment data • Claims/encounter data	• ITS	• Annually/ Monthly	• Gender • Age • Race/ethnicity • Geography • Criminal justice involvement • High-cost high needs
3-2: Total cost of SUD, SUD-IMD, and SUD-Other, and Non-SUD, by setting, including claims data (IP, OP, Rx, LTC)	• N/A	• State eligibility and enrollment data • Claims/encounter data	• ITS	• Annually/ Monthly	• Gender • Age • Race/ethnicity • Geography • Criminal justice involvement • High-cost high needs

Measure(s)	Comparison Group(s)	Data Source(s)	Analytic Approach	Frequency	Stratification
3-3: Total cost of MH diagnosis by IMD and Other, by setting, including claims data (IP, OP, Rx, LTC)	• N/A	• State eligibility and enrollment data • Claims/encounter data	• ITS	• Annually/ Monthly	• Gender • Age • Race/ethnicity • Geography • Criminal justice involvement • High-cost high needs

[1] For each of the key informant interviews, two rounds of interviews will take place to assess change in responses throughout key intervals of the evaluation period.
Note: ADD: attention deficit disorder; ADHD: attention deficit hyperactivity disorder; ASD: autism spectrum disorder; BRFSS: Behavioral Risk Factor Surveillance System; CDC: Centers for Disease Control; CUBS: Childhood Understanding Behaviors Survey; DBH: Department of Behavioral Health; DiD: difference-in-difference; ED: emergency department; ID: identification; IMD: Institutions for Mental Diseases; IP: inpatient; ITS: interrupted time series; LTC: long term care; MH: mental health; NQF: National Quality Forum; NSDUH: National Survey on Drug Use and Health; OB/GYN: obstetrician/gynecologist; OP: outpatient; OUD: opioid use disorder; PCP: primary care practitioner; PRAMS: Pregnancy Risk Assessment Monitoring System; Rx: medical prescription; SUD: substance use disorder; WONDER: Wide-ranging Online Data for Epidemiologist Research

Data Sources

Multiple data sources will be used to evaluate the eight research questions for the evaluation.

- Administrative Data
  - Medicaid claims and eligibility data
  - Provider enrollment data
  - Vital records
- National, State, and Beneficiary Surveys
  - Survey of Alaska Medicaid members
  - BRFSS data
  - Alaska Childhood Understanding Behaviors Survey (CUBS) data
  - National Survey of Children's Health
  - National Survey on Drug Use and Health (NSDUH) data
  - Pregnancy Risk Assessment Monitoring System (PRAMS) data
- Additional Data Sources
  - Centers for Disease Control and Prevention (CDC) Wide-Ranging Online Data for Epidemiologist Research (WONDER)
  - Children's Bureau
- Key Informant Interviews

Data will be collected from beneficiary surveys regarding beneficiaries' experiences with improvements in care coordination and integration, as well as their experiences with ease of access to healthcare, care quality, and health improvements. The beneficiary surveys utilized questions from the Consumer Assessment of Healthcare Providers and Systems (CAHPS)®[3-1] survey and will include additional questions customized to assess beneficiary knowledge of SUD and BH services in the State. Additional data will be collected from interviews with provider stakeholder, non-provider stakeholders, and THOs regarding interviewees' perspectives on the continuation of SUD and BH services, program sustainability, and program successes and challenges.

Administrative Data

Administrative data supplied by DBH will be utilized to calculate most measures identified in the evaluation design. These data include fee-for-service (FFS) claims, recipient eligibility and demographic data, and provider information. Due to changes in the processing of SUD Medicaid claims in 2020, multiple claims data sources may be combined to provide the most complete picture of Alaska Medicaid claims possible.[3-2] For the Interim Evaluation Report, DBH supplied three primary data sources: data used for the legislative audit; quarterly data from the State's administrative services organization (ASO), Optum, and weekly financial data. Beginning on January 1, 2025, DBH will manage all in a single Medicaid Management Information System (MMIS) data source.

3-1 CAHPS® is a registered trademark of the Agency for Healthcare Research and Quality (AHRQ).

3-2 Alaska Medicaid Section 1115 SUD Demonstration Status Report. 2020. Available at: https://www.medicaid.gov/medicaid/section-1115-demonstrations/downloads/state-annual-report-demostration-yr2-deliverable.pdf. Accessed on: Jul 31, 2024.

Administrative data will be cleaned, validated, and transformed to be suitable for each analysis. The exact data validation processes will vary across the specific data sources to be used for the evaluation, depending on the nature of the data being evaluated as well as the analysis for which the data are being prepared. Data quality will be assessed through:

- **Completeness**: The completeness of data is assessed through the degree to which required fields or measures are fully populated with data. Data that are reported as Not Available or Not Reportable may be considered complete depending on the specific nature of the data fields.
- **Validity**: The validity of data sets is assessed through the degree to which data are clinically and mathematically within required constraints. Data fields will be verified to ensure they are within an appropriate and credible range through a comparison of values to valid value tables as well as national and regional averages as appropriate to the data field.
- **Reliability**: The reliability of the data is assessed through the degree to which equivalent fields in different data sets contain the same information. This will involve performing cross-field checks, ensuring that data fields and data sets contain similar values where appropriate.
- **Comprehensiveness**: The comprehensiveness of data sets is assessed through the degree to which required fields or measures are present in the data. When required measures or data are not present, additional data may be requested.

National and Beneficiary Surveys

Beneficiary Surveys

State beneficiary surveys will be used to assess beneficiaries' ability to obtain timely appointments, satisfaction and experience with healthcare, and perception that their personal doctor seemed informed about the care they received from other providers. CAHPS surveys are often used to assess satisfaction with provided healthcare services and are adapted to elicit information addressing the research hypotheses related to beneficiaries' continuity of healthcare coverage, and overall health status and utilization. Results will be compared against national benchmarks where available. The survey sampling frame will be identified through eligibility and enrollment data, with specific enrollment requirements being finalized upon inspection of the data. Typically, beneficiaries are drawn from those enrolled continuously during the last six months of the measurement period, with no more than a one-month gap in enrollment.

To the extent possible, the independent evaluator will align multiple surveys to be conducted concurrently to increase response rates across all programs with overlapping populations. A range of sampling protocols will be considered including simple random samples; stratified random samples; multistage stratifications (i.e., cluster); and targeted oversamples.

The State and its independent evaluator will seek to streamline survey administration to minimize the number of separate survey rounds required, thereby minimizing the burden on beneficiaries and maximizing the response rate. Due to Alaska's frontier geography and seasonal factors, a telephone survey will be completed in May and June 2026 and 2028, to maximize the response rate.[3-3,3-4] The sampling strategy described above may be revised

[3-3] Bradburn NM, Sudman, S, Wansink B. *Asking Questions: The Definitive Guide to Questionnaire Design*. San Francisco, CA, Jossey-Bass; 2004.

[3-4] Tourangeau, R, Rips, LJ, Rasinski, K. *The Psychology of Survey Response*. New York: Cambridge University Press; 2000.

based on enrollment. Two survey instruments will utilize questions from CAHPS surveys and will include additional questions customized to assess beneficiary knowledge of SUD and BH services within the State. The expected sample will provide sufficient statistical power to detect a difference in a rate of at least 10.2 percentage points with 95 percent confidence and 80 percent power. Assuming a response rate of approximately 15 percent with a 10 percent oversample, the maximum number of surveys to be sent is 2,500.

BRFSS

The independent evaluator will utilize the BRFSS as an out-of-state comparison group using member-level data if data are available and complete enough to support rigorous statistical testing of outcomes. BRFSS is a health-focused telephone survey developed by the Centers for Disease Control and Prevention (CDC) that collects data from approximately 400,000 adults annually across all 50 states, Washington D.C., and three territories.[3-5] The questionnaire generally consists of two components: a core component and an optional component. Beneficiary surveys will be used to assess Measure 2-8 (*Percentage of beneficiaries who rate their overall mental or emotional health as "very good" or "excellent"*) among the Waiver population; however, Measure 2-12 (*Percentage of beneficiaries who indicated poor mental health in the last 30 days*) will allow the independent evaluator to benchmark rates of Alaska Medicaid members against statewide and national rates from the BRFSS core module. The Medicaid coverage indicator from the optional/core (depending on the year) Healthcare Access module may be used to identify responses among individuals similar to Alaska Medicaid members. However, fewer than a dozen states included the optional Healthcare Access module in a given year historically, which may limit the availability and selection of potential benchmark states. This survey is further limited by the inability to identify Waiver-specific beneficiaries, meaning that the rate for the entire Medicaid population will be reported. For these measures, BRFSS results from other states will be used as a benchmark to provide context and triangulate findings to other states' Medicaid populations. Contingent on the availability of data, respondents to the BRFSS survey from all other states may serve as a comparison group to Waiver members.

To provide an understanding of the capabilities of the data for performing statistical analyses, the independent evaluator will calculate the statistical power associated with any out-of-state comparison group data and report the results.

CUBS

Data from the CUBS instrument will be utilized to assess parenting behaviors; social supports; and child safety, experiences, and development. CUBS is a research project sponsored by the Alaska Department of Health, Division of Public Health, and serves as a three-year follow-up to the Alaska PRAMS of mothers who completed PRAMS and are still living in Alaska. The independent evaluator will request to obtain anonymized beneficiary-level information using current and prior survey phases, as data are available, and questionnaires are consistent. Due to periodic changes in the survey instrument, some survey items will be added, removed, or the language will be substantively revised, which limits the ability to assess these items for the full time frame.[3-6] An advantage of the CUBS survey is that it includes an indicator for Medicaid coverage, which allows the independent evaluator to limit survey results to the Medicaid population, if the sample size is large enough to produce reliable estimates.

3-5 Centers for Disease Control and Prevention. About BRFSS. Available at: https://www.cdc.gov/brfss/about/index.htm. Accessed on: Aug 21, 2024.

3-6 Alaska Department of Health. Alaska Childhood Understanding Behaviors Survey (CUBS). Available at: https://health.alaska.gov/dph/wcfh/Pages/mchepi/cubs/default.aspx; Accessed on: Aug 21, 2024.

Some CUBS survey measures included in the evaluation design indicate that PRAMS or the National Survey of Children's Health (NSCH) data will also be utilized. In these instances, PRAMS or NSCH data will be preferred; however, due to lags in reporting or the availability of optional modules, CUBS data may be used if national data sources are not available.

NSCH

The NSCH produces data relating to children who are exposed to adverse events, such as drug abuse or domestic violence. This survey will provide insight into the change in rates of adverse events among Alaska children as well as provide national rates as a comparison.

NSDUH

The NSDUH surveys individuals regarding drug use and treatment among adults nationally. Data from this survey will be used to identify rates of substance use disorder treatment by treatment setting among Alaska respondents and those nationally, where other state-level data are available.

PRAMS

Data from the PRAMS instrument will be utilized to assess depression, domestic violence, and drug use among new mothers in Alaska. The survey is administered through the CDC and allows for the identification of Alaska residents as well as comparisons to national rates. The independent evaluator will work with the State to obtain anonymized beneficiary-level information for Alaska and obtain beneficiary-level information for other states, if feasible. Although the independent evaluator will work to develop state or national comparison groups using the PRAMS data, this effort may be limited by the extensive process required to request PRAMS data for multiple states.

Additional Data Sources

CDC WONDER

In the event that vital records data from ADOH is not available, or in the case that the independent evaluator wants to compare results among Alaska residents to those nationally, data from CDC Wonder may be used. CDC Wonder provides county and state-level data on overdose mortality. This data will be used to support the calculation of Measure 2-25, *Rate of overdose deaths, specifically overdose deaths due to any opioid* by identifying overdose deaths overall, and those attributable to opioids.

Children's Bureau

The Children's Bureau, acting under the authority of the Department of Health and Human Services (HHS), develops annual child maltreatment reports, which include state-level data on the types of maltreatment perpetrated against child victims and the risk factors among primary caregivers of child victims. These data may be available in an individual-level file; however, due to the sensitivity of the information and the inability to link children to Medicaid members, the aggregate rates included in the annual maltreatment reports will likely be utilized.

Vital Records

Vital Records data from Alaska Department of Health (ADOH) will be used to calculate Measure 2-25 *Rate of overdose deaths, specifically overdose deaths due to any opioid* by identifying overdose deaths overall, and those attributable to opioids.

Key Informant Interviews

Key informant interviews will be conducted with State administrators, providers, THOs, and consumer advocates to add depth to the quantitative results and to gather information that would otherwise be unavailable through administrative data.[3-7,3-8, 3-9] Specifically, data will be collected from interviews with provider stakeholders, non-provider stakeholders, and THOs regarding interviewees' perspectives on the continuation of SUD and BH services, program sustainability, and program successes and challenges. Key informant interviews will be conducted using a semi-structured interview protocol to allow informants the ability to provide open-ended feedback on Behavioral Health Reform.[3-10] Key informant interviews will be transcribed and imported into MAXQDA where the data will be coded to permit qualitative analysis. The transcripts, coding methodologies, and coded data will be used to answer the appropriate Hypotheses.

Analytic Methods

DiD

A DiD analysis will be performed on all non-claims based measures for which a suitable comparison group can be identified. The analysis will compare rates to the weighted national average of participating states to rates among Alaska beneficiaries. Further, rates will be compared to national rates and the rates of eligible beneficiaries who are not participating in Behavioral Health Reform, where possible. For example, this analysis may be utilized to compare Medicaid beneficiaries to the statewide average or to make comparisons among Medicaid beneficiaries. This approach will compare the changes in outcome rates between the baseline period and the evaluation period, across the intervention and comparison groups. For the DiD analysis to be valid, the comparison group must accurately represent the change in outcomes that would have been experienced by the intervention group in the absence of the program. The DiD analysis will be conducted with beneficiary-level rates, using a logistic regression model for measures with binary outcomes.

The logistic regression form of the DiD model is:

$$\ln\left(\frac{Y_{it}}{1-Y_{it}}\right) = \beta_0 + \beta_1 T + \beta_2\, post + \beta_3(post \times T) + \gamma \mathrm{D'}_{it} + \varepsilon$$

3-7 Neuman WL. *Social Research Methods: Qualitative and Quantitative Approaches.* 7th ed. Edinburgh Gate: Pearson Education Limited. 2014.

3-8 Bradley K, Heeringa J, Pohl V, et al. Selecting the Best Comparison Group and Evaluation Design: A Guidance Document for State Section 1115 Demonstration Evaluations. *Centers for Medicare & Medicaid Services* and *Mathematica Policy Research.* 2020.

3-9 Creswell JW, Creswell JD. *Research Design: Qualitative, Quantitative, and Mixed Methods Approaches*. 5th ed. Thousand Oaks, CA: SAGE; 2018.

3-10 Ibid.

where Y is the probability of an outcome for group *i* in year *t*, *T* is a binary indicator of the intervention group, *post* is a binary indicator for the evaluation period, the vector D' represents any observed confounding variables that may account for differences between the intervention and comparison groups (described in additional detail below), γ is a coefficient vector, and ε is an error term. The intercept β_0 represents the log-odds of an outcome for the comparison group during the baseline. The coefficient β_1 identifies the average difference in the log-odds of an outcome between the groups during the baseline period prior to Waiver implementation. The time period dummy coefficient β_2 captures the change in the log-odds of an outcome between the baseline and evaluation time periods for the non-intervention group. The coefficient on the interaction term β_3 represents the DiD estimate of interest in this evaluation. In other words, it is how the log-odds of an outcome for the intervention group is changed in the implementation period compared to the pre-implementation period.

The DiD approach will be used where possible, as it controls for any factors external to the program that are applied equally to both groups. When data from 2020–2021 are utilized, controls for the coronavirus disease 2019 (COVID-19) public health emergency (PHE) will be utilized. However, the method is still susceptible to external factors that may have differentially impacted one group and not the other. If sufficient pre-intervention data are available, it is possible to test whether external factors are applied equally to the intervention and comparison groups by visually verifying that both groups exhibit parallel trends in the baseline period. In the absence of treatment, the intervention and comparison groups used in DiD should experience similar changes, manifested as parallel lines during the baseline period. If the parallel trend assumption does not hold, the two-period DiD may still be useful as data during the baseline and evaluation periods will be aggregated into a single pre-intervention and post-intervention average, respectively. Furthermore, the proposed DiD model estimates a single average treatment effect, under the assumption that any heterogeneity in the treatment effect is due to random variation. This assumption is explicit in the model set-up as the DiD treatment effect is represented by a single coefficient (β_3), and therefore any heterogeneity in treatment effects between individuals cannot be modeled. The independent evaluator recognizes the limitations of this approach and will therefore consider estimating additional models such as panel data models, fixed and random effects models, or hierarchical models. Results from adjusted models will be presented and interpreted, keeping in mind the limitations of each approach. Sensitivity testing will allow the independent evaluator to better estimate program impacts by assessing a variety of logistic regression specifications, including control variables as needed to improve the model fit or assess changes in key variables of interest.

If a valid comparison group cannot be constructed, the most rigorous method supported by the data will be utilized.

ITS

When a suitable comparison group cannot be found and data can be collected at multiple points in time before and after the implementation of the program, an ITS methodology can be used. This analysis is quasi-experimental in design and will compare a trend in outcomes between the baseline period and the evaluation period for those who were subject to the program.

In ITS, the measurements taken before a demonstration was initiated are used to predict the outcome if the demonstration did not occur. The measurements collected after the demonstration are then compared to the predicted outcome to evaluate the impact the demonstration had on the outcome.

The ITS model is:

$$Y_t = \beta_0 + \beta_1\, time + \beta_2\, post + \beta_3\, time \times post + \gamma \mathbf{D'}_{it} + \mu_t$$

where Y_t is the outcome of interest for the time period *t, time* represents a linear time trend, *post* is a dummy variable to indicate the time periods post-implementation, $time \times post$ is the linear time trend variable for the post-implementation time period, the matrix **D'** represents any observed confounding variables that may account for differences between the intervention and comparison groups, and γ is a coefficient vector. For ITS analyses utilizing aggregate-level data, confounding variables will take the form of average values in the population, such as average age, average risk score, or percent female. For analysis utilizing individual-level data, control variables may include age, sex, race/ethnicity, county of residence, Chronic Illness and Disability Payment System (CDPS) risk score, dual eligibility status, or duration of Medicaid enrollment. The intercept, β_0, identifies the starting level of outcome *Y*, β_1 is the slope of the outcome between the measurements before the program, β_2 is the change in the outcome when the program began, β_3 is the change in the slope for the measurements after the program, and μ_t is the error term.

Comparative ITS may be used to assess measures where there are sufficient pre-implementation data points and a valid comparison group. This analysis will be estimated using linear regression modeling of the following comparative ITS equation:

$$Y_t = \beta_0 + \beta_1 T + \beta_2 X_t + \beta_3 T X_t + \beta_4 Z + \beta_5 ZT + \beta_6 Z X_t + \beta_7 Z X_t T + \varepsilon$$

Where *Y* is the measure rate, *T* is time, *X* is study phase (pre- or post-interruption), *XT* is time after interruption, *Z* is treatment or control, *ZT* is time for treatment, *ZX* is study phase for treatment, and *ZXT* is time after interruption for treatment.

Assuming that the measurements taken after the implementation of Behavioral Health Reform would have been equal to the expectation predicted from the measurements taken before Behavioral Health Reform in the absence of the intervention, any changes in the observed rates after implementation can be attributed to the program. However, as the ITS approach relies on a pre- and post-period, it is unable to differentiate between mechanisms that may have impacted observed changes; it is possible that external events could have occurred simultaneously with Behavioral Health Reform and influenced the outcomes of interest. When data from 2020–2021 are utilized, the independent evaluator will rely on best practices to mitigate the potentially confounding effect of simultaneously occurring confounding events such as the coronavirus 2019 (COVID-19) public health emergency (PHE) as well as post-PHE Medicaid "unwinding" by including the use of dummy variables for each time period. In the context of SUD/MH services, the COVID-19 PHE made access to care more difficult due to the strain on the SUD/MH system. Specifically, MH and substance use disorders worsened during this time period due to increased isolation, depression, and anxiety.[3-11] The increased need of SUD/MH services coupled with physical distancing measures, complicated treatment delivery. To account for the impact of the COVID-19 PHE, ITS models will incorporate dummy variables to adjust for the confounding effects if sufficient data are available. An indicator variable for quarter 2 (Q2) 2020 will represent the initial wave of the COVID-19 PHE-related shutdowns and stay-at-home orders, and a separate indicator variable for Q3 2020 through the end of Q1 2021 will reflect subsequent Alaska-specific public health orders. For measures calculated annually, an indicator variable for 2020 will be included in the model to adjust for the COVID-19 PHE. As Behavioral Health Reform overlaps with the COVID-19 PHE as well as post-PHE Medicaid "unwinding," the independent evaluator will explore how the results change when excluding the years most impacted by these external events, or when estimating program effects separately by each year, rather than aggregating baseline years and evaluation years. A

[3-11] Lin, C., Clingan, S., Valdez, J., Mooney, L., and Hser, Y. (2022). The impact of COIVD-19 on substance use disorder treatment in California: Service providers' perspectives. *Journal of Substance Abuse Treatment*, 133:108544. Available at: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC8702565/. Accessed on September 9, 2024.

similar approach will be taken to account for the "unwinding" period during which the authorized Medicaid continuous enrollment condition ended and Alaska's Department of Health (DOH) began redeterminations of eligibility. Furthermore, the independent evaluator will consider several sensitivity analyses to test the robustness of the main model results. These tests may include modifying regression specifications and control variables to better estimate program impact and/or assess the degree to which findings materially change given alternative specifications. One example of sensitivity testing is the inclusion and specification of COVID-19 controls, where applicable. The most appropriate controls for each analysis utilizing an ITS approach will be identified.

A second assumption of the proposed ITS model is that the expected mean of the error term is zero; however, if current observations are correlated with prior observations, this regression assumption would be violated. The independent evaluator will test this assumption by examining error autocorrelation; if subsequent error terms are highly correlated, then parameter estimates and variance obtained from the model may be biased, resulting in misleading conclusions. During analyses, the independent evaluator will take steps to test for autocorrelation and assess the model fit. If the model is a poor fit for the data, additional procedures will be explored such as transforming the model to remove autocorrelation or estimating an autoregressive model.

A limitation of ITS is the need for sufficient data points both before and after program implementation.[3-12, 3-13, 3-14] To facilitate this methodology, the independent evaluator may consider additional baseline data points using prior year calculations, and/or calculating quarterly rates where feasible, if multiple years both pre- and post-implementation are available to control for seasonality.

Synthetic Controls

The synthetic control method will compare changes in outcome rates between the baseline and evaluation periods for both the intervention and synthetic comparison group. This method allows the independent evaluator to construct a comparison group from states with similar pre-implementation outcomes that did not implement similar policies during the post-implementation period. States are selected as "donors" based on their observed outcomes and are assigned weights according to how closely their pre-implementation outcomes align with those of the intervention state.

The form of the synthetic control model is:

$$\hat{\tau}_{1t} = Y_{1t} - \sum_{j=2}^{J+1} w_j * Y_{jt}$$

Where $\hat{\tau}_{1t}$ represents the difference between the estimate of outcome on the treated unit and the synthetic unit at year *t*. Y_{jt} represents the outcome for the intervention state *j* at in year *t*. The model then estimates the outcome for

3-12 Baicker K, Svoronos, T, (2019) Testing the Validity of the Single ITS Design. *NBER Working Paper 26080.* Available at: https://www.nber.org/papers/w26080.pdf. Accessed on: Aug 21, 2024.

3-13 Bernal JL, Cummins S, Gasparrini A. (2017) Interrupted time series regression for the evaluation of public health interventions: a tutorial. *International Journal of Epidemiology,* 46(1): 348-355. Available at: https://doi.org/10.1093/ije/dyw098. Accessed on: Aug 21, 2024.

3-14 Penfold RB, Zhang F. (2013) Use of Interrupted Time Series Analysis in Evaluating Health Care Quality Improvements. *Academic Pediatrics*, 13(6): S38 - S44. Available at: https://doi.org/10.1016/j.acap.2013.08.002. Accessed on: Aug 21, 2024.

"donor" states in the analysis, represented as the sum of $J + 1$, assuming j=1 represents the intervention state. The weight for "donor" state j is then multiplied by the outcome for the "donor" state at year t.[3-15]

This model requires that donor states did not implement a similar intervention and that there is sufficient pre-implementation data to construct a reliable counterfactual. However, meeting this assumption may be challenging, as other states with a relatively high prevalence of SUD, MH disorders, or other outcomes of interest, such as opioid deaths, may also enact a SUD/MH Medicaid waiver or other similar policies to increase identification of and access to treatment services. Additionally, for measures that rely on survey data, survey instruments can change over time and may not always be comparable with prior years. This may limit the pre and post implementation data that can be included in the analysis, and therefore, the viability of the synthetic control method. To further refine the synthetic control model, the independent evaluator may conduct sensitivity analyses to identify and include appropriate control variables.

The synthetic control method may be used on measures that utilize national survey data and Measure 2-25 *Rate of overdose deaths, specifically overdose deaths due to any opioid*, to compare the rate of overdoses among Alaska residents to those included in the synthetic control group.

To develop a control that reflects Alaska's trend prior to the implementation of Behavioral Health Reform, the independent evaluator will select and weigh "donor" states that closely align with pre-implementation outcomes for the state of Alaska. These states will be weighted based on observed characteristics such as rates of substance use and BH disorders, demographic composition, and the absence of similar policies during the post-implementation period.

Pre-test/post-test

For measures with consistent specifications over time for which national or regional benchmarks are not available, and which have too few observations to support an ITS analysis, rates will be calculated and compared both before and after waiver renewal.[3-16] Statistical testing will be conducted through a Chi-square analysis. A Chi-square test allows for comparison between two groups that have a categorical outcome, such as survey results or numerator compliance, to determine if the observed counts differ from the expectation. Specifically, comparisons will be made using this model:

$$Y = \beta_0 + \beta_1 * post$$

where Y is the rate of the outcome being measured each year, β_0 captures the average rate in the baseline years, and the coefficient β_1 for the dummy variable, *post*, representing the evaluation years, captures the change in average outcome between the baseline and evaluation time periods. For measures that utilized pre/post-testing, a weighted average of the evaluation period is also presented and represents a pooled average of the numerator and denominator counts across all three evaluation years.

3-15 Abadie, A, (2021). Using Synthetic Controls: Feasibility, Data Requirements, and Methodological Aspects. *Journal of Economic Literature*, 59(2), 391-425. Available at: Using Synthetic Controls: Feasibility, Data Requirements, and Methodological Aspects (aeaweb.org). Accessed on Aug 20, 2024.

3-16 Because measures are calculated on an annual reporting period, the post-implementation period during the current demonstration approval period of three years is insufficient to support an ITS analysis.

Binomial logistic regression may be utilized to evaluate measures that are binary outcomes or presented as rates. Due to limited options for comparison groups, it is difficult to conclude whether the changes in rates are a direct result of the specific program, as simultaneous external factors occurring during the same time period may have also had an impact that could not be accounted for. When possible, control variables will be utilized to better isolate program impact. These variables may include controls for confounding events, such as the COVID-19 PHE.

Non-Inferiority Testing

To support testing of hypotheses that suggest program impacts will "maintain or improve" or "maintain or reduce," the independent evaluator may consider employing noninferiority statistical testing. Specifically, this approach can be utilized for measures that employ a pre-test/post-test, ITS, or DiD framework.

For measures that include a pre/post or ITS framework, non-inferiority testing can be performed to determine whether measure rates in the evaluation period were meaningfully different from rates in the baseline period (i.e., to statistically test whether rates were "the same or better" than baseline rates). Non-inferiority testing allows for an assessment of meaningful difference in rates by comparing the change in rates between the baseline and evaluation period to a predetermined threshold. This threshold represents the greatest difference between the baseline and evaluation period that can exist while still being considered "equivalent." Specifically, the predetermined threshold (δ) will be calculated using the following variation of the Cohen's *h* equation:

$$\delta = P_2 - \sin\left(\frac{2 * \arcsin(\sqrt{P_2}) \pm h}{2}\right)^2$$

Where P_2 is the baseline average rate and *h* is the chosen Cohen's *h* effect size. While an effect size of 0.20 has commonly been deemed to represent a "small" effect as originally suggested by Cohen, Cohen writes, "the terms 'small,' 'medium,' and 'large' are relative, not only to each other, but to the area of behavioral science or even more particularly to the specific content and research method being employed in any given investigation" (p. 25).[3-17] Because the application of effect size in this context is to identify a minimum acceptable difference between proportions while still considering them "equal" for practical purposes, a stricter threshold than what may be typically used is appropriate. Therefore, δ for each measure was calculated based off Cohen's *h* of 0.05 (differences between proportions).

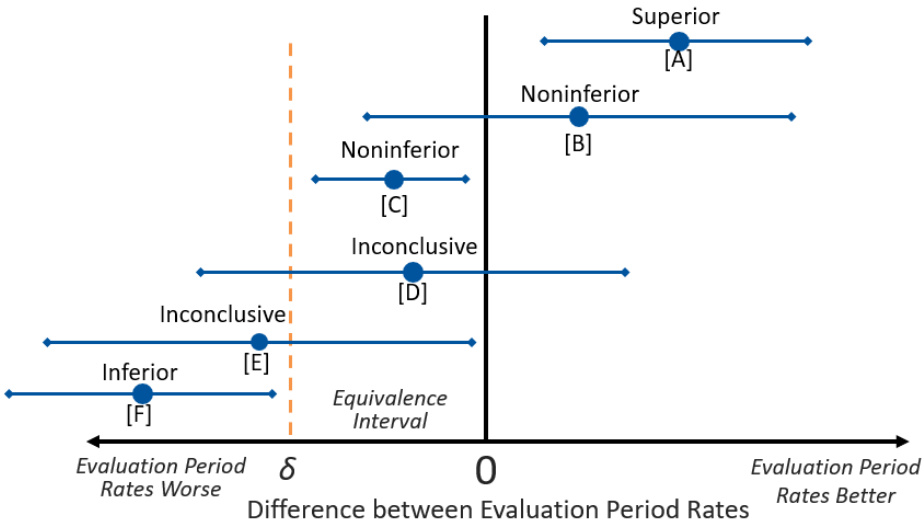
Statistical testing will be conducted by assessing whether the observed difference between the average baseline and evaluation period rates is different from δ. The calculated change in rate threshold will be compared to the 95 percent confidence intervals (CI) from performed pre-test/post-test results to determine whether rates were meaningfully different in the demonstration period.

Non-inferiority testing characterizes results in one of four ways as shown in Figure 3-1: superior, non-inferior, inconclusive, or inferior. Superior results [A] indicate the CI from the pre-test/post-test is entirely above both the predefined threshold value and zero (i.e., the pre-test/post-test is found to be statistically significant). Non-inferior findings [B/C] indicate that while results from statistical testing may be inconclusive or significantly worsening, non-inferiority testing shows any worsening in rates are not practically/clinically significant and therefore can be characterized as being not inferior to baseline rates. Inconclusive findings [D/E] occur when the 95 percent CI

[3-17] Cohen, J. (1988). Statistical Power Analysis for the Behavioral Sciences.

captures the non-inferiority threshold value. Inferior results [F] indicate the CI from the pre-test/post-test is entirely below the predefined threshold value. Figure 3-1 presents both the technical terms and the simplified terms utilized to interpret non-inferiority findings. The results utilize the simplified terms for ease of interpretability.

Figure 3-1—Non-Inferiority Testing



For measures that use a DiD framework and are hypothesized to perform at least as well as or better than a comparison group, a prespecified fraction (δ) of the change in the comparison group (coefficient on time, β_2) is used to define an "equivalence range" which would conclude that the treatment group performed as well as the comparison group. The equivalence range is bounded by the change in rates for the comparison group, plus or minus 10 percent of the change in the comparison group. The change in the treatment group will be compared against this equivalence range using a 95 percent confidence interval. Figure 3-2 illustrates how the equivalence window will be calculated and how statistical significance will be determined.

Figure 3-2—Illustration of Non-Equivalence Testing Procedure

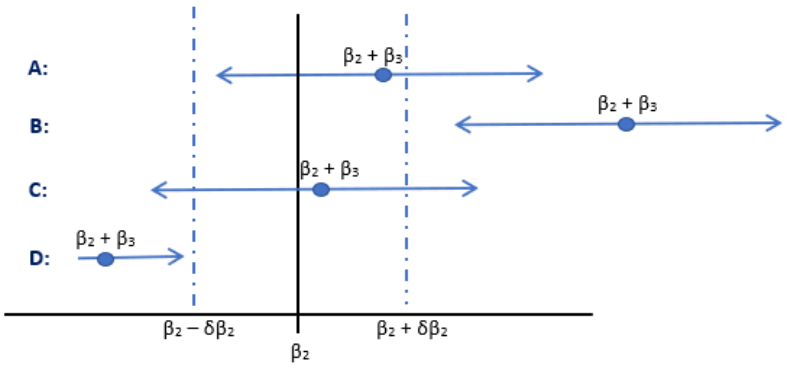


Table 3-4 defines the equivalence intervals used for each scenario in Figure 3-2.

Table 3-4—Noninferiority Equivalence Intervals

Desired Direction	Equivalence Interval	Noninferiority Threshold
Higher is better and $\beta_2 > 0$ OR Lower is better and $\beta_2 < 0$	$(\beta_2 - \delta\beta_2)$ to β_2	$(\beta_2 - \delta\beta_2)$
Lower is better and $\beta_2 > 0$ OR Higher is better and $\beta_2 < 0$	β_2 to $(\beta_2 + \delta\beta_2)$	$(\beta_2 + \delta\beta_2)$

In Figure 3-2, given a measure in which higher is better, the confidence interval in Scenario A, denoted by the arrows, includes β_2 but not the noninferiority threshold, $(\beta_2 - \delta\beta_2)$. Therefore, evidence supports the finding that the treatment group is not inferior to the comparison group. The confidence interval in Scenario B is above β_2, which suggests that the treatment group is superior to the comparison group. The confidence interval in scenario C spans both β_2 and $(\beta_2 - \delta\beta_2)$. Therefore, there is insufficient evidence to establish noninferiority and the results are inconclusive. The confidence interval in Scenario D falls below the noninferiority threshold $(\beta_2 - \delta\beta_2)$ and supports the finding that the treatment group is inferior to the comparison group.

Health Equity Analysis

To provide a more comprehensive view of the difference in rates among demographic groups, a health equity analysis will be conducted. A detailed assessment of changes in health disparities across time will be the primary analytic approach for assessing health equity. Outcome measures for relevant demographic subgroups (e.g., age, sex, race, ethnicity, etc.) will be compared to a reference group and assessed for statistically significant differences as well as clinically meaningful differences in relative percentages and effect sizes. Statistical significance will be calculated through a two-tailed t-test among the reference and comparison groups, while clinically meaningful outcomes will be assessed through effect sizes and relative percentage differences between the groups of interest.

Cohen's *h* or Hedge's *g* will be utilized to determine the effect size between comparison and reference group rates. Effect sizes can fall into small, medium, or large categories.[3-18] This method is applicable to measures where the rate is bounded between 0 and 1.

The formula for Cohen's *h* is given by:

$$h = \left(2 * arcsin\sqrt{P_1}\right) - \left(2 * arcsin\sqrt{P_2}\right)$$

Where P_1 is the annual rate for the comparison group and P_2 is the annual rate for the reference group. The effect size will be displayed with shaded boxes indicating the magnitude and direction of the results.

The formula for Hedge's *g* is given by:

[3-18] Cohen, J. Statistical Power Analysis for the Behavioral Sciences, 2nd Ed. Hillsdale, N.J.: L. Erlbaum Associates; 1988:25

$$g = \frac{\bar{y}_1 - \bar{y}_2}{s_p}$$

$$s_p = \sqrt{\frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{(n_1 - 1) + (n_2 - 1)}}$$

Where g represents the effect size between two group means, calculated as the mean of the comparison and reference group, $\bar{y}_1$ and $\bar{y}_2$, respectively, divided by the pooled standard deviation, s_p.[3-19]

For measures where the rates are not bounded between 0 and 1, the relative percent difference between each demographic stratification and the appropriate reference category will be calculated. The relative percent difference is calculated by subtracting the reference group rate from the comparison group rate and then dividing by the reference group rate. The relative percent difference will be displayed using arrows indicating the magnitude and direction of the results.

For each measure that supports a health equity analysis approach, the rates will be compared across reference and comparison demographic groups where data are available, accurate, and relevant. The independent evaluator may limit reporting to groups that have either statistically significant or clinically meaningful differences. At the time of writing this Evaluation Design it is anticipated that demographic data will be available for the following: gender, age, race, ethnicity, and geography. The independent evaluator will work with the DBH to develop a method to identify and report results by criminal justice involvement and high-cost high needs.[3-20] The measure specifications identify the demographic stratification groups for each measure, based on the anticipated available demographic data fields. Subgroup analyses will be conducted to assess program impacts by each demographic group. This allows the independent evaluator to take an exploratory approach in identifying disparities.

In accordance with the Centers for Medicare & Medicaid Services (CMS) suppression guidance, rates with a numerator or denominator greater than 1, but less than 10 will be suppressed due to potentially unreliable rate calculation and to ensure anonymity.[3-21] Furthermore, rates may be suppressed in accordance with the Healthcare Effectiveness Data and Information Set (HEDIS) general guidelines, which requires rates with denominator counts less than 30 to be suppressed to ensure reliability of reporting.[5-22] The most stringent suppression method will be used when suppressing each rate. Sample sizes will reflect the denominator counts for each subgroup by measure. The feasibility of reporting each subgroup will be dependent on numerator and denominator counts meeting suppression criteria and overall sample size.

3-19 National Institute of Standards and Technology. Available at: https://www.itl.nist.gov/div898/software/dataplot/refman2/auxillar/hedgeg.htm#:~:text=Description%3A%20The%20Hedge's%20g%20statistic,the%20pooled%20standard%20deviation%2C%20respectively. Accessed on Sept 9, 2024.

3-20 DBH and the State's criminal justice system operate as distinct entities, which complicates data-sharing to identify Behavioral Health Reform beneficiaries who are also involved with the criminal justice system.

3-21 U.S. Department of Health & Human Services. CMS Cell Suppression Policy. Available at: https://www.hhs.gov/guidance/document/cms-cell-suppression-policy. Accessed on Nov 15, 2024.

5-22 HEDIS® is a registered trademark of the National Committee for Quality Assurance (NCQA).

Comparison to National Benchmarks

To provide additional context around rates and changes in rates due to Behavioral Health Reform, the independent evaluator will compare post-implementation rates with both historical rates prior to the program and against national benchmarks without necessarily conducting formal statistical testing (e.g., DiD or pre-test/post-test approaches). By combining reference points from historical rates with contemporaneous national benchmarks, rates calculated for Behavioral Health Reform can be reported in the context of historical Alaska-specific performance in addition to performance nationally, thus triangulating an impact of Behavioral Health Reform on outcomes. Although statistical testing through a DiD or pre-test/post-test approach would be preferable, these comparisons may be necessary if the level of data for the comparison group is not granular enough to support such statistical testing.

Qualitative Synthesis

To evaluate the successes, challenges, unintended consequences, and changes to the quality of and access to care, a series of semi-structured key informant interviews with representatives of DOH, providers, and THOs will be conducted to obtain qualitative results. A qualitative synthesis will be utilized to evaluate Behavioral Health Reform.

An informative sample of key informant interviewees will be recruited from nominees identified by DBH. Interviews will invite input from appropriate individuals identified by DBH as having experience and subject matter expertise regarding the development and implementation of SUD and MH services. Each of these individuals will be requested to participate in a 60-minute interview session to provide insights into Behavioral Health Reform. Respondents from DOH, providers, and THOs will provide coverage for a comprehensive assessment across all the stakeholders involved in the planning and implementation of Behavioral Health Reform. The interviews will be conducted in two waves, in state fiscal year (SFY) 2027 and SFY 2029. 15 providers, seven State administrators, and 12 THOs will be interviewed in each wave. The number of interviews conducted is subject to change based on the saturation of feedback obtained during the interim evaluation. These proposed numbers are based on past evaluations conducted. For the second wave of interviews, the independent evaluator will attempt to re-interview the same person. If the same providers are unavailable, the independent evaluator will seek to interview another provider from the same facility or providers from a similar facility. Interviews with State administrators will cover the same administrative roles in the two waves, recognizing that these roles may or may not be filled by the same individuals.

A flexible protocol will be developed for semi-structured interviews with a sample of subjects with knowledge of the specific strategies developed and implemented as a result of Behavioral Health Reform, as well as the successes and barriers encountered during the implementation of SUD and MH services. Interview questions will be developed to seek information about stakeholders' experiences with the implementation of Behavioral Health Reform, including:

- Providers' reported successes and barriers to maintaining the integration of SUD and MH services.
- Administrators' reported successes and barriers to maintaining SUD and MH services.
- THOs' reported continued changes in quality of care and access to care in providing SUD and MH services.

Early interviews will inform the development and choice of topics and help inform the selection of additional interview subjects to round out the list of individuals to be interviewed for this project. Open-ended questions will be used to maximize the diversity and richness of responses and ensure a more holistic understanding of the

subject's experience. Probing follow-up questions will be used as appropriate to elicit additional detail and understanding of critical points, terminology, and perspectives. The sessions will be recorded and transcribed with participant consent.

The information obtained from these interviews will be synthesized with the results from other quantitative data analyses providing an in-depth discussion of each of the domains/objectives to be considered. As the key informant interviews are being conducted, the independent evaluator will perform ongoing and iterative review of the interview responses and notes to identify overall themes and common response patterns. Unique responses that are substantively interesting and informative will also be noted and may be used to develop probing questions for future interviews. The results of these preliminary analyses will be used to document the emergent and overarching themes related to each hypothesis. The documentation of emergent themes will be reviewed iteratively to determine if responses to interview questions are continuing to provide new perspectives and answers, or if the responses are converging on a common set of response patterns indicating saturation on a particular interview question. As additional interview data are collected, the categories, themes, and relationships will be adjusted to reflect the broader set of concepts and different types of relationships identified. The documentation of emergent themes will also be used as an initial starting point for organizing the analysis of the interview data once all interviews are completed.

Following the completion of the key informant interviews, the interview notes and transcripts will be reviewed using standard qualitative analysis techniques. The data will first be examined through open coding to identify key concepts and themes that may not have been captured as emergent themes during previous analyses. After identifying key concepts, axial coding techniques will be used to develop a more complete understanding of the relationships among categories identified by respondents in the data. The result of the open and axial coding analysis will be an account of the scope of issues raised by respondents, and a synthesis of how those concepts are related as presented in the participants' own words and experiences. This thematic coding process will ensure a thorough qualitative analysis with direct links to respondents' exact responses. The open and axial coding will be performed with a focus on identifying the dimensionality and breadth of responses to the hypothesis posed for the overall project. The open and axial coding will aim to identify additional themes and develop a more complete understanding of the themes and relationships among categories identified by respondents. Interviewee responses will be identified through the analysis to illustrate and contextualize the conclusions drawn from the research and will be used to support the development of the final report. The responses to questions related to mitigating barriers to maintaining and providing SUD and MH resources will be used to add context to the quantitative findings regarding those subjects. Opinions identifying opportunities for improving the efficacy of the Behavioral Health Reform more effective in increasing access and quality will inform the *Lessons Learned and Recommendations* chapter in the evaluation reports.

Cost Effectiveness Analysis

The cost effectiveness analysis is designed to analyze the differences between actual and projected costs for the evaluation period. Note that the cost analyses do not refer to or attempt to replicate the formal Budget Neutrality test required for Section 1115 demonstration waivers, which sets a fixed target under which waiver expenditures must fall that was set at the time the waiver was approved. The methodology for analyzing the costs for the

Behavioral Health Reform is adapted from CMS' guidance for assessing the costs of SUD or serious mental illness (SMI) evaluations.[3-23]

Cost of care for Behavioral Health Reform beneficiaries based on FFS reimbursement amounts will be calculated for each beneficiary in each month. To identify the source of treatment cost drivers for beneficiaries, total costs will be stratified by the categories of service presented in Table 3-5 for SUD and MH beneficiaries, SUD only beneficiaries, and MH only beneficiaries. Data will be aggregated across all beneficiaries to calculate per-member per-month (PMPM) costs for each month of Behavioral Health Reform and 24 months prior.[3-24] ITS analyses will be conducted for total cost of care, as well as for each of the three levels of cost stratification mentioned above. This method will project the costs incurred by the Behavioral Health Reform population during the baseline period, prior to Behavioral Health Reform, and during the evaluation period. The projected costs will represent a counterfactual estimate of the costs of the Behavioral Health Reform population during the evaluation period as if Behavioral Health Reform had never been renewed. Thus, the method will compare the actual costs of the Behavioral Health Reform population in the evaluation period to the projected counterfactual costs of this population in the evaluation period. Seasonality indicators and variables indicating time periods affected by the COVID-19 PHE and post-pandemic Medicaid "unwinding" will be included in the model to control for these factors.

Table 3-5—Categories of Service

SUD and MH Beneficiaries	• Total • IP • OP (ED and Non-ED) • LTC	• Professional • Dental • Pharmacy
SUD Only Beneficiaries	• SUD IMD • SUD Other	• Non-SUD
MH Only Beneficiaries	• MH IMD • MH Other	• Non-MH

Note: ED: emergency department; IMD: Institutions for Mental Disease; IP: inpatient; LTC: long-term care; MH: mental health; OP: outpatient; SUD: substance use disorder

As Behavioral Health Reform will provide additional coverage and services to beneficiaries, it is possible that initial costs will increase. The independent evaluator will also review the overall cost-effectiveness of the program, contrasting and comparing any additional costs incurred through the program to observed benefits of the program. The cost-effectiveness analysis will not involve a direct comparison of costs and savings as benefits of the program may be non-pecuniary in nature, such as provision of new services that previously were unavailable, increased employment opportunities leading to improved financial well-being, lower mortality rates, and

3-23 United States Department of Health and Human Services. Appendix C: Approaches to Analyzing Costs Associated with Section 1115 Demonstrations for Beneficiaries with Serious Mental Illness/Serious Emotional Disturbance or Substance Use Disorders. Available at: https://www.hhs.gov/guidance/document/appendix-c-analyzing-costs-associated-demonstrations-smised-or-sud-0. Accessed on: Aug 21, 2024.

3-24 CMS guidance describes constructing an ITS with member-level controls. However, due to a low prevalence of costs for most beneficiaries—especially when stratified by category of service—robust statistical analysis at the member-level was not feasible. CMS guidance references literature on evaluating healthcare expenditures using a two-part model as one mechanism to account for this issue; however, the method described in the literature is not applied in an ITS framework, which relies on assessing trends in costs. Given the frequency of months in which beneficiaries did not incur any costs and the unbalanced nature of the panel dataset, member-level trends could not be reliably estimated.

improved health outcomes overall. Furthermore, benefits may manifest over the long term and may not be measurable at the time of the evaluation.

When appropriate, supplemental analyses will also be conducted to assess issues that emerge during the course of the evaluation period and respond to stakeholder queries and quality improvement needs.

Disentangling Confounding Events

It is possible that co-interventions or other events coinciding with the demonstration may have confounded measure rates; as such, a comparison of rates during the baseline period to the evaluation period would not be able to disentangle those effects from demonstration effects. These effects may include policy changes at the State or federal level or changes to the substance use or the MH system. Known confounding effects will be controlled for using appropriate methods during the analysis of the demonstration.

4. Methodological Limitations

Despite the planned rigor of the evaluation, several limitations may impact the ability of the evaluation to attribute changes in performance metrics to the Behavioral Health Reform Section 1115 Demonstration Waiver.

Data Source Limitations

During the prior evaluation, claims data were provided through multiple avenues and did not always include consistent data. Beginning in January 2025, Alaska will enter into an agreement with a new claims vendor who will report all claims in a single data source. This change to administrative claims data may come with challenges for standardizing the data across prior years but will assist in the consolidation of data sources. The weekly financial claims data provided for the previous demonstration's Interim Evaluation Report did not include the data elements necessary for measure calculation. Nearly 20 percent of the Optum substance use disorder (SUD) claims were missing, resulting in issues with referential integrity between the Optum and weekly financial files. The independent evaluator will work with the Department of Behavioral Health (DBH) to validate data quality, including determining the frequency and accuracy of required fields (i.e., diagnosis code or member identification number). The independent evaluator and DBH will resolve discrepancies and report limitations in the data when applicable.

Another limitation is the use of claims data to identify screening among Behavioral Health Reform beneficiaries. Screening data may be difficult to identify in claims due to low reporting among providers. As such, the rates identified through screening may not be an accurate representation of program performance. When these limitations are present, the independent evaluator will describe them in the results.

Analytic Methods Limitations

No in-state comparison groups are available for claims-based measures because the renewal demonstration, Behavioral Health Reform, was implemented for all targeted beneficiaries in the State simultaneously and will continue to operate for all Alaska beneficiaries diagnosed with an SUD or mental health (MH) disorder.[4-1] Alaska does not operate an all-payer claims database, which eliminates the possibility of comparing Medicaid beneficiaries to other low-income beneficiaries in the State who were ineligible for Medicaid. Additionally, data from another state with similar population characteristics and Medicaid policies and procedures in place is unlikely to be available due limitations in interstate data sharing.[4-2] While Transformed Medicaid Statistical Information System (T-MSIS) data from the Centers for Medicare & Medicaid Services (CMS) has been suggested to create a viable comparison group, use of these data were not feasible at the time of this evaluation design. T-MSIS data may become available for use in forming a counterfactual comparison group for the waiver population by the time the Interim Evaluation Reports or the Summative Evaluation Report are developed. Therefore, the counterfactual comparison identified is the comparison of measure rates across the baseline and evaluation periods. For many measures, only a pre-post comparison of outcomes prior to the renewal of the

[4-1] In this evaluation, BH refers to both SUD and MH services, while 'SUD' and 'MH' refer specifically to their respective services.

[4-2] HSAG has approached other states regarding mutual data exchanges for the purposes of Section 1115 Waiver evaluations. No state has expressed interest in any such arrangement.

Behavioral Health Reform to outcomes post-demonstration implementation will be possible. However, a pre-post comparison of rates does not allow for causal inference of program effects.

Where sufficient data points are available, the independent evaluator will utilize an interrupted times series (ITS) analysis to make comparisons while accounting for underlying seasonal trends in the outcome. The results will indicate whether the measure rates increased or decreased, and whether the results represented statistically significant changes in performance. Furthermore, it is possible that co-interventions or other events coinciding with the demonstration may have confounded measure rates; as such, a comparison of rates during the baseline period to rates during the evaluation period would not be able to disentangle those effects from demonstration effects.

Comparisons to similarly situated out-of-state beneficiaries will be identified through national surveys. These national surveys include the Behavioral Risk Factor Surveillance System (BRFSS), National Survey of Drug Use and Health (NSDUH), Pregnancy Risk Assessment Monitoring System (PRAMS), and the National Survey of Children's Health (NSCH). A drawback of these surveys is that the instruments can change over time, which can complicate comparisons across years and there may be significant delays in data availability. Additionally, receiving beneficiary-level data for all states may not be feasible. The independent evaluator will present results from these surveys when the data are consistent. Furthermore, for all national data sources other than BRFSS, there is no mechanism to identify Medicaid-specific populations; therefore, the statewide rate for Alaska will be presented. BRFSS results will be limited to Medicaid respondents if the sample size is large enough for accurate reporting and if sufficient comparison states include the survey question. All surveys will be reported in the context that data collection is reliant on self-reported responses, therefore, the data derived from surveys may not always be accurate. Finally, for survey measures to utilize a synthetic control method, sufficient pre-implementation time periods must be available, and the independent evaluator must be able to identify comparison states that have similar pre-implementation outcomes, but do not implement a similar policy in the post-implementation period. Another limitation pertains to the health equity analysis. The independent evaluator recognizes that health equity is a complex subject, with significant ongoing discussions in the broader scientific community about how to measure it effectively. Since no single approach to evaluating health equity is without limitations, this evaluation utilizes multiple methods to address health equity. The proposed analysis is designed to provide an overview of changes in health disparities during the Behavioral Health Reform study period. However, it acknowledges the primary limitation that any observed changes in disparities cannot be causally attributed to the Behavioral Health Reform, as external factors may also impact the outcomes.

Finally, the evaluation of the Behavioral Health Reform is also limited by the relatively short time frame available to assess the program's long-term impact. For example, the ability of beneficiaries to access needed care may lead to health improvements in future years and tangential outcomes such as improvements in education, housing, employment, and involvement with the criminal justice system.[4-3] Although some outcomes, such as a reduction in emergency department visits, may experience an immediate decline, outcomes related to treatment adherence, SUD/MH-related health outcomes, and declines in mortality may not be evident until decade(s) later. For example, individuals with cirrhosis have been found to experience a higher likelihood of survival upon abstinence from alcohol, when compared to their counterparts who did not abstain. These outcomes may not be present for years after exposure to Behavioral Health Reform, complicating the ability to track and attribute these successes

[4-3] Berk LE. *Development Through the Lifespan*. 7th Ed. Hoboken, NJ: Pearson; 2018.

to the demonstration.[4-4] Additionally, a study conducted by Substance Abuse and Mental Health Services (SAMSHA) found that, on average, it takes nine years from the first treatment before individuals experience a full year of abstinence from substance use, which may indicate that active engagement in a SUD program may take years to achieve the desired outcome.[4-5] This further limits the ability to capture improvements in treatment adherence and health outcomes. While the renewal extends the evaluation period by five years, some outcomes may actualize on a longer time frame or may fall outside of the scope of this demonstration. To address these limitations, the independent evaluator will leverage quantitative and qualitative methods to assess changes in outcomes during the evaluation period.

4-4 Vuittonet CL, Halse M, Leggio L, Fricchione SB, Brickley M, Haass-Koffler CL, Tavares T, Swift RM, Kenna GA. (2014). Pharmacotherapy for alcoholic patients with alcoholic liver disease. *Am J Health Syst Pharm* 71(15):1265-76. Available at: 10.2146/ajhp140028. PMID: 25027533; PMCID: PMC4170837. Accessed on: Sep 6, 2024.

4-5 Beaulieu, M., Tremblay, J., Baudry, C., Pearson, J., and Bertrand, K. (2021). A systematic review and meta-analysis of the efficacy of the long-term treatment and support of substance use disorders. *Social Science and Medicine* (285): 114289. A systematic review and meta-analysis of the efficacy of the long-term treatment and support of substance use disorders - ScienceDirect. Accessed on: Sep 5, 2024.

Appendix A. Attachments

Independent Evaluator

The Department of Behavioral Health (DBH) will select an independent evaluator with experience and expertise to conduct a scientifically and rigorous Medicaid Section 1115 waiver evaluation that meets all the requirements specified in the Special Terms and Conditions (STCs).[A-1] The independent evaluator will be required to have the following qualifications:

- Knowledge of public health programs and policy
- Experience in healthcare research and evaluation
- Understanding of Alaska's programs and populations
- Expertise with conducting complex program evaluations
- Relevant work experience
- Skills in data management and analytic capacity
- Medicaid experience and technical knowledge

Based on State protocols, DBH will follow established policies and procedures to acquire an independent entity or entities to conduct the waiver evaluation. In addition, DBH will ensure that the selected independent evaluator does not have any conflicts of interest and will require the independent evaluator to sign a "No Conflict of Interest" statement.

Evaluation Budget

Table A-1 presents the cost estimate for the evaluation of the Behavioral Health Reform.

Table A-1—Evaluation Budget

Task	SFY26	SFY27	SFY28	SFY29	SFY30	SFY31	Total
Project Management (Semi-Annual Progress Reports)	$ 18,911	$ 20,046	$ 21,248	$ 22,523	$ 23,875	$ 11,777	**$ 118,380**
Key Informant Interviews (Instrument, Administration, Analysis)	$ 34,665	$ 36,742	$ -	$ 41,394	$ -	$ -	**$ 112,801**
Beneficiary Surveys (Instrument, Administration, Analysis)	$ 32,329	$ 25,652	$ -	$ 35,000	$ 30,552	$ -	**$ 123,533**
Mid-Point Assessment, Draft	$ -	$ 68,572	$ -	$ -	$ -	$ -	**$ 68,572**
Mid-Point Assessment, Final	$ -	$ 25,040	$ -	$ -	$ -	$ -	**$ 25,040**
Interim Evaluation Report, Draft	$ 23,505	$ 87,311	$ 80,009	$ -	$ -	$ -	**$ 190,825**
Interim Evaluation Report, Final	$ -	$ -	$ 54,116	$ -	$ -	$ -	**$ 54,116**
Summative Evaluation Report, Draft	$ -	$ -	$ -	$ 102,163	$ 104,757	$ -	**$ 206,920**
Summative Evaluation Report, Final	$ -	$ -	$ -	$ -	$ -	$ 69,071	**$ 69,071**
Total	**$ 109,410**	**$ 263,363**	**$ 155,373**	**$ 201,080**	**$ 159,184**	**$ 80,848**	**$ 969,258**

A-1 Centers for Medicare & Medicaid Services. Special Terms and Conditions. Available at: https://www.medicaid.gov/medicaid/section-1115-demonstrations/downloads/ak-behavioral-health-refm-aprvl-03262024.pdf. Accessed on: Aug 21, 2024.

Timeline and Major Milestones

Table A-2 is the proposed evaluation timeline for the Behavioral Health Reform Demonstration Waiver. This timeline is preliminary and subject to change based on approval of the Evaluation Design.

Table A-2—Evaluation Timeline

Task	SFY2025		SFY2026				SFY2027				SFY2028				SFY2029				SFY2030	
	CY2025			CY2026				CY2027				CY2028				CY2029				CY2030	
	Q2	Q3	Q4	Q1	Q2	Q3	Q4	Q1	Q2	Q3	Q4	Q1	Q2	Q3	Q4	Q1	Q2	Q3	Q4	Q1	Q2
Prepare and Implement Study Design																					
Conduct kick-off meeting																					
Prepare workplan																					
Data Collection																					
Obtain claims/encounters																					
Obtain member, provider, and eligibility/enrollment data																					
Obtain financial data																					
Perform gap analysis and other data quality checks																					
Conduct Analysis																					
Key Informant Interviews																					
Develop protocols																					
Conduct interviews																					
Conduct analyses																					
Non-Survey Analyses																					
Prepare and calculate metrics																					
Conduct statistical testing and comparison																					
Survey Analyses																					
Develop survey instrument																					
Field survey																					
Conduct survey analyses																					
Reporting																					
Draft Mid-Point Assessment																					
Final Mid-Point Assessment																					
Draft Interim Evaluation Report																					
Final Interim Evaluation Report																					
Draft Summative Evaluation Report																					
Final Summative Evaluation Report																					

Note: CY: calendar year; SFY: state fiscal year; Q: quarter

Proposed Measure Specifications

The tables in this section provide the detailed measure specifications for the Behavioral Health Reform evaluation.

Hypothesis 1: Behavioral Health Reform will maintain or increase the identification of and access to substance use disorder (SUD) and mental health (MH) disorder treatment services by increasing access to community-based care.

Research Question 1.1: Does Behavioral Health Reform increase the number of beneficiaries in the waiver population who are screened, referred to, and engaged in treatment for SUD and MH disorders?

Number of beneficiaries screened for symptoms of SUD using industry-recognized, evidence-based screening instruments (Measure 1-1)	
Numerator/Denominator	Numerator: Number of waiver recipients screened for symptoms of SUD Denominator: Number of waiver recipients
Comparison Population	N/A
Stratification	Gender, age, race/ethnicity, geography, criminal justice involvement, and high-cost high needs

Number of beneficiaries screened for symptoms of SUD using industry-recognized, evidence-based screening instruments (Measure 1-1)	
Measure Steward	N/A
Measure Name	N/A
Data Source	• State eligibility and enrollment data • Claims/encounter data
Desired Direction	Higher is better
Analytic Approach	Pre-test/post-test
Frequency	Annual

Number of beneficiaries screened for symptoms of MH disorders using industry recognized, evidence-based screening instruments (Measure 1-2)	
Numerator/Denominator	Numerator: Number of waiver recipients screened for symptoms of MH disorders Denominator: Number of waiver recipients
Comparison Population	N/A
Stratification	Gender, age, race/ethnicity, geography, criminal justice involvement, and high-cost high needs
Measure Steward	N/A
Measure Name	N/A
Data Source	• State eligibility and enrollment data • Claims/encounter data
Desired Direction	Higher is better
Analytic Approach	Pre-test/post-test
Frequency	Annual

Number of beneficiaries in the waiver population with a SUD or BM diagnosis, by setting (Measure 1-3)	
Numerator/Denominator	Numerator: Number of waiver recipients diagnosed with a SUD or MH disorder, stratified by the following settings for SUD: Any setting, Early Intervention, Outpatient, Intensive Inpatient, Residential and Inpatient. Stratified by the following settings for MH: Any setting, IOP or PH, IP, OP, ED, Telehealth. Denominator: Number of waiver recipients
Comparison Population	N/A
Stratification	Gender, age, race/ethnicity, geography, criminal justice involvement, and high-cost high needs
Measure Steward	N/A
Measure Name	N/A
Data Source	• State eligibility and enrollment data • Claims/encounter data
Desired Direction	No desired direction
Analytic Approach	Pre-test/post-test
Frequency	Annual

Number of child beneficiaries who are taking medication for ADD/ADHD, ASD, difficulties with emotions, concentration, or behavior (Measure 1-4)	
Numerator/Denominator	Numerator: Number of respondents indicating that they are taking medication for ADD/ADHD, ASD, difficulties with emotions, concentration, or behavior Denominator: Number of respondents to the survey age 3-17 years old
Comparison Population	National/state rates
Stratification	N/A
Measure Steward	National Survey of Children's Health
Measure Name	N/A
Data Source	National Survey of Children's Health
Desired Direction	No change or an increase in the rate supports the hypothesis
Analytic Approach	• Pre-test/post-test • DiD • Synthetic Controls
Frequency	Annual

Percentage of beneficiaries who had initiation of SUD treatment (Measure 1-5)	
Numerator/Denominator	Numerator: Number of waiver recipients in the denominator who had initiation of SUD treatment within 14 days of the index episode. Stratified by: Alcohol, Opioid, Other Denominator: Number of waiver recipients aged 13 and over during the measurement year with an SUD diagnosis and 194 days continuous enrollment prior to the episode and 47 days after the index episode.
Comparison Population	N/A
Stratification	Gender, age, race/ethnicity, geography, criminal justice involvement, and high-cost high needs
Measure Steward	CMS Adult Core Set
Measure Name	IET
Data Source	• State eligibility and enrollment data • Claims/encounter data
Desired Direction	No change or an increase in the rate supports the hypothesis
Analytic Approach	• Pre-test/post-test • ITS
Frequency	Annually/Monthly

Percentage of beneficiaries who had engagement of SUD treatment (Measure 1-6)	
Numerator/Denominator	Numerator: Number of waiver recipients in the denominator who had initiation of SUD treatment within 14 days of the index episode and two or more engagement episodes within 34 days of the initiation episode. Stratified by: Alcohol, Opioid, Other Denominator: Number of waiver recipients aged 13 and over during the measurement year with an SUD diagnosis and 194 days continuous enrollment prior to the episode and 47 days after the index episode
Comparison Population	N/A
Stratification	Gender, age, race/ethnicity, geography, criminal justice involvement, and high-cost high needs

Percentage of beneficiaries who had engagement of SUD treatment (Measure 1-6)	
Measure Steward	CMS Adult Core Set
Measure Name	IET
Data Source	• State eligibility and enrollment data • Claims/encounter data
Desired Direction	No change or an increase in the rate supports the hypothesis
Analytic Approach	• Pre-test/post-test • ITS
Frequency	Annually/Monthly

Follow-up after discharge within 7-days and 30-days from ED visits for SUD, and specifically for OUD (Measure 1-7)	
Numerator/Denominator	Numerator: Number of ED visits in the denominator with a follow-up visit for SUD. Rates are reported for follow-up visits within 7 and 30 days of the ED visit. Denominator: Number of ED visits for waiver recipients 18 years of age and older with a principal diagnosis of SUD and were continuously enrolled from the date of the ED visit through 30 days after the ED visit
Comparison Population	N/A
Stratification	Gender, age, race/ethnicity, geography, criminal justice involvement, and high-cost high needs
Measure Steward	CMS Adult Core Set
Measure Name	FUA-AD
Data Source	• State eligibility and enrollment data • Claims/encounter data
Desired Direction	No change or an increase in the rate supports the hypothesis
Analytic Approach	• Pre-test/post-test • ITS
Frequency	Annually/Monthly

Follow-up after discharge within 7-days and 30-days from ED visits for a MH disorder (Measure 1-8)	
Numerator/Denominator	Numerator: Number of ED visits in the denominator with a follow-up visit for mental illness. Rates are reported for follow-up visits within 7 and 30 days of an ED visit for mental illness. Denominator: Number of ED visits for waiver recipients 18 years of age and older with a principal diagnosis of mental illness or intentional self-harm with continuous enrollment from the date of the ED visit through 30 days after the ED visit
Comparison Population	N/A
Stratification	Gender, age, race/ethnicity, geography, criminal justice involvement, and high-cost high needs
Measure Steward	CMS Adult Core Set
Measure Name	FUM-AD
Data Source	• State eligibility and enrollment data • Claims/encounter data
Desired Direction	No change or an increase in the rate supports the hypothesis

Follow-up after discharge within 7-days and 30-days from ED visits for a MH disorder (Measure 1-8)	
Analytic Approach	• Pre-test/post-test • ITS
Frequency	Annually/Monthly

Number of Medicaid qualified SUD providers (identified by provider ID numbers) who bill for SUD services (Measure 1-9)	
Numerator/Denominator	Numerator: Number of providers billing for SUD services, stratified by region Denominator: N/A
Comparison Population	N/A
Stratification	N/A
Measure Steward	N/A
Measure Name	N/A
Data Source	• Provider data • Claims/encounter data
Desired Direction	Higher is better
Analytic Approach	Pre-test/post-test
Frequency	Annual

Number of Medicaid qualified professionals licensed in the State to provide MH who bill for MH disorder services (Measure 1-10)	
Numerator/Denominator	Numerator: Number of providers billing for MH services, stratified by region Denominator: N/A
Comparison Population	N/A
Stratification	N/A
Measure Steward	N/A
Measure Name	N/A
Data Source	• Provider data • Claims/encounter data
Desired Direction	Higher is better
Analytic Approach	Pre-test/post-test
Frequency	Annual

Number of respondents who received substance use treatment (Measure 1-11)	
Numerator/Denominator	Numerator: Number of respondents who indicated that they received substance use treatment Denominator: Number of respondents to the survey
Comparison Population	• State/National rates • Synthetic control group
Stratification	N/A
Measure Steward	NSDUH
Measure Name	N/A

Number of respondents who received substance use treatment (Measure 1-11)	
Data Source	National Survey on Drug Use and Health (NSDUH)
Desired Direction	No desired direction
Analytic Approach	• Pre-test/post-test • DiD • Synthetic Controls
Frequency	Annual

Number of respondents who were classified as needing substance use treatment (Measure 1-12)	
Numerator/Denominator	Numerator: Number of respondents that indicated that they needed substance use treatment if they met Diagnostic and Statistical Manual of Mental Disorders, 5th edition (DSM-5) criteria for a drug or alcohol use disorder or received treatment for a drug or alcohol use Denominator: Number of respondents to the survey
Comparison Population	• State/National rates • Synthetic control group
Stratification	N/A
Measure Steward	NSDUH
Measure Name	N/A
Data Source	National Survey on Drug Use and Health (NSDUH)
Desired Direction	No desired direction
Analytic Approach	• Pre-test/post-test • DiD • Synthetic Controls
Frequency	Annual

Number of respondents who received mental health treatment (Measure 1-13)	
Numerator/Denominator	Numerator: Number of respondents who indicated that they received mental health treatment Denominator: Number of respondents to the survey
Comparison Population	• State/National rates • Synthetic control group
Stratification	N/A
Measure Steward	NSDUH
Measure Name	N/A
Data Source	National Survey on Drug Use and Health (NSDUH)
Desired Direction	No desired direction
Analytic Approach	• Pre-test/post-test • DiD • Synthetic Controls
Frequency	Annual

Research Question 1.2: Does Behavioral Health Reform mitigate barriers to maintaining and providing SUD and BH resources?

Providers' experience providing SUD and MH services (Measure 1-14)	
Numerator/Denominator	Numerator: N/A Denominator: N/A
Comparison Population	N/A
Stratification	N/A
Measure Steward	N/A
Measure Name	N/A
Data Source	Key informant interviews
Desired Direction	N/A
Analytic Approach	Qualitative Synthesis
Frequency	Two rounds (Interim Evaluation Report, Summative Evaluation Report)

Providers' reported barriers maintaining the integration of SUD and MH services (Measure 1-15)	
Numerator/Denominator	Numerator: N/A Denominator: N/A
Comparison Population	N/A
Measure Steward	N/A
Measure Name	N/A
Data Source	Key informant interviews
Desired Direction	N/A
Analytic Approach	Qualitative Synthesis
Frequency	Two rounds (Interim Evaluation Report, Summative Evaluation Report)

Provider's reported successes maintaining the integration of SUD and MH services (Measure 1-16)	
Numerator/Denominator	Numerator: N/A Denominator: N/A
Comparison Population	N/A
Stratification	N/A
Measure Steward	N/A
Measure Name	N/A
Data Source	Key informant interviews
Desired Direction	N/A
Analytic Approach	Qualitative Synthesis
Frequency	Two rounds (Interim Evaluation Report, Summative Evaluation Report)

Providers' reported changes in ability to provide care after DBH procurement of a BH organization (Measure 1-17)	
Numerator/Denominator	Numerator: N/A Denominator: N/A
Comparison Population	N/A
Stratification	N/A
Measure Steward	N/A
Measure Name	N/A
Data Source	Key informant interviews
Desired Direction	N/A
Analytic Approach	Qualitative Synthesis
Frequency	Two rounds (Interim Evaluation Report, Summative Evaluation Report)

Administrators' reported barriers maintaining SUD and MH services (Measure 1-18)	
Numerator/Denominator	Numerator: N/A Denominator: N/A
Comparison Population	N/A
Stratification	N/A
Measure Steward	N/A
Measure Name	N/A
Data Source	Key informant interviews
Desired Direction	N/A
Analytic Approach	Qualitative Synthesis
Frequency	Two rounds (Interim Evaluation Report, Summative Evaluation Report)

Administrators' reported successes maintaining SUD and MH services (Measure 1-19)	
Numerator/Denominator	Numerator: N/A Denominator: N/A
Comparison Population	N/A
Stratification	N/A
Measure Steward	N/A
Measure Name	N/A
Data Source	Key informant interviews
Desired Direction	N/A
Analytic Approach	Qualitative Synthesis
Frequency	Two rounds (Interim Evaluation Report, Summative Evaluation Report)

Administrators' plan for program sustainability and anticipated challenges (Measure 1-20)	
Numerator/Denominator	Numerator: N/A Denominator: N/A
Comparison Population	N/A
Stratification	N/A
Measure Steward	N/A
Measure Name	N/A
Data Source	Key informant interviews
Desired Direction	N/A
Analytic Approach	Qualitative Synthesis
Frequency	Two rounds (Interim Evaluation Report, Summative Evaluation Report)

Alaska tribal entities' reported continued changes in quality of care and access to care providing SUD and MH services (Measure 1-21)	
Numerator/Denominator	Numerator: N/A Denominator: N/A
Comparison Population	N/A
Stratification	N/A
Measure Steward	N/A
Measure Name	N/A
Data Source	Key informant interviews
Desired Direction	N/A
Analytic Approach	Qualitative Synthesis
Frequency	Two rounds (Interim Evaluation Report, Summative Evaluation Report)

Research Question 1.3: Does Behavioral Health Reform decrease utilization of ED, IP, or institutional settings within the beneficiary population?

IP admissions for SUD, and specifically for OUD (Measure 1-22)	
Numerator/Denominator	Numerator: Total number of SUD inpatient stays, stratified by Any SUD and OUD Denominator: Number of beneficiary months among waiver recipients in the waiver population, divided by 1,000
Comparison Population	N/A
Stratification	Gender, age, race/ethnicity, geography, criminal justice involvement, and high-cost high needs
Measure Steward	N/A
Measure Name	Modified IPU
Data Source	• State eligibility and enrollment data • Claims/encounter data
Desired Direction	No desired direction
Analytic Approach	• Pre-test/post-test • ITS
Frequency	Annually/Monthly

IP admissions for MH disorders (Measure 1-23)	
Numerator/Denominator	Numerator: Total number of MH inpatient stays Denominator: Number of beneficiary months among waiver recipients in the waiver population, divided by 1,000
Comparison Population	N/A
Stratification	Gender, age, race/ethnicity, geography, criminal justice involvement, and high-cost high needs
Measure Steward	N/A
Measure Name	Modified IPU
Data Source	• State eligibility and enrollment data • Claims/encounter data
Desired Direction	No desired direction
Analytic Approach	• Pre-test/post-test • ITS
Frequency	Annually/Monthly

ED visits for SUD (Measure 1-24)	
Numerator/Denominator	Numerator: Total number of SUD ED admissions Denominator: Number of beneficiary months among waiver recipients in the waiver population, divided by 1,000
Comparison Population	N/A
Stratification	Gender, age, race/ethnicity, geography, criminal justice involvement, and high-cost high needs
Measure Steward	N/A
Measure Name	Modified AMB-ED
Data Source	• State eligibility and enrollment data • Claims/encounter data
Desired Direction	No desired direction
Analytic Approach	• Pre-test/post-test • ITS
Frequency	Annually/Monthly

ED visits for OUD (Measure 1-25)	
Numerator/Denominator	Numerator: Total number of OUD ED admissions Denominator: Number of beneficiary months among waiver recipients in the waiver population, divided by 1,000
Comparison Population	N/A
Stratification	Gender, age, race/ethnicity, geography, criminal justice involvement, and high-cost high needs
Measure Steward	N/A
Measure Name	Modified AMB-ED

ED visits for OUD (Measure 1-25)	
Data Source	• State eligibility and enrollment data • Claims/encounter data
Desired Direction	No desired direction
Analytic Approach	• Pre-test/post-test • ITS
Frequency	Annually/Monthly

Number of ED visits among beneficiaries who had at least one ED admission for SUD (Measure 1-26)	
Numerator/Denominator	Numerator: Total number of SUD ED admissions Denominator: Number of beneficiary months among waiver recipients in the waiver population with at least one ED admission for SUD, divided by 1,000
Comparison Population	N/A
Stratification	Gender, age, race/ethnicity, geography, criminal justice involvement, and high-cost high needs
Measure Steward	N/A
Measure Name	Modified AMB-ED
Data Source	• State eligibility and enrollment data • Claims/encounter data
Desired Direction	No desired direction
Analytic Approach	• Pre-test/post-test • ITS
Frequency	Annually/monthly

Number of ED visits for SUD among high utilizing beneficiaries (Measure 1-27)	
Numerator/Denominator	Numerator: Total number of SUD ED admissions Denominator: Number of beneficiary months among waiver recipients in the waiver population who fall above the 95th percentile of ED admissions for SUD, divided by 1,000
Comparison Population	N/A
Stratification	Gender, age, race/ethnicity, geography, criminal justice involvement, and high-cost high needs
Measure Steward	N/A
Measure Name	Modified AMB-ED
Data Source	• State eligibility and enrollment data • Claims/encounter data
Desired Direction	No desired direction
Analytic Approach	• Pre-test/post-test • ITS
Frequency	Annually/monthly

ED visits for MH disorders (Measure 1-28)	
Numerator/Denominator	Numerator: Total number of MH ED admissions Denominator: Number of beneficiary months among waiver recipients in the waiver population, divided by 1,000
Comparison Population	N/A
Stratification	Gender, age, race/ethnicity, geography, criminal justice involvement, and high-cost high needs
Measure Steward	N/A
Measure Name	Modified AMB-ED
Data Source	• State eligibility and enrollment data • Claims/encounter data
Desired Direction	No desired direction
Analytic Approach	• Pre-test/post-test • ITS
Frequency	Annually/monthly

Number of ED visits among beneficiaries who had at least one ED admission for MH (Measure 1-29)	
Numerator/Denominator	Numerator: Total number of MH ED admissions Denominator: Number of beneficiary months among waiver recipients in the waiver population with at least one ED admission for MH, divided by 1,000
Comparison Population	N/A
Stratification	Gender, age, race/ethnicity, geography, criminal justice involvement, and high-cost high needs
Measure Steward	N/A
Measure Name	N/A
Data Source	• State eligibility and enrollment data • Claims/encounter data
Desired Direction	No desired direction
Analytic Approach	• Pre-test/post-test • ITS
Frequency	Annually/monthly

Number of ED visits for MH among high utilizing beneficiaries (Measure 1-30)	
Numerator/Denominator	Numerator: Total number MH SUD ED admissions Denominator: Number of beneficiary months among waiver recipients in the waiver population who fall above the 95th percentile of ED admissions for MH, divided by 1,000
Comparison Population	N/A
Stratification	Gender, age, race/ethnicity, geography, criminal justice involvement, and high-cost high needs
Measure Steward	N/A
Measure Name	N/A

Number of ED visits for MH among high utilizing beneficiaries (Measure 1-30)	
Data Source	• State eligibility and enrollment data • Claims/encounter data
Desired Direction	No desired direction
Analytic Approach	• Pre-test/post-test • ITS
Frequency	Annually/monthly

Mean length of stay among IMDs measured from admission date to discharge date (Measure 1-31)	
Numerator/Denominator	Numerator: Number of days before discharge Denominator: Number of waiver recipients admitted to an IMD
Comparison Population	N/A
Stratification	Gender, age, race/ethnicity, geography, criminal justice involvement, and high-cost high needs
Measure Steward	N/A
Measure Name	N/A
Data Source	• State eligibility and enrollment data • Claims/encounter data
Desired Direction	Lower is better
Analytic Approach	Pre-test/post-test
Frequency	Annual

30-day readmission rate to IP facilities following hospitalization for an SUD-related diagnosis (Measure 1-32)	
Numerator/Denominator	Numerator: Number of hospitalizations for an SUD-related diagnosis during measurement year that were followed by an unplanned readmission to an inpatient facility Denominator: Number of waiver recipients 18 years old and older in the waiver population who were continuously enrolled with no more than one gap in enrollment of up to 45 days during the year prior to the discharge date and no gap during the 30 days following the discharge date
Comparison Population	N/A
Stratification	Gender, age, race/ethnicity, geography, criminal justice involvement, and high-cost high needs
Measure Steward	N/A
Measure Name	Modified PCR
Data Source	• State eligibility and enrollment data • Claims/encounter data
Desired Direction	Lower is better
Analytic Approach	Pre-test/post-test
Frequency	Annual

30-day readmission rate to IP facilities following hospitalization for a MH-related diagnosis (Measure 1-33)	
Numerator/Denominator	Numerator: Number of hospitalizations for an MH-related diagnosis during measurement year that were followed by an unplanned readmission to an inpatient facility Denominator: Number of waiver recipients 18 years old and older in the waiver population who were continuously enrolled with no more than one gap in enrollment of up to 45 days during the year prior to the discharge date and no gap during the 30 days following the discharge date
Comparison Population	N/A
Stratification	Gender, age, race/ethnicity, geography, criminal justice involvement, and high-cost high needs
Measure Steward	N/A
Measure Name	Modified PCR
Data Source	• State eligibility and enrollment data • Claims/encounter data
Desired Direction	Lower is better
Analytic Approach	Pre-test/post-test
Frequency	Annual

Research Question 1.4: Does Behavioral Health Reform increase the percentage of beneficiaries who adhere to treatment for SUD and MH disorders?

Number of beneficiaries with a SUD diagnosis including those with OUD who used services in the last month or year, by service or benefit type (Measure 1-34)	
Numerator/Denominator	Numerator: Waiver recipients in the denominator, stratified by the following settings: Any setting, Early Intervention, Outpatient, Intensive Inpatient, Residential and Inpatient, Withdrawal Management, MAT Denominator: Number of waiver recipients with an SUD diagnosis
Comparison Population	N/A
Stratification	Gender, age, race/ethnicity, geography, criminal justice involvement, and high-cost high needs
Measure Steward	N/A
Measure Name	N/A
Data Source	• State eligibility and enrollment data • Claims/encounter data
Desired Direction	No desired direction
Analytic Approach	• Pre-test/post-test • ITS
Frequency	Annually/Monthly

Number of beneficiaries with a MH diagnosis who used services in the last month or year, by service or benefit type (Measure 1-35)	
Numerator/Denominator	Numerator: Waiver recipients in the denominator, stratified by the following settings: Any service, IOP or PH, IP, OP, ED, Telehealth Denominator: Number of waiver recipients with a MH diagnosis
Comparison Population	N/A
Stratification	Gender, age, race/ethnicity, geography, criminal justice involvement, and high-cost high needs

Number of beneficiaries with a MH diagnosis who used services in the last month or year, by service or benefit type (Measure 1-35)	
Measure Steward	N/A
Measure Name	N/A
Data Source	• State eligibility and enrollment data • Claims/encounter data
Desired Direction	No desired direction
Analytic Approach	• Pre-test/post-test • ITS
Frequency	Annually/Monthly

Time to treatment among beneficiaries who access SUD treatment (Measure 1-36)	
Numerator/Denominator	Numerator: Number of days between the index episode start date and first date of treatment. Rates stratified by: Alcohol, Opioid, Other Denominator: Number of waiver recipients aged 13 and over during the measurement year with an SUD diagnosis and 194 days continuous enrollment prior to the episode and 47 days after the index episode
Comparison Population	N/A
Stratification	Gender, age, race/ethnicity, geography, criminal justice involvement, and high-cost high needs
Measure Steward	N/A
Measure Name	Modified IET
Data Source	• State eligibility and enrollment data • Claims/encounter data
Desired Direction	No change or an increase in the rate supports the hypothesis
Analytic Approach	• Pre-test/post-test • ITS
Frequency	Annually/Monthly

Hypothesis 2: Beneficiaries receiving BH services will experience maintained or improve health and well-being outcomes.

Research Question 2.1: Does Behavioral Health Reform increase the percentage of beneficiaries with SUD or a MH disorder who experience care for comorbid conditions?

Percentage of adults who accessed preventive/ambulatory health services (Measure 2-1)	
Numerator/Denominator	Numerator: Number of waiver recipients with an ambulatory or preventive care visit Denominator: Number of waiver recipients 20 years and older continuously enrolled throughout the measurement year with no more than one gap in enrollment of up to 45 days
Comparison Population	N/A
Stratification	Gender, age, race/ethnicity, geography, criminal justice involvement, and high-cost high needs

Percentage of adults who accessed preventive/ambulatory health services (Measure 2-1)	
Measure Steward	HEDIS[A-1]
Measure Name	AAP
Data Source	• State eligibility and enrollment data • Claims/encounter data
Desired Direction	Higher is better
Analytic Approach	Pre-test/post-test
Frequency	Annual

Percentage of beneficiaries 3–21 years of age with a well-care visit with a PCP or OB/GYN (Measure 2-2)	
Numerator/Denominator	Numerator: Waiver recipients with one or more well-care visit during the measurement year. Denominator: Number of waiver recipients aged 3-21 years who are continuously enrolled during the measurement year with no more than one gap in enrollment of up to 45 days
Comparison Population	N/A
Stratification	Gender, age, race/ethnicity, geography, criminal justice involvement, and high-cost high needs
Measure Steward	HEDIS
Measure Name	WCV
Data Source	• State eligibility and enrollment data • Claims/encounter data
Desired Direction	Higher is better
Analytic Approach	Pre-test/post-test
Frequency	Annual

Screening for chronic conditions relevant to State Medicaid population (Measure 2-3)	
Numerator/Denominator	Numerator: Number of waiver recipients screened for chronic conditions Denominator: Number of waiver recipients
Comparison Population	N/A
Stratification	Gender, age, race/ethnicity, geography, criminal justice involvement, and high-cost high needs
Measure Steward	N/A
Measure Name	N/A
Data Source	• State eligibility and enrollment data • Claims/encounter data
Desired Direction	Higher is better
Analytic Approach	Pre-test/post-test
Frequency	Annual

A-1 HEDIS® is a registered trademark of the National Committee for Quality Assurance (NCQA).

Screening for co-morbidity of MH disorders and SUDs (Measure 2-4)	
Numerator/Denominator	Numerator: Number of waiver recipients screened for chronic conditions. Two numerators are reported: 1. Number of waiver recipients with MH disorder who also have SUD 2. Number of waiver recipients with SUD who also have MH disorder Denominator: Two denominators are reported: 1. Number of waiver recipients with MH disorder 2. Number of waiver recipients with SUD
Comparison Population	N/A
Stratification	Gender, age, race/ethnicity, geography, criminal justice involvement, and high-cost high needs
Measure Steward	N/A
Measure Name	N/A
Data Source	• State eligibility and enrollment data • Claims/encounter data
Desired Direction	Higher is better
Analytic Approach	Pre-test/post-test
Frequency	Annual

Prevention Quality Chronic Composite (Measure 2-5)	
Numerator/Denominator	Numerator: Hospital discharges for long and short-term diabetes complications, COPD, hypertension, heart failure, uncontrolled diabetes, asthma, and lower-extremity amputation among waiver recipients with diabetes Denominator: Population age 18 years and older in metropolitan area or county among waiver recipients
Comparison Population	N/A
Stratification	Gender, age, race/ethnicity, geography, criminal justice involvement, and high-cost high needs
Measure Steward	AHRQ
Measure Name	PQI 92
Data Source	• State eligibility and enrollment data • Claims/encounter data
Desired Direction	Higher is better
Analytic Approach	Pre-test/post-test
Frequency	Annual

Pediatric Quality Chronic Composite (Measure 2-6)	
Numerator/Denominator	Numerator: Hospital discharges for asthma or diabetes complications Denominator: Population ages 6 to 17 years in metropolitan area or county among waiver recipients
Comparison Population	N/A
Stratification	Gender, age, race/ethnicity, geography, criminal justice involvement, and high-cost high needs
Measure Steward	AHRQ
Measure Name	PDI 92

Pediatric Quality Chronic Composite (Measure 2-6)	
Data Source	• State eligibility and enrollment data • Claims/encounter data
Desired Direction	Higher is better
Analytic Approach	Pre-test/post-test
Frequency	Annual

Percentage of beneficiaries who have a high rating of their healthcare quality (8, 9, or 10 on a scale of 0–10) (Measure 2-7)	
Numerator/Denominator	Numerator: Two rates are reported: 1. Adults rating the quality of their healthcare as very good or excellent 2. Children rating the quality of their healthcare as very good or excellent Denominator: Two denominators are calculated: 1. Adult survey question respondents 2. Child survey question respondents
Comparison Population	National benchmarks
Stratification	Gender, age, race/ethnicity[A-2]
Measure Steward	CAHPS[A-3]
Measure Name	N/A
Data Source	Beneficiary survey
Desired Direction	Higher is better
Analytic Approach	Pre-test/post-test
Frequency	Two rounds of surveys (Interim Evaluation Report, Summative Evaluation Report)

Percentage of beneficiaries who rate their overall mental or emotional health as "very good" or "excellent" (Measure 2-8)	
Numerator/Denominator	Numerator: Two rates are reported: 1. Adults rating their mental health as very good or excellent 2. Children rating their mental health as very good or excellent Denominator: Two denominators are calculated: 1. Adult survey question respondents 2. Child survey question respondents
Comparison Population	National benchmarks
Stratification	Gender, age, race/ethnicity
Measure Steward	CAHPS
Measure Name	N/A
Data Source	Beneficiary survey
Desired Direction	Higher is better
Analytic Approach	Pre-test/post-test

A-2 To minimize survey burden and ensure that stratified rates are high enough to meet reporting standards, survey measures will be limited to stratifications by gender, age, and race/ethnicity.

A-3 CAHPS® is a registered trademark of the Agency for Healthcare Research and Quality (AHRQ).

Percentage of beneficiaries who rate their overall mental or emotional health as "very good" or "excellent" (Measure 2-8)	
Frequency	Two rounds of surveys (Interim Evaluation Report, Summative Evaluation Report)

Percentage of beneficiaries who demonstrate "very good" or "excellent" knowledge of available SUD/MH treatment and services (Measure 2-9)	
Numerator/Denominator	Numerator: Two rates are reported: 1. Adults indicating that they know where to find SUD or BH services if needed 2. Children indicating that they know where to find Sud or BH services if needed Denominator: Two denominators are calculated: 1. Adult survey question respondents 2. Child survey question respondents
Comparison Population	N/A
Stratification	Gender, age, race/ethnicity
Measure Steward	N/A
Measure Name	N/A
Data Source	Beneficiary survey
Desired Direction	Higher is better
Analytic Approach	Pre-test/post-test
Frequency	Two rounds of surveys (Interim Evaluation Report, Summative Evaluation Report)

Percentage of beneficiaries who are knowledgeable of the number of SUD and MH services available (Measure 2-10)	
Numerator/Denominator	Numerator: Rates for two groups are reported: Adults, Children Among these groups rates are reported among respondents indicating that they knew where to receive all types of treatments mentioned in the survey. Stratified by the following: SUD services, MH services SUD services listed include: Group therapy, One on one, Peer support, Family therapist, Residential, MAT MH services listed include: Group therapy, One on one, Telemedicine, Family therapist, Residential, Peer support Denominator: Two denominators are calculated: 1. Adult survey respondents 2. Child survey respondents
Comparison Population	N/A
Stratification	Gender, age, race/ethnicity
Measure Steward	N/A
Measure Name	N/A
Data Source	Beneficiary survey
Desired Direction	Higher is better
Analytic Approach	Pre-test/post-test
Frequency	Two rounds of surveys (Interim Evaluation Report, Summative Evaluation Report)

Percentage of mothers who often or always felt depressed since their new baby was born (Measure 2-11)	
Numerator/Denominator	Numerator: Number of respondents indicating they are often or always depressed Denominator: Number of survey question respondents
Comparison Population	• State/National rates • Synthetic control group
Stratification	N/A
Measure Steward	PRAMS
Measure Name	N/A
Data Source	PRAMS
Desired Direction	• Pre-test/post-test • DiD • Synthetic Controls
Analytic Approach	Pre-test/post-test
Frequency	Annual

Percentage of beneficiaries who indicated poor mental health in the last 30 days (Measure 2-12)	
Numerator/Denominator	Numerator: Number of respondents that their mental health was not good in the past 30 days Denominator: Number of survey question respondents
Comparison Population	• Synthetic control group • State/National rates
Stratification	N/A
Measure Steward	BRFSS
Measure Name	N/A
Data Source	BRFSS
Desired Direction	Lower is better
Analytic Approach	• Pre-test/post-test • DiD • Synthetic Controls
Frequency	Annual

Percentage of mothers who indicate that they have someone who would help them while sick (Measure 2-13)	
Numerator/Denominator	Numerator: Number of respondents indicating they have someone who would help them while sick Denominator: Number of survey question respondents
Comparison Population	N/A
Stratification	N/A
Measure Steward	CUBS
Measure Name	N/A
Data Source	CUBS
Desired Direction	Higher is better

Percentage of mothers who indicate that they have someone who would help them while sick (Measure 2-13)	
Analytic Approach	Pre-test/post-test
Frequency	Annual

Desire to obtain SUD/MH treatment and obtainment of SUD/MH treatment in the past three months (Measure 2-14)	
Numerator/Denominator	Numerator: Number of respondents who indicated they sought SUD/MH treatment. Stratified by the following: – Those who desired to obtain treatment – Those who obtained treatment Denominator: Number of survey question respondents
Comparison Population	N/A
Stratification	N/A
Measure Steward	CUBS
Measure Name	N/A
Data Source	CUBS
Desired Direction	Higher is better
Analytic Approach	Pre-test/post-test
Frequency	Annual

Research Question 2.2: Does Behavioral Health Reform reduce the number of beneficiaries who experience or are exposed to adverse events?

Percentage of mothers who reported that during the past 12 months, their husband or partner pushed, hit, slapped, kicked, choked, or physically hurt them in any other way (Measure 2-15)	
Numerator/Denominator	Numerator: Number of respondents indicating that their husband or partner pushed, hit, slapped, kicked, choked, or physically hurt them in any other way Denominator: Number of survey question respondents
Comparison Population	• Synthetic control group • State/National rates
Stratification	N/A
Measure Steward	• CUBS • PRAMS
Measure Name	N/A
Data Source	• CUBS • PRAMS
Desired Direction	Lower is better
Analytic Approach	• Pre-test/post-test • DiD • Synthetic Controls
Frequency	Annual

Percentage of mothers who reported that in the past 12 months, their husband or partner threatened them, limited their activities against their will or made them feel unsafe in any other way (Measure 2-16)	
Numerator/Denominator	Numerator: Number of respondents indicating that their husband or partner threatened them, limited their activities against their will or make them feel unsafe in any other way Denominator: Number of survey question respondents
Comparison Population	Single state comparison: New York[A-4]
Stratification	N/A
Measure Steward	• CUBS • PRAMS
Measure Name	N/A
Data Source	• CUBS • PRAMS
Desired Direction	Lower is better
Analytic Approach	• Pre-test/post-test • DiD • Synthetic Controls
Frequency	Annual

Percentage of respondents whose child lived with someone who had a problem with alcohol or drugs (Measure 2-17)	
Numerator/Denominator	Numerator: Number of respondents who indicated their child lived with someone who had a problem with alcohol or drugs Denominator: Number of survey question respondents
Comparison Population	National benchmarks
Stratification	N/A
Measure Steward	• CUBS • NSCH
Measure Name	N/A
Data Source	• CUBS • NSCH
Desired Direction	Lower is better
Analytic Approach	• Pre-test/post-test • DiD
Frequency	Annual

[A-4] This is a supplemental PRAMS question that is included in the Alaska and New York surveys. The independent evaluator may elect to report the New York rate for context.

Percentage of respondents whose child lived with someone who was mentally ill, suicidal, or severely depressed (Measure 2-18)	
Numerator/Denominator	Numerator: Number of respondents who indicated their child lived with someone who was mentally ill, suicidal, or severely depressed Denominator: Number of survey question respondents
Comparison Population	National benchmarks
Stratification	N/A
Measure Steward	• CUBS • NSCH
Measure Name	N/A
Data Source	• CUBS • NSCH
Desired Direction	Lower is better
Analytic Approach	• Pre-test/post-test • DiD
Frequency	Annual

Percentage of respondents whose child witnessed violence or physical abuse between household beneficiaries (Measure 2-19)	
Numerator/Denominator	Numerator: Number of respondents who indicated their child witnessed violence or physical abuse between household members Denominator: Number of survey question respondents
Comparison Population	N/A
Stratification	N/A
Measure Steward	CUBS
Measure Name	N/A
Data Source	CUBS
Desired Direction	Lower is better
Analytic Approach	Pre-test/post-test
Frequency	Annual

Percentage of respondents who reported that their child saw or heard parents or adults slap, hit, kick, punch one another in the home (Measure 2-20)	
Numerator/Denominator	Numerator: Number of respondents who their child saw or heard parents or adults slap, hit, kick, punch one another in the home Denominator: Number of survey question respondents
Comparison Population	National/State rates
Stratification	N/A
Measure Steward	NSCH
Measure Name	N/A
Data Source	NSCH
Desired Direction	Lower is better

Percentage of respondents who reported that their child saw or heard parents or adults slap, hit, kick, punch one another in the home (Measure 2-20)	
Analytic Approach	• Pre-test/post-test • DiD
Frequency	Annual

Maltreatment types among victims (Measure 2-21)	
Numerator/Denominator	Numerator: Number of child victims, reported by maltreatment type: Medical neglect, Neglect, Other, Physical Abuse, Psychological maltreatment, Sexual abuse, Sex trafficking, Unknown Denominator: Number of child victims
Comparison Population	National/state rates
Stratification	N/A
Measure Steward	Children's Bureau
Measure Name	N/A
Data Source	Children's Bureau
Desired Direction	No desired direction
Analytic Approach	• Pre-test/post-test • DiD
Frequency	Annual

Caregiver risk factors among child victims (Measure 2-22)	
Numerator/Denominator	Numerator: Number of child victims, reported by risk factor: Alcohol Abuse, Domestic Violence, Drug Abuse, Inadequate housing Denominator: Number of child victims
Comparison Population	National/state rates
Stratification	N/A
Measure Steward	Children's Bureau
Measure Name	N/A
Data Source	Children's Bureau
Desired Direction	No desired direction
Analytic Approach	• Pre-test/post-test • DiD
Frequency	Annual

Maternal use of marijuana or cannabis in any form (Measure 2-23)	
Numerator/Denominator	Numerator: Number of respondents indicating that they used marijuana or hash. Stratified by time: – During the 12 months before I got pregnant – During my most recent pregnancy – Since my new baby was born Denominator: Number of survey question respondents
Comparison Population	• Synthetic control group • National/state rates
Stratification	N/A
Measure Steward	PRAMS
Measure Name	N/A
Data Source	PRAMS
Desired Direction	Lower is better
Analytic Approach	• Pre-test/post-test • DiD • Synthetic Controls
Frequency	Annual

Frequency of maternal marijuana or cannabis use in the past 30 days (Measure 2-24)	
Numerator/Denominator	Numerator: Number of days Denominator: Number of survey question respondents
Comparison Population	N/A
Stratification	N/A
Measure Steward	CUBS
Measure Name	N/A
Data Source	CUBS
Desired Direction	Lower is better
Analytic Approach	Pre-test/post-test
Frequency	Annual

Research Question 2.3: Does Behavioral Health Reform decrease the rate of drug overdoses and overdose deaths due to opioids?

Rate of overdose deaths, specifically overdose deaths due to any opioid (Measure 2-25)	
Numerator/Denominator	Numerator: Number of overdose deaths attributable to opioids Denominator: Number of Alaska residents
Comparison Population	• Synthetic control group • State/National rates

Rate of overdose deaths, specifically overdose deaths due to any opioid (Measure 2-25)	
Stratification	Gender, age, race/ethnicity, geography, criminal justice involvement, and high-cost high needs[A-5]
Measure Steward	N/A
Measure Name	N/A
Data Source	• State eligibility and enrollment data • Claims/encounter data • Vital stats • CDC WONDER
Desired Direction	Lower is better
Analytic Approach	• Pre-test/post-test • DiD • Synthetic Controls
Frequency	Annual

Non-fatal overdoses (all cause) (Measure 2-26)	
Numerator/Denominator	Numerator: Number of non-fatal overdoses Denominator: Number of Alaska residents
Comparison Population	N/A
Stratification	Gender, age, race/ethnicity, geography, criminal justice involvement, and high-cost high needs
Measure Steward	N/A
Measure Name	N/A
Data Source	• State eligibility and enrollment data • Claims/encounter data
Desired Direction	Lower is better
Analytic Approach	Pre-test/post-test
Frequency	Annual

Use of opioids at high dosage in persons without cancer (Measure 2-27)	
Numerator/Denominator	Numerator: Number of waiver recipients in the denominator who received prescriptions for opioids with an average daily dosage greater than or equal to 90 morphine milligram equivalents (MME) over a period of 90 days or more. Denominator: Number of waiver recipients aged 18 and older with two or more prescriptions for opioids on different days with a cumulative days' supply of 15 or more
Comparison Population	N/A
Stratification	Gender, age, race/ethnicity, geography, criminal justice involvement, and high-cost high needs
Measure Steward	HEDIS
Measure Name	OHD

[A-5] This measure will only be stratified if vital statistics data are available. Data from CDC Wonder cannot be stratified.

Use of opioids at high dosage in persons without cancer (Measure 2-27)	
Data Source	• State eligibility and enrollment data • Claims/encounter data
Desired Direction	Lower is better
Analytic Approach	Pre-test/post-test
Frequency	Annual

Use of opioids from multiple providers (Measure 2-28)	
Numerator/Denominator	Numerator: Number of waiver recipients receiving prescription opioids for => 15 days during the measurement year, who received opioids from multiple providers. Stratified by the following: – Multiple prescribers – Multiple pharmacies – Multiple prescribers and multiple pharmacies Denominator: Number of waiver recipients aged 18 and older with no more than one gap in continuous enrollment of up to 45 days and have two or more opioid dispensing events on different dates of service with => 15 total days covered by opioids
Comparison Population	N/A
Stratification	Gender, age, race/ethnicity, geography, criminal justice involvement, and high-cost high needs
Measure Steward	HEDIS
Measure Name	UOP
Data Source	• State eligibility and enrollment data • Claims/encounter data
Desired Direction	Lower is better
Analytic Approach	Pre-test/post-test
Frequency	Annual

Risk of continued opioid use (Measure 2-29)	
Numerator/Denominator	Numerator: Number of waiver recipients who have a new episode of opioid use that puts them at risk for continued opioid use. Stratified by the following: – At least 15 days of prescription opioids in a 30-day period – At least 31 days of prescription opioids in a 62-day period Denominator: Number of waiver recipients aged 18 and older with no more than one gap in continuous enrollment of up to 45 days who were enrolled at least 180 days prior to their index prescription start date (IPSD) and 61 days after the IPSD
Comparison Population	N/A
Stratification	Gender, age, race/ethnicity, geography, criminal justice involvement, and high-cost high needs
Measure Steward	HEDIS
Measure Name	COU
Data Source	• State eligibility and enrollment data • Claims/encounter data
Desired Direction	Lower is better

Risk of continued opioid use (Measure 2-29)	
Analytic Approach	Pre-test/post-test
Frequency	Annual

Hypothesis 3: Behavioral Health Reform will maintain or reduce the cost of Medicaid for Alaska and the federal government.

Research Question 3.1: Does Behavioral Health Reform maintain or reduce Alaska's per capita Medicaid BH costs?

Total costs of healthcare (sum of parts below), by State and federal share (Measure 3-1)	
Numerator/Denominator	Numerator: Total costs of healthcare. Stratified by the following: Total costs, IP, OP (ED OP and non-ED OP), LTC, Professional, Dental, Pharmacy Denominator: Total number of beneficiary months, divided by 1,000
Comparison Population	N/A
Stratification	Gender, age, race/ethnicity, geography, criminal justice involvement, and high-cost high needs
Measure Steward	N/A
Measure Name	N/A
Data Source	• State eligibility and enrollment data • Claims/encounter data
Desired Direction	No change or lower is better
Analytic Approach	ITS
Frequency	Annually/monthly

Total cost of SUD, SUD-IMD and SUD-Other and Non-SUD, by setting, including claims data (IP, OP, Rx, LTC) (Measure 3-2)	
Numerator/Denominator	Numerator: Total cost of SUD services. Stratified by: SUD-IMD, SUD-Other, Non-SUD Denominator: Total number of beneficiary months, divided by 1,000
Comparison Population	N/A
Stratification	Gender, age, race/ethnicity, geography, criminal justice involvement, and high-cost high needs
Measure Steward	N/A
Measure Name	N/A
Data Source	• State eligibility and enrollment data • Claims/encounter data
Desired Direction	No change or lower is better
Analytic Approach	ITS
Frequency	Annually/monthly

Total cost of MH diagnosis by IMD and Other, by setting, including claims data (IP, OP, Rx, LTC) (Measure 3-3)	
Numerator/Denominator	Numerator: Total cost of BH services. Stratified by: MH-IMD, MH- Other, Non-MH Denominator: Total number of beneficiary months, divided by 1,000
Comparison Population	N/A
Stratification	Gender, age, race/ethnicity, geography, criminal justice involvement, and high-cost high needs
Measure Steward	N/A
Measure Name	N/A
Data Source	• State eligibility and enrollment data • Claims/encounter data
Desired Direction	No change or lower is better
Analytic Approach	ITS
Frequency	Annually/monthly