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

State Demonstrations Group

July 28, 2023

Jacey Cooper
State Medicaid Director
Chief Deputy Director, Health Care Programs
California Department of Health Care Services
1501 Capitol Avenue, 6th Floor, MS 0000
Sacramento, CA 95814

Dear Director Cooper:

The Centers for Medicare & Medicaid Services (CMS) completed its review of the Drug Medi-Cal Organized Delivery System (DMC-ODS) Evaluation Design[1], which is required by the Special Terms and Conditions (STCs), specifically, STC #15.3, of the section 1115 demonstration, "California Advancing and Innovating Medi-Cal (CalAIM) Demonstration" (Project No: 11-W-00193/9), effective through December 31, 2026. CMS has determined that the Evaluation Design, which was submitted on June 27, 2022 and June 26, 2023, meets the requirements set forth in the STCs and our evaluation design guidance, and therefore approves the state's DMC-ODS Evaluation Design.

CMS has added the approved DMC-ODS Evaluation Design to the demonstration's STCs as Attachment T. A copy of the STCs, which includes the new attachment, is enclosed with this letter. In accordance with 42 CFR 431.424, the approved Evaluation Design may now be posted to the state's Medicaid website within 30 days. CMS will also post the approved Evaluation Design as a standalone document, separate from the STCs, on Medicaid.gov.

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

[1] The DMC-ODS Evaluation Design is inclusive of the state's substance use disorder and contingency management programs.

We appreciate our continued partnership with California on the CalAIM section 1115 demonstration. If you have any questions, please contact your CMS demonstration team.

Sincerely,

Paula M. Kazi -S Digitally signed by Paula M. Kazi -S
Date: 2023.07.28 00:13:42 -04'00'

Paula M. Kazi
Acting Director
Division of Demonstration Monitoring and Evaluation

cc: Cheryl Young, State Monitoring Lead, CMS Medicaid and CHIP Operations Group

California Drug Medi-Cal Organized Delivery System

Evaluation Design

Part of California Advancing and Innovating Medi-Cal (CalAIM)

A Section 1115 Demonstration Waiver Evaluation

and

Including the Medi-Cal 2020 Calendar Year 2021 Temporary Extension

Revised 6/9/2023

UCLA David Geffen School of Medicine
Integrated Substance Abuse Programs

Table of Contents

General Background Information

The Drug Medi-Cal Organized Delivery System (DMC-ODS) 1115 demonstration waiver was created by the California Department of Health Care Services (DHCS) with the intent of addressing many previously existing limitations in the DMC system. Prior to the DMC-ODS, the system was comprised of fragmented services, creating gaps that undermined client access and quality of care. The continuum of substance use disorder (SUD) services was uncoordinated, making it difficult for clients to navigate the system. SUD treatment providers indicated that many important services they provided or wished to provide for clients were not billable, were only reimbursable if delivered by a limited number of provider types or were too limited to provide proper care to clients. Providers were not necessarily required to deliver evidence-based practices in line with current research, and counties lacked the authority to fully ensure the quality and accountability of their local providers.

The DMC-ODS was created to test the impact of organizing SUD services to improve service delivery to Medicaid-eligible individuals with SUDs. The intent was to demonstrate that organized SUD care improves quality, access, and coordination/integration of treatment for beneficiaries while decreasing other health care system costs. Under the DMC-ODS waiver, care is organized according to the American Society of Addiction Medicine (ASAM) Criteria for SUD services. The ASAM Criteria are a set of guidelines developed by ASAM to set a standard for appropriate assessment, placement, and treatment planning of clients with SUD and co-occurring disorders as well as to a set standard for SUD providers. Services under the DMC-ODS waiver also create a continuum of care and create requirements allowing for local control, accountability, and greater administrative oversight.

The DMC-ODS waiver was originally approved by CMS in August 2015, and later became part of California's larger Medi-Cal 2020 Waiver, which ended December 31, 2021. It is now part of California Advancing and Innovating Medi-Cal (CalAIM), which is being implemented through a combination of 1115 and 1915b waivers starting January 1, 2022 and continuing through December 31, 2026. Most DMC-ODS services are now covered in the California Medicaid State Plan. This evaluation covers DMC-ODS under CalAIM as an extension of DMC-ODS under Medi-Cal 2020, including an evaluation of the Medi-Cal 2020 calendar year 2021 temporary extension, and will continue to evaluate the impact of DMC-ODS since its inception.

The population targeted by DMC-ODS is Medicaid-eligible individuals with SUDs. As described in the DMC-ODS waiver's Special Terms and Conditions (STCs),[1] for counties that opt-in to the DMC-ODS waiver, beneficiaries must meet the medical necessity criteria and reside in a participating county to receive waiver services. Currently, the DMC-ODS waiver is implemented

[1] https://www.dhcs.ca.gov/provgovpart/Documents/CalAIM-1115-Approval-Letter-and-STCs.pdf

in 37 counties that cover 95.9% of the state's population.[2] It is anticipated that currently non-participating counties will be given the option to opt-in to DMC-ODS during the CalAIM demonstration. If they do, they will also become part of the DMC-ODS population for evaluation purposes.

To address rapidly rising stimulant overdoses, the DMC-ODS will also cover Contingency Management (CM) under a new pilot program known as the Recovery Incentives Program: California's Contingency Management Benefit. This program began implementation in March 2023. Stimulant-related overdose death rates in California are 7.2 times higher today than they were 10 years ago, putting stimulants approximately on par with opioids in terms of total overdose-related deaths[3]. Methamphetamine use is also associated with hypertension, myocardial infarction, stroke, aortic dissection, and heart failure (Manja et al., 2023). Currently, no Food and Drug Administration-approved medications exist for the treatment of Stimulant Use Disorders (StimUD), but studies have repeatedly supported the use of CM as a highly effective evidence-based practice in the treatment of StimUD, particularly in reducing drug use (De Crescenzo et al., 2018; Farrell et al., 2019; AshaRani et al., 2020; Brown & DeFulio, 2020; Ronsley et al., 2020). County participation in the Recovery Incentives Program is optional.

The previous DMC-ODS evaluation plan was approved by CMS on June 20, 2016. The resulting evaluation documented DMC-ODS implementation and found that the DMC-ODS waiver has improved access to treatment, treatment quality, and coordination of care, and met the initial goals of the DMC-ODS (Urada et al., 2016, 2017, 2018, 2019, 2021, 2022).[4] Health disparities were identified in treatment placement, however. Under the new waiver, aside from the addition of the Recovery Incentives Program, DMC-ODS remains mostly intact with the addition of changes to clarify or streamline billing, benefit rules, and facilitate Health IT. The current evaluation design will look for any effect of the new changes but is otherwise focused on monitoring maintenance of the measured improvements found during the initial waiver, identifying emerging trends, determining opportunities to facilitate further progress, evaluating health equity, and evaluating the new Recovery Incentives Program.

One component of the waiver is still under review: DMC-ODS services that are provided by Traditional Healers and Natural Helpers. If this benefit is approved in the future, the evaluation team will bring methods already being employed in other parts of the evaluation (e.g., analysis of claims, provider interviews, and client perception surveys described below) to bear on these services.

---

[2] Projections Prepared by Demographic Research Unit, California Department of Finance, January 2021: https://www.dof.ca.gov/Forecasting/Demographics/Estimates/e-4/2010-21/documents/E-4_2021InternetVersion.xlsx

[3] Based on 12-month rolling averages from Q2 2021 and Q2 2011 data from: https://skylab.cdph.ca.gov/ODdash/. Total overdose deaths based on combination of psychostimulant and cocaine-related deaths.

[4] Due to data availability the 2022 report covered partial data for CY 2021. Analyses of 2021 data will be incorporated into the current evaluation as described in the methodology section.

If substantial external contextual issues arise in the future, the evaluation team will also measure and discuss these impacts, as the team has in the past with COVID-19 (Bass et al., 2022). Examples of potential contextual changes might include a waning (or increasing) impact of COVID-19, changes in the availability of fentanyl and high-potency stimulants, workforce shortages, increasing use of peers, and an expected IRS ruling that could have an impact on the size and total amount of incentives available to beneficiaries.

Evaluation Questions and Hypotheses

The evaluation will examine whether the DMC-ODS continues to achieve the following six goals as required by STC 46, an additional seventh goal on health disparities in the pursuit of CalAIM's goal of improving health equity, and an eighth goal based on STC 57e requirements specific to a contingency management evaluation. The Recovery Incentives Program also shares some overlapping goals with the rest of DMC-ODS (e.g., increased adherence and retention in treatment, reduced overdose deaths).

1. Increased rates of identification, initiation, and engagement in treatment;
2. Increased adherence to and retention in treatment;
3. Reductions in overdose deaths, particularly those due to opioids;
4. Reduced utilization of emergency departments and inpatient hospital settings for treatment where the utilization is preventable or medically inappropriate through improved access to other continuum of care services;
5. Fewer readmissions to the same or higher level of care where the readmission is preventable or medically inappropriate;
6. Improved access to care for physical health conditions among beneficiaries.
7. Improved health equity across DMC-ODS performance and outcome measures.
8. An effective contingency management program, including cost-effectiveness and effects on beneficiary health outcomes.

UCLA will also coordinate with DHCS to leverage the monitoring metrics[5] that DHCS is reporting to CMS to incorporate these metrics into the evaluation UCLA will conduct more in-depth analyses and additional quantitative and qualitative data collection to provide important context, insights, and recommendations beyond these metrics.

A short summary of the approaches for each of these goals follows. Additional details on the measures can be found in Table 1.

[5] https://www.medicaid.gov/medicaid/section-1115-demo/downloads/evaluation-reports/sud-monitoring-metrics.pdf

Increased rates of identification, initiation, and engagement in treatment

UCLA will calculate identification using a combination of data from ASAM level of care screenings and assessments, Managed Care Plan / Fee-for-Service (MCP/FFS), and Drug Medi-Cal claims. Separately, DHCS will report on related metrics: Metric 1 - Assessed for SUD Treatment Needs Using a Standardized Screening Tool, Metric 2 - Medicaid Beneficiaries with Newly Initiated SUD Treatment/Diagnosis, Metric 3 - Medicaid Beneficiaries with SUD Diagnosis (monthly) and Metric 4 - Medicaid Beneficiaries with SUD Diagnosis (annually). DHCS will report the initiation and engagement monitoring metric as required (Metric 15 – Initiation and Engagement of Alcohol and Other Drug Abuse or Dependence Treatment). However, there are data quality limitations to the initiation rate due to low rates of SUD diagnosis coding in the MCP/FFS delivery system.

Therefore, UCLA will enhance DHCS' and CMS' understanding of true initiation rates in the DMC-ODS evaluation by conducting more in-depth analyses that take these limitations into consideration. For example, UCLA can calculate initiation among DMC beneficiaries who were referred after an ASAM brief screening to assess the effectiveness of the DMC referral process. Separately, UCLA can calculate clients initiating DMC-ODS treatment after identification in physical health settings (merging DMC claims and MCP/FFS data) as a measure of coordination between the two systems. Trends in referrals to SUD treatment from health care sources will also continue to be monitored, and data on medications prescribed outside of specialty care settings will be reported for context.

Engagement rates as defined by NCQA can be accurately computed using claims data. Engagement has generally been steady over time among DMC-ODS clients (Padwa et al., 2022).

Earlier evaluation reports described increasing admissions and high levels of engagement in the DMC-ODS. About 23% of beneficiaries who had an ASAM-based brief screening received their indicated level of care within 30 days, leaving room for improvement. However, about 88% of clients who started treatment went on to engage in it by attending at least two more sessions (Padwa et al., 2022). Challenges to increase access included a shortage of qualified medical directors, licensed practitioners of the healing arts, bilingual staff, as well as difficulties in expanding medical withdrawal management, youth treatment, and understanding how to take advantage of the recovery services benefit. Penetration rates were likely limited in part due to the national phenomenon that 97.5% of people who need treatment usually do not recognize that need, and a smaller percentage do not seek treatment despite recognizing the need (SAMHSA, 2021). UCLA has recommended increasing outreach and screening in primary care and other non-specialty care settings as a result (Urada et al., 2022; Bass et al., 2022). SUD treatment referrals from health care sources have been flat, however, since the pre-DMC-ODS period (Lee et al., 2022). This may be in part due to increasing buprenorphine prescribing in primary care settings (Darfler et al., 2020). UCLA will continue to monitor these trends and conduct

stakeholder surveys and interviews to further investigate and recommend the best ways to close gaps in the number of people needing and receiving treatment.

Increased adherence to and retention in treatment

UCLA will analyze DMC-ODS claims to calculate length of stay, produce more in-depth analyses (e.g., by county, primary drug, race/ethnicity), and generate recommendations based on these results. The goal of improving overall retention may be complicated by the goal of reducing the statewide average residential length of stay to 30 days (STC 46). However, UCLA has provided recommendations to reduce the residential average length of stay without compromising quality (Urada et al., 2022, p. 108) and will continue monitoring trends. UCLA will also continue to track transitions in care. A slight increase in the rate of residential transitions to outpatient within 14 days was found among DMC-ODS counties, increasing from 7.1% in 2016 to 9.5% in 2020, as rates fell from 7.0% to 2.8% in state plan counties. Transitions from residential withdrawal management to residential treatment rose slightly from 17.0% to 20.2% in DMC-ODS counties from 2016 to 2020, while they rose from 3.2% to 8.0% in state plan counties (Lee et al., 2022).

Reductions in overdose deaths, particularly those due to opioids

While DHCS will be calculating and reporting required monitoring metrics[6] for this topic, specifically metric 26 – Overdose Deaths (count) and 27 – Overdose Deaths (rate), based on data provided to DHCS from CDPH, it will be important to place these metrics in context and control for them to the extent possible. In recent years, overdoses have risen in California despite the DMC-ODS due to external factors such as increasing availability of fentanyl and high-potency stimulants and the onset of the COVID-19 pandemic. In the future, potential emergence of new substances may further affect overdoses. For example, if xylazine use emerges in California, reducing the effectiveness of naloxone,[7] this could increase overdoses and have important policy ramifications for the use of this important tool. To the extent possible, UCLA will collaborate with DHCS and CDPH to examine the effect of treatment on overdose deaths and conduct in-depth analyses, e.g., by county, primary drug, race/ethnicity, and collect supplementary data from stakeholder surveys and/or interviews to generate recommendations. In addition to opioid overdose deaths, stimulant overdose deaths will be a particular focus of the evaluation of the Recovery Incentives Program.

[6] https://www.medicaid.gov/medicaid/section-1115-demo/downloads/evaluation-reports/sud-monitoring-metrics.pdf
[7] https://www.sciencedirect.com/science/article/pii/S037687162200117X

Reduced utilization of emergency departments and inpatient hospital settings for treatment where the utilization is preventable or medically inappropriate through improved access to other continuum of care services

While DHCS will examine this among all Medi-Cal beneficiaries (metric 23 – Emergency Department Utilization for SUD and metric 24 – Inpatient Stays for SUD), similar to overdose deaths, it is likely that DMC-ODS effects may be overwhelmed by external trends. In both cases, difference in difference analyses will be employed where possible to separate the DMC-ODS effect (see analytic methods below). A decrease in recurring overdoses were observed for a subset of counties following residential treatment under DMC-ODS compared to pre-waiver period and State Plan counties (Khurana et al., 2022, p. 115-117). UCLA will continue to analyze data among people who received SUD treatment under the DMC-ODS to determine whether utilization of emergency departments and inpatient hospital settings decreased relative to the pre-waiver period and will conduct cost analyses to determine whether savings (if any) in these settings offset increased SUD treatment expenses.

Fewer readmissions to the same or higher level of care where the readmission is preventable or medically inappropriate

UCLA will continue to use a measure adapted to DMC-ODS settings by focusing on readmissions to withdrawal management within 30 and 90 days of discharge. In 2020, the UCLA evaluation recently found that 17.5% of withdrawal management clients were readmitted within 90 days, down from 20% in 2019 (Padwa et al., 2022, p 64). UCLA will also describe residential readmissions with the understanding that not all readmissions are negative outcomes and examine whether transitions to outpatient treatment reduce residential readmissions.

Improved access to care for physical health conditions among beneficiaries

UCLA will examine improved access to physical health care among clients who participate in DMC-ODS treatment using annual client-reported ratings and administrative data. In 2020, 86% of clients agreed with the UCLA's Treatment Perceptions Survey item: "Staff here work with my physical health care providers to support my wellness," (Padwa et al., p. 66). UCLA will also analyze Medi-Cal MCP/FFS billing data to quantify increases in physical health care following admission to treatment. DHCS will also report metric 32 – Access to Preventative Ambulatory Health Services for Adult Medicaid Beneficiaries with SUD, which identifies the percentage of beneficiaries with ambulatory or preventative care visits.

Additional details on how each measure will be collected and how the hypotheses will be tested are included in the methodology section that follows. A driver diagram for the evaluation can be found in Appendix A.

Improved Health Equity

Past analyses have found that DMC-ODS treatment admissions did trend higher among all racial/ethnic groups (Bass et al., 2022) after DMC-ODS implementation. However, disparities in timely linkage to care have been detected for youth, older adult, Black, and Hispanic Medi-Cal enrollees (Padwa et al., 2022). Also, once admitted, treatment engagement increased among younger clients but decreased among older ones. DHCS plans to use quality improvement efforts via the External Quality Review Organization, for example, to reduce or eliminate disparities. UCLA will also continue to closely examine trends in health equity within each measure included in the six goals previously described above, track these findings over time, investigate causes of any disparities found (e.g., through interviews and surveys), summarize findings, and generate recommendations. At a minimum, groups of interest will include race, ethnicity, age, gender, and location. UCLA and DHCS are examining the feasibility of adding other groups including sexual orientation based on the data availability.

CMS is currently reviewing the addition of Traditional Healers and Natural Helpers to the DMC-ODS. If approved, UCLA will also evaluate the impact of this change, particularly on the American Indian/Alaska Native (AI/AN) population.

An effective contingency management program, including cost-effectiveness and effects on beneficiary health outcomes.

Due to the large number of studies and systematic reviews that have established the efficacy of CM, the primary goal of the Recovery Incentives Program evaluation is not to conduct research aimed at further re-establishing effectiveness, but rather to evaluate the effectiveness of real-world implementation in the California Recovery Incentives Program, document efforts to scale this proven treatment in a large state, and to facilitate quality improvement. A range of hypotheses will be tested as shown in Table 1.

Consistent with STC 57e, to the extent feasible, the state will conduct evaluation analyses stratified by StimUD and other types of SUD. However, the Recovery Incentives Program is currently aimed exclusively at beneficiaries who have StimUD.

Since the Recovery Incentives Program is part of DMC-ODS, the overall DMC-ODS evaluation and all analyses are inclusive of the participating Recovery Incentives Program treatment sites and clients. However, more in-depth data collection and analysis will be specifically applied to the Recovery Incentives Program, including efforts to measure the effects of this program above and beyond that of DMC-ODS, e.g., comparing Recovery Incentives Program StimUD clients to non- Recovery Incentives Program StimUD clients in DMC-ODS.

Summary of key evaluation questions, hypotheses, data sources, and analytic approaches

Table 1 below summarizes the questions, hypotheses, and measures to be used in this study. As previously noted, UCLA will also coordinate with DHCS to incorporate established monitoring metrics[8] that DHCS is reporting to CMS separately. The measures below are meant to supplement DHCS-reported measures to answer remaining DMC-ODS-related questions, often by using data specific to California and DMC-ODS (e.g., California's ASAM LOC Placement data, Incentive Manager data, UCLA-administered surveys).

[8] https://www.medicaid.gov/medicaid/section-1115-demo/downloads/evaluation-reports/sud-monitoring-metrics.pdf

Table 1. Summary of key evaluation questions, hypotheses, data sources, and analytic approaches

Driver	Potential Measures	Measure Steward, Endorsement	Numerator	Denominator	Data Source(s)	Analytic approach
Question: Are rates of overdose deaths impacted by the demonstration? Goal: Reduction in overdose deaths, particularly those due to opioids. Hypothesis: People with opioid use disorders (OUD) who receive MAT and people with StimUD who participate in the Recovery Incentives Program will be less likely to have an overdose death compared to people with OUD and StimUD who do not receive these services, respectively.						
Primary Driver: Reduce overdose deaths	Overdose deaths overall and among opioids and stimulants separately	None	N/A	N/A	California Comprehensive Death File, CA Department of Public Health matched to DMC Claims	Compare individuals with StimUD who participated in the Recovery Incentives Program to those who did not Time period: Start of Recovery Incentives Program (2023) through end of waiver (2026) contingent on data availability Compare individuals with OUD who received MAT to those who did not and determine whether access to MAT increased under DMC-ODS (2015-2026, contingent on data availability) Quasi-experimental causal inference designs including but not limited to difference-in-differences augmented with propensity score matching as needed, synthetic controls, etc.

Table 1. Summary of key evaluation questions, hypotheses, data sources, and analytic approaches						
Driver	**Potential Measures**	**Measure Steward, Endorsement**	**Numerator**	**Denominator**	**Data Source(s)**	**Analytic approach**
Question: Does the demonstration increase access to and utilization of SUD treatment services? Goal: Increased rates of identification, initiation, and engagement in SUD treatment services. Hypothesis: Counts or rates will be maintained at benchmark year* levels or higher.						
Primary Driver: increased rates of identification, initiation, and engagement in treatment	Number of ASAM level of care screenings and assessments	None	Number of ASAM LOC screenings and assessments	N/A	ASAM LOC Placement data	Descriptive statistics using parametric and/or non-parametric tests of statistical significance and/or regression analysis to confirm identification, IET rates, and timely admission to the indicated level of care are maintained or improve between comparison & waiver periods (2020-2026, contingent on data availability)
	Initiation among beneficiaries with an ASAM brief screening	NQF #0004 adaptation	Number of beneficiaries who initiated treatment within 14 days of the index episode start date	Number of beneficiaries with an ASAM brief screening with a level of care recommendation	DMC Claims, ASAM LOC Placement data	
	Engagement in treatment among DMC-ODS clients	NQF #0004	Initiation of tx and two or more encounters with any SUD diagnosis within 30 days after initiation	Number of beneficiaries (above) who initiated treatment	DMC Claims, ASAM LOC Placement data	
Secondary Driver: Ensure appropriate and timely placement according to ASAM criteria	Timely admission to the indicated level of care within 30 days of ASAM Criteria-based brief screenings	None	Admission within 30 days of an ASAM Criteria-based brief screening	Beneficiaries with an ASAM brief screening with a level of care recommendation	DMC Claims, ASAM LOC Placement data	Descriptive Statistics (2020-2026, contingent on data availability)

Table 1. Summary of key evaluation questions, hypotheses, data sources, and analytic approaches						
Driver	**Potential Measures**	**Measure Steward, Endorsement**	**Numerator**	**Denominator**	**Data Source(s)**	**Analytic approach**
Secondary Driver: Ensure clients are satisfied with services	UCLA Client Treatment Perceptions Survey ratings, % of clients providing a 4 or higher rating on all questions	UCLA	Clients providing a 4 or 5 rating	All TPS participants	UCLA Client Treatment Perceptions Survey	Descriptive statistics (2020-2025)
Secondary Driver: Quality improvement efforts	UCLA County administrator survey questions on the impact of QI activities and the EQRO	None	N/A	N/A	County administrator survey	Descriptive statistics (2020-2026)
Question: Do enrollees receiving SUD services adhere to and remain in treatment? Goal: Increased adherence to and retention in treatment. Hypothesis: Adherence and retention will be maintained at benchmark year* levels or higher.						
Primary Driver: Adherence to and retention in treatment	Days in treatment	None	N/A	N/A	DMC Claims CalOMS-Tx	Descriptive statistics using parametric and/or non-parametric tests of statistical significance, and/or regression analysis and quasi-experimental causal inference designs including but not limited to difference-in-differences augmented with propensity score matching as needed, synthetic controls, etc. (2020-2026, contingent on data availability)

Table 1. Summary of key evaluation questions, hypotheses, data sources, and analytic approaches						
Driver	**Potential Measures**	**Measure Steward, Endorsement**	**Numerator**	**Denominator**	**Data Source(s)**	**Analytic approach**
Secondary Driver: Improve care coordination and transitions between levels of care	Transition to specialty care after withdrawal management	None	Withdrawal management discharges followed by DMC-ODS treatment within 7 or 14 days	WM discharges	DMC claims, MCP/FFS data**	Descriptive statistics using parametric and/or non-parametric tests of statistical significance, and/or regression analysis (2020-2026, contingent on data availability)
Question: Do enrollees receiving SUD services experience improved health outcomes? Goal: Reduced utilization of emergency departments and inpatient hospital settings for treatment where the utilization is preventable or medically inappropriate through improved access to other continuum of care services. Hypothesis: DMC-ODS implementation will be associated with reductions in utilization of emergency departments and inpatient hospital settings for treatment where the utilization is preventable or medically inappropriate through improved access to other continuum of care services.						
Primary driver: Reduced utilization of ED and inpatient hospital settings	Utilization (e.g., days)	None	Clients who received DMC-ODS treatment who had any ED and inpatient hospital visits during and after treatment	Clients who receive DMC-ODS treatment	DMC Claims MCP/FFS data**	Descriptive statistics using parametric and/or non-parametric tests of statistical significance, and/or regression analysis. Quasi-experimental causal inference designs including but not limited to difference-in-differences augmented with propensity score matching as needed, synthetic controls, etc. (2015-2026, contingent on data availability)
	Paid claim amounts	None	N/A	N/A	DMC Claims MCP/FFS data**	

Table 1. Summary of key evaluation questions, hypotheses, data sources, and analytic approaches						
Driver	**Potential Measures**	**Measure Steward, Endorsement**	**Numerator**	**Denominator**	**Data Source(s)**	**Analytic approach**
Question: Does the demonstration reduce withdrawal management readmissions? Goal: Fewer readmissions to the same or higher level of care where the readmission is preventable or medically inappropriate. Hypothesis: DMC-ODS implementation will be associated with fewer readmissions to the same or higher level of care where the readmission is preventable or medically inappropriate.						
Primary driver: Readmissions to withdrawal management	Re-admissions within 30 days of discharge	None	Clients re-admitted to withdrawal management within 30 days of discharge from withdrawal management	Clients discharged from withdrawal management	DMC Claims CalOMS-Tx	Descriptive statistics using parametric and/or non-parametric tests of statistical significance, and/or regression analysis. Quasi-experimental causal inference designs including but not limited to difference-in-differences augmented with propensity score matching as needed, synthetic controls, etc. (2015-2026, contingent on data availability)

Table 1. Summary of key evaluation questions, hypotheses, data sources, and analytic approaches						
Driver	**Potential Measures**	**Measure Steward, Endorsement**	**Numerator**	**Denominator**	**Data Source(s)**	**Analytic approach**
Question: Does the demonstration improve coordination of care? Goal: Improved access to care for physical health conditions among beneficiaries. Hypothesis: DMC-ODS implementation will be associated with improved access to care for physical health conditions among beneficiaries.						
Primary driver: Ensure client satisfaction with services	Treatment Perceptions Survey item: "Staff here work with my physical health care providers to support my wellness."	None	Clients providing a rating of 4 or 5	All clients responding to the TPS survey	Treatment Perceptions Survey	Confirm client satisfaction w/ coordination is at benchmark year* levels/higher. Descriptive statistics using parametric and/or non-parametric tests of statistical significance, and/or regression analysis (2020-2025)
Secondary driver: Improve care coordination	Percentage of clients with ambulatory or preventive care visits before and following treatment	NCQA adaptation	Number of clients with SUD who had an ambulatory or preventive care visit during the measurement period	Number of beneficiaries with DMC-ODS treatment	MCP/FFS data** DMC claims	Compare ambulatory or preventative care visits before & after treatment. Descriptive statistics using parametric and/or non-parametric tests of statistical significance, and/or regression analysis (2015-2026, contingent on data availability)

Table 1. Summary of key evaluation questions, hypotheses, data sources, and analytic approaches						
Driver	**Potential Measures**	**Measure Steward, Endorsement**	**Numerator**	**Denominator**	**Data Source(s)**	**Analytic approach**
Question: Does the demonstration reduce health disparities? Goal: Improved health equity Hypothesis: Health disparities will decrease.						
Primary Driver: Improve health equity	Timely admission to indicated level of care	None	Clients admitted to their indicated level of care within 30 days of ASAM brief screening	Clients who received an ASAM brief screening	ASAM LOC Placement data DMC Claims	Compare rates by race, ethnicity, and age. Descriptive statistics using parametric and/or non-parametric tests of statistical significance, and/or regression analysis. Quasi-experimental causal inference designs including but not limited to difference-in-differences augmented with propensity score matching as needed, synthetic controls, etc. (2017-2026)
	Treatment engagement	NQF #0004	Initiation of treatment and two or more encounters with any SUD diagnosis within 30 days after initiation	Number of beneficiaries who initiated treatment	DMC Claims	
	Any other measures on which meaningful disparities emerge					

Table 1. Summary of key evaluation questions, hypotheses, data sources, and analytic approaches						
Driver	**Potential Measures**	**Measure Steward, Endorsement**	**Numerator**	**Denominator**	**Data Source(s)**	**Analytic approach**
Question: Has the Recovery Incentives Program been effectively implemented? Goal: An effective contingency management program, including cost-effectiveness and effects on beneficiary health outcomes. Hypothesis: Effective implementation will lead to improvements in client retention, discharge status, self-reported outcomes, drug test results, deaths, and healthcare utilization among clients participating in the Recovery Incentives Program.						
Primary driver: Improvements in Recovery Incentives Program outcomes	Days in treatment, engagement, discharge status, self-reported satisfaction and improvement in health, SUD, arrests, ED and inpatient hospital utilization, costs, deaths	None	N/A	N/A	Client surveys, DMC claims, MCP/FFS data,** CalOMS-Tx, Death data	Compare outcomes between clients with StimUD participating in the Recovery Incentives Program and those in non-Recovery Incentives Program treatment programs (where available), controlling for background characteristics. Comparisons by demographics. Descriptive statistics using parametric and/or non-parametric tests of statistical significance, and/or regression analysis Compare rates of positive, negative, and missed drug screens among individuals with StimUD in the Recovery Incentives Program and compare rates to those found in the literature using a one-sample t-test or analogous procedure (2023-2026)
	Rates of positive, negative, and missed drug screens	None	Negative urinalysis outcomes	Sum of all possible tests over the planned course of treatment	Stimulant drug tests / incentive manager vendor	

Table 1. Summary of key evaluation questions, hypotheses, data sources, and analytic approaches

Driver	Potential Measures	Measure Steward, Endorsement	Numerator	Denominator	Data Source(s)	Analytic approach
Primary driver: Fidelity to the CM model	Drug screen results, Days in treatment, Discharge status, Self-reported improvement, Overdose rates, ED and inpatient hospital utilization (SUD or all diagnoses)	None	N/A	N/A	Data from incentive manager vendor, fidelity assessments, provider surveys, client surveys, CalOMS-Tx, DMC-ODS claims	Compare outcomes (e.g., drug screen results, days in treatment, discharge status, self-reported improvement, overdose rates, ED utilization, inpatient utilization) between higher- and lower-fidelity providers according to measures developed by UCLA Descriptive statistics using parametric and/or non-parametric tests of statistical significance, and/or regression analysis. (2023-2026)
Secondary driver: Implementation of an effective and accessible CM program	Newly developed survey questions adapted from an existing questionnaire and qualitative interviews	None	N/A	N/A	Provider surveys and interviews	Descriptive analyses from survey to track implementation challenges and successes over time and qualitative analyses of interview transcripts
	Use of CM based on DMC claims	None	Clients receiving CM	Clients with StimUD in eligible levels of care	DMC-ODS claims CalOMS-Tx	Track percentage of people in treatment for StimUD who participate in the Recovery Incentives Program; Descriptive statistics using parametric and/or non-parametric tests of statistical significance, and/or regression analysis. (2023-2026)

* Benchmark year is expected to be 2021 but may be adjusted if appropriate. The benchmark for evaluating 2021 will be 2020 or an alternative (see methodology section). Where pre-DMC-ODS data do not exist and maintenance is hypothesized, the starting year is 2020. Where pre-DMC-ODS data do exist, the starting year is set at 2015 to take advantage of this data. Analyses based on recovery incentives-specific data start in 2023 when collection of the relevant data begins.

** ED, hospital, and associated cost data come from MCP/FFS data is historically subject to reporting delays of about 3 years.

In addition to the hypothesis testing described above, the study team may describe emerging facilitators and barriers to DMC-ODS implementation, e.g., associated with implementation of peer support specialists, potential impacts from payment reform, and other emerging issues. For example, between 2015 and 2021 issues such as COVID-19, rising overdose deaths from fentanyl and stimulants, and increasing rates of homelessness were incorporated into DMC-ODS evaluation reports as special topics as these issues took on increased urgency.

To the extent possible, UCLA will also examine total costs as well as cost drivers measured on a Per Member Per Month (PMPM) basis before and during the demonstration periods (2015-2026 contingent on data availability), e.g., total Medicaid costs and total federal Medicaid costs, 2) SUD-IMD costs, other SUD costs and non-SUD costs, and 3) inpatient costs, non-ED outpatient costs, and ED outpatient costs.

Methodology

Evaluation Design Summary

The evaluation uses a mixed-methods design that takes advantage of different comparisons based on the measure in question.

Where appropriate, administrative data from Drug Medi-Cal (DMC) claims and CalOMS-Tx will be used for a difference-in-difference design to account for different county implementation periods, consistent with CMS recommendations for strong evaluation designs.[9] This approach essentially combines pre-post comparisons and comparisons across counties to test whether changes are detected when counties "go live" but not at the same time in other counties. In other cases, data (e.g., stakeholder surveys, interviews, ASAM Criteria-based Level of Care Placement data) will only be available post-implementation, in which case post-only analyses will be conducted.

Evaluation of the Recovery Incentives Program is focused on initial implementation of a specific set of new practices targeted at a specific set of clients in specific settings, in marked contrast to evaluation of the broader DMC-ODS program that has been in place for several years and affects the entire continuum of care. The evaluation approach for the Recovery Incentives Program therefore necessarily has a different focus, organized around the RE-AIM framework (Glasgow, 1999):

[9] Reschovsky, J.D. and Bradley, K. (2019). Planning Section 1115 Demonstration Implementation to Enable Strong Evaluation Designs. Available at: https://www.medicaid.gov/medicaid/section-1115-demo/downloads/evaluation-reports/enable-strng-eval-dsgn.pdf

1. **Reach**. This will be measured as the percentage of people in treatment for StimUD who participate in the Recovery Incentives Program. UCLA will also evaluate whether there are disparities in its reach to different beneficiary populations (e.g., race, ethnicity, gender, age, county).

2. **Effectiveness.** Effectiveness will be based on results of drug testing, treatment retention, and treatment engagement.

3. **Adoption.** Adoption will be measured by evaluating how many provider agencies deliver Recovery Incentives Program services.

4. **Implementation.** Implementation will be evaluated by the degree to which CM is implemented with fidelity to the Recovery Incentives Program protocols and by tracking adaptations made. Perceptions of challenges and areas for potential improvement will also be collected from provider staff and participants.

5. **Maintenance.** Maintenance will be measured by evaluating the degree to which programs implementing the Recovery Incentives Program continue providing the service throughout the evaluation period, and information from surveys and interviews focusing on factors that could promote or impede the continued delivery of Recovery Incentives Program services after the end of the pilot period.

Target and Comparison Populations

The population targeted by the DMC-ODS is Medicaid-eligible individuals with SUD. Where appropriate, state plan counties and variation in introduction of the DMC-ODS waiver across counties in California over time will be exploited for comparison purposes as described in the analytic methods section below.

In some cases, particularly when analyzing datasets that did not exist prior to DMC-ODS implementation, the evaluation design is focused on monitoring maintenance of previously measured improvements. In these cases, the waiver year 2021 is proposed as a benchmark year to measure maintenance of improvements as CalAIM extends DMC-ODS into 2022 and beyond. However, COVID-19 or other future trends may eventually make another year more appropriate. For example, the DMC-ODS evaluation previously found that COVID-19 reduced admissions (Bass et al., 2022), so if pandemic-driven trends dissipate in the future and DMC-ODS treatment admissions return to pre-pandemic levels, the pre-pandemic year 2019 could become a more appropriate comparison year to avoid confounding the effects of CalAIM with recovery from the pandemic. If a year other than 2021 is adopted as a baseline year, sensitivity analyses will be performed to quantify the effect of this change. To evaluate the year 2021, 2020 will be used as a comparison, with the understanding that COVID-19 may affect both years. Alternatives to 2020 and 2021 including average benchmarks based on the time series of data available for each outcome variable will also be explored.

In other cases, where improvements have not previously been established, data will be analyzed to establish whether the initial waiver was associated with or caused improvements, as well as whether those improvements have been maintained during the current CalAIM waiver.

As a result of the above considerations, time periods in Table 1 differ by measure according to the following rules: 1. Where maintenance is hypothesized, the starting year is 2020 (to provide a comparison for 2021), though 2021 may then serve as a benchmark for the ensuing years. 2. Where administrative data exist prior to DMC-ODS, the starting year is 2015 to provide two years or more (depending on county) of pre-DMC-ODS data to serve as a baseline. 3. Analyses based on data collected specifically for the Recovery Incentives Program starts in 2023 when data collection begins. Although aspirational 2026 end dates are listed, full 2026 data may not always be available for inclusion in the report due in December 2026. In some cases, e.g. county administrator surveys, this is under the evaluator's control and will be complete in 2026. In the case of administrative datasets, cutoff dates will be determined by data availability which may range from partial 2026 data to a much earlier cutoff in the case of MCP/FFS.

The primary target population for the Recovery Incentives Program evaluation will be clients who receive CM for the treatment of StimUD. The comparison population will consist of clients who receive treatment for StimUD but do not receive CM. Administrative data on this population will be available for the treatment programs participating in the Recovery Incentives Program in both the pre-and post-Recovery Incentives Program periods and will be available for other treatment programs that are not participating in the Recovery Incentives Program.

During the DMC-ODS waiver period, the IRS is expected to make a ruling on whether CM incentives are considered income. Should the IRS determine that it is not income, the current $599 annual cap on incentives provided to individuals would increase by amount to be determined. If this were to occur, in addition to the $599 Recovery Incentives Program group and non-Recovery Incentives Program comparison group, a third, higher-dose Recovery Incentives Program group would be created and evaluated separately from the $599 group but using the same methods.

Evaluation Period

DMC-ODS under CalAIM is considered an extension of DMC-ODS under the previous Medi-Cal 2020 waiver. Therefore, the evaluation period will extend from the date the first counties implemented DMC-ODS on February 1, 2017 through the end of the CalAIM waiver on December 31, 2026. However, exact dates will differ by analysis depending on data availability, normal data reporting lag times, and hypotheses. The first DMC-ODS report (mid-point assessment) will also include previously unreported analyses of 2021 data. The evaluation period for Recovery Incentives Program evaluation will have the same end date, but implementation began in March 2023.

Data Sources

Administrative data sources

California Outcome Measurement System, Treatment (CalOMS-Tx)

CalOMS-Tx is California's existing data collection and reporting system for all clients in publicly funded SUD treatment services. Treatment providers collect information from clients at admission and discharge and send this data to DHCS each month. CalOMS-Tx provides California's contribution to the Treatment Episode Dataset (TEDS) maintained by the Substance Abuse and Mental Health Services Administration (SAMHSA). CalOMS includes client background (e.g. demographics, source of referral, number of prior treatment episodes, housing, employment, criminal justice status, number of children), treatment information (e.g.treatment discharge status, use of medications), and 30-day measures at admission and discharge (e.g. number of arrests & jail days, family conflicts, social support). This makes CalOMS-Tx data richer in many respects than other data sources (e.g. claims), though it has its own limitations (see limitations section). More information on CalOMS-Tx can be found at: http://www.dhcs.ca.gov/provgovpart/Pages/CalOMS-Treatment.aspx

Death Data

The California Department of Public Health (CDPH) provides data from their California Comprehensive Death File to DHCS for all Medi-Cal beneficiaries. UCLA will collaborate with DHCS and CDPH to use this data to identify overdose deaths as a key outcome measure. All-cause deaths will also be examined if the data allows.

Drug Medi-Cal Claims (DMC Claims)

In California, Medicaid-funded SUD treatment is paid for through DMC claims. DMC is a carve-out for specialty care SUD treatment. For the UCLA evaluation, DMC claims data provides information on patient demographics, access to treatment after DMC-ODS waiver implementation, types of services provided, and costs. New billing procedures under development are expected to record the delivery of CM services and potentially positive or negative drug test results. DMC claims data provides detailed data on services received and is likely to be more complete than other datasets like CalOMS-Tx but is limited in scope to billing-related data.

Incentive Manager Vendor Data

The incentive manager vendor for the Recovery Incentives Program, under contract with DHCS, will collect data on incentive payments while administering these incentives. The following data elements are expected to be collected:

- Beneficiary name (recipient's full name: last, first, and middle initial)
- Beneficiary Client Identification Number (CIN) (recipient's unique identification number established by DHCS
- Provider name (billing and/or rendering provider name)
- National provider identifier (billing and/or rendering provider number)
- Date of service (date drug test was performed, incentive disbursed if test was negative for stimulants, excused or unexcused absence)
- Drug test results (positive or negative for stimulants)
- Calculated incentive amount on date of service (incentive amount owed to client)
- Disbursed incentive amount on date of service
- Cumulative disbursed incentive amounts, per client per calendar year (total incentive amounts disbursed to each beneficiary enrolled in the Recovery Incentives Program per calendar year)
- *Other data to be determined by DHCS*

Managed Care Plan/ Fee-for-Service Data (MCP/FFS)

In California, Medicaid-funded medical care (excluding SUD and serious mental illness) is paid for either through managed care plans or fee-for-service reimbursement. For the UCLA evaluation, MCP/FFS data provides information on client demographics, types of services, and costs.

Mental Health (MH) Claims

In California, Medicaid-funded MH treatment is paid for through Short Doyle Medi-Cal claims (SD/MC). SD/MC is a carve-out for serious mental illness treatment services to persons eligible for Medi-Cal. For the UCLA evaluation, SD/MC claims data provides information on the dates, types, and quantities of MH services provided for beneficiaries accessing services for SMI.

Medi-Cal Eligibility Data System (MEDS)

The MEDS database provides information on all California Medi-Cal beneficiaries. These data, particularly the MEDS Monthly Extract File (MMEF), are used to calculate penetration rates.

Master Provider File (MPF)

The MPF is DHCS's comprehensive list of SUD treatment programs in the state of California. The MPF includes information on all SUD treatment facilities, including mailing addresses and DMC certification and decertification dates, among other provider-level information. In combination with lists of IMD facilities, MPF can be used to identify provider identification numbers for these facilities, therefore enabling IMD-specific analyses using CalOMS and DMC Claims data.

UCLA evaluation data collection activities

ASAM Level of Care (LOC) Placement Data

Given that The ASAM Criteria are a defining feature of the DMC-ODS waiver, a large new data collection effort was initiated across DMC-ODS waiver counties to collect data on the use of ASAM Criteria-based LOC brief initial screenings, initial assessments, reassessments, and services delivered. This endeavor has been a collaborative effort between UCLA, DHCS, and counties to collect these data. DHCS Information Notice 17-035 describing the requirements and procedures to collect ASAM Criteria-based LOC data was released in September 2017 and was superseded by Information Notice 18-046 on October 1, 2018. These data include the date of screening or assessment, type (brief initial screen, initial assessment, follow-up assessment), indicated LOCs (per screener or assessment result), actual placement decision(s), the reason for the difference between indicated and actual LOCs (if any), and the reason for delays in placement (if any). Data on three types of screenings or assessments are possible, defined as follows on the data collection instrument.

- Brief Initial Screen: a brief initial screening that preliminarily determines an LOC placement until a full assessment can be performed
- Initial Assessment: a longer comprehensive assessment meant to determine the LOC recommendation and establish medical necessity
- Follow-up Assessment: following an initial assessment, any re-assessment of the client occurring during the same treatment episode

Up to three indicated and actual levels of care could be recorded. Indicated and actual levels of care defined as:

- Indicated LOC. This is the initially recommended LOC according to the screening/assessment instrument prior to taking client preference into account. For example, this would be listed under "Final Level of Care Recommendations" if using CONTINUUM™ software.
- Actual LOC/Withdrawal Management placement decision. This is the actual LOC decided upon after client input and the LOC where the client is referred.

The options for LOC, as worded in the LOC reporting template, are listed below. These include broad To Be Determined (TBD) options to allow for the results of brief initial screenings that may indicate a general treatment modality the client should report to for further assessment (e.g., outpatient) without specifying the exact LOC to be received there (e.g., 1-outpatient or 2.1-intensive outpatient). The list also includes Withdrawal Management (WM) levels of treatment, which can be combined with other levels of care.

Level of Care

None
Outpatient/Intensive Outpatient (OP/IOP), exact level TBD
Residential, exact level TBD
Withdrawal Management (WM), exact level TBD
Ambulatory WM, exact level TBD
Residential/Inpatient WM, exact level TBD
Narcotic Treatment program/Opiate Treatment program (NTP/OTP)
0.5 Early Intervention
1.0 OP
2.1 IOP
2.5 Partial Hospitalization
3.1 Clinically Managed Low-Intensity Residential
3.3 Clinically Managed Population-Specific High-Intensity Residential
3.5 Clinically Managed High-Intensity Residential Services
3.7 Medically Monitored Intensive Inpatient Services
4.0 Medically Managed Intensive Inpatient Services
1-WM Ambulatory WM without Extended Onsite Monitoring
2-WM Ambulatory WM with Extended Onsite Monitoring
3.2-WM Clinically Managed Residential WM
3.7-WM Medically Monitored Inpatient WM
4-WM Medically Managed Intensive Inpatient WM

If at least one of the indicated and actual levels of care do not match, providers are asked to select the reason for the difference. The options are:

Reason for difference

Not applicable - no difference
Clinical judgment
Lack of insurance/payment source
Legal issues
Level of care not available
Managed care refusal
Client preference
Geographic accessibility
Family responsibility
Language
Used two residential stays in a year already.
Other

County Administrator Surveys/Interviews

UCLA will continue to develop and distribute online surveys to obtain information and insights from county SUD/behavioral health administrators participating in the delivery of services under

the DMC-ODS system of care. Surveys will be conducted annually to address DMC-ODS-related perceptions, barriers, and facilitators. Past topics have included, for example, access to care, screening and placement practices, training needs, quality of care, coordination, and integration of services. Additional topics, including on the Recovery Incentives Program, will be included as driven by the evaluation measures and other new issues/external factors as they emerge. UCLA will also conduct in-depth interviews with stakeholders on an as-needed basis to further inform and understand the findings from the administrative and survey data. Surveys will continue to be administered online (e.g., Qualtrics), and will be sent to either all DMC-ODS counties (currently 37) or all counties (58 counties, 57 surveys because Yuba and Sutter counties share a single administrator). Historically nearly all county administrators have responded (most recently 36 out of 37, 97%), eliminating the need for stratification.

Treatment Perceptions Survey (TPS)

The TPS was developed by UCLA as part of the activities for the initial DMC-ODS waiver evaluation activities in 2017. The TPS for adults was based on San Francisco County's Treatment Satisfaction Survey; and the TPS for youth was based on Los Angeles County's Treatment Perceptions Survey (Youth). (Both survey questionnaires include items from the Mental Health Statistics Improvement Program, MHSIP.) Input on the survey development was solicited from and provided by DHCS, the Substance Abuse Prevention Treatment+ Committee (SAPT+) of the County Behavioral Health Director's Association (CBHDA) of California, the DMC-ODS External Quality Review Organization (EQRO) Clinical Committee, Behavioral Health Concepts (BHC), the Youth System of Care Evaluation Team at Azusa Pacific University, and other stakeholders. The tool has since been validated (Teruya et al, 2022) and data collection has occurred annually during a five-day survey period among counties participating in the DMC-ODS waiver since 2018. The TPS data serves multiple purposes. 1) it fulfills counties' EQRO requirement to conduct a client satisfaction survey at least annually using a validated tool, 2) it addresses the data collection needs for the CMS required evaluation of the DMC-ODS waiver, and 3) supports DMC-ODS quality improvement efforts and provides key information on the impacts of the DMC-ODS waiver.

The TPS is administered annually as part of a major statewide undertaking by UCLA, counties, and providers during a specified five-day survey period. Providers are directed to administer the survey to every client receiving services both in-person or via tele-health during this time. During the most recent (2021) data collection period, 16,628 surveys were collected. Among adults, the smallest two racial groups were Native Hawaiian/Pacific Islander (n=259) and Asian (n=410). These sample sizes are sufficient to detect a small effect size (D) of 0.16 with a two-sided alpha of .05 and power (beta) of .80 using an independent samples t-test. TPS response rates have historically been estimated at about 60% but dipped during 2020, during the COVID-19 pandemic. If needed, the sample can be weighted for survey nonresponse to match the demographic profile of clients with DMC claims during the survey period. However, since no large differences were found in ratings between demographic groups in 2021, such weighting may have a minimal impact.

The survey for adults includes 14 statements addressing client perceptions of access, quality, care coordination, outcome, and general satisfaction. The survey for youth includes 18 statements in the same five domains as the adult survey plus an additional domain: therapeutic alliance. Survey respondents indicate the extent to which they disagree or agree with statements using a 5-point Likert scale (1= Strongly disagree and 5= Strongly agree). The survey also collects demographic information (i.e., gender, age, race/ethnicity, and length of time receiving services at the treatment program).

TPS Adult Survey Items by Domain

Access

1. The location was convenient (public transportation, distance, parking, etc.).
2. Services were available when I needed them.

Quality

3. I chose the treatment goals with my provider's help.
4. Staff gave me enough time in my treatment sessions.
5. Staff treated me with respect.
6. Staff spoke to me in a way I understood.
7. Staff were sensitive to my cultural background (race, religion, language, etc.).

Care Coordination

8. Staff here work with my PH care providers to support my wellness.
9. Staff here work with my MH care providers to support my wellness.

Outcome

10. As a direct result of the services I am receiving, I am better able to do things that I want to do.

General Satisfaction

11. I felt welcomed here.
12. Overall, I am satisfied with the services I received.
13. I was able to get all the help/services that I needed.
14. I would recommend this agency to a friend or family member

TPS Youth Survey Items by Domain

Access

1. The location of services was convenient for me.
2. Services were available at times that were convenient for me.
3. I had a good experience enrolling in treatment.

Therapeutic Alliance

4. My counselor and I work on treatment goals together.
5. I feel my counselor took the time to listen to what I had to say.
6. I developed a positive, trusting relationship with my counselor.
7. I feel my counselor was sincerely interested in me and understood me.
8. I like my counselor here.
9. My counselor is capable of helping me.

Quality

10. I received the right services.
11. Staff treated me with respect.
12. Staff were sensitive to my cultural background (race/ethnicity, religion, language, etc.).
13. My counselor provided necessary services for my family.

Care Coordination

14. Staff here make sure that my health and emotional health needs are being met (physical exams, depressed mood, etc.).
15. Staff here helped me with other issues and concerns I had related to legal/probation, family, and educational systems.

Outcome

16. As a result of the services I received, I am better able to do things I want to do.

General Satisfaction

17. Overall, I am satisfied with the services I received.
18. I would recommend the services to a friend who is need of similar help.

TPS survey forms for both adults and youth are available in 13 languages (English, Spanish, Chinese, Tagalog, Farsi, Arabic, Russian, Hmong, Korean, Eastern Armenian, Western Armenian, Vietnamese, Cambodian) and in one-page and two-page (larger font) versions. The relevant MHSUD Information Notices, survey instructions, forms in multiple threshold languages, and other materials (i.e., Frequently Asked Questions, TPS Codebook, sample county and program summary reports) are available online at http://www.uclaisap.org/dmc-ods-eval/html/client-treatment-perceptions-survey.html.

County administrators coordinate the survey administration and data collection within their provider network and submit the paper forms or electronic data files to UCLA for processing. The data are analyzed, and county- and provider-level summary reports are prepared and made available to participating counties. Counties are also given access to their raw data files and respondents' written comments.

Recovery Incentives Program-Specific Data Collection

Recovery Incentives Program Client Surveys

Recovery Incentives Program treatment providers will be asked to distribute a link to an online survey to participating clients. Surveys will be conducted in multiple waves:

- Baseline survey: At the beginning of Recovery Incentives Program treatment (e.g., intake), providers will be asked to provide clients with a link to a UCLA survey and encourage participation. Clients who go to that survey will receive information about the evaluation, provide consent to participate, and will be asked for contact information and a small number of baseline questions.
- Follow-up surveys: Follow-up surveys will be sent to all clients who completed the baseline survey and provided contact information and consent to be contacted for the follow-up survey. The follow-up surveys will occur early and late in treatment, for example five and 13 weeks after the client began treatment. It will capture information on client perceptions of the Recovery Incentives Program, client functioning (e.g., drug use, use of emergency room and hospital services, etc.) including success stories and perceptions of Recovery Incentives Program implementation needing improvement. These surveys will include clients who are still in treatment and those who left treatment. Participants who have left treatment may be more forthcoming in disclosing what aspects of the pilot program did not work well and monitoring for fraud, e.g., if they indicate they weren't using drugs but were recruited by the agency to participate for money. If resources allow, a small cohort may be selected for brief weekly follow-ups to collect information on client perceptions that may help refine the incentive algorithm.

A sample of up to 60 participating provider sites will be asked to provide the baseline survey link and QR code to all new clients who start participating in the Recovery Incentives Program until they reach a target N. Each provider's quota will be based on estimates of Recovery Incentives Program participation that each site provided during the initial application process, or on estimates based on CalOMS-Tx data. The goal of using quotas is to ensure a representative sample across providers, rather than a potentially biased sample from high-performing providers. The total sample will be approximately 600

Participants will be paid a small incentive (e.g., $10) by UCLA to participate. Survey participant eligibility will be verified against data from the incentive manager vendor to avoid participation by people who are not participating in the Recovery Incentives Program. Initially, incentives will likely take the form of an electronic gift card handled separately from the incentive manager. [10]

[10] Although payments would ideally be handled through the incentive manager, this can only occur if the incentive management vendor is able to implement this and the Internal Revenue Service rules that the Recovery Incentives Program incentives are not income. This ruling is pending. If the IRS does not provide a ruling or rules that the incentives are income, then survey incentives cannot be provided through the incentive manager because it could put the participant over the $599 limit and subject the client to income taxes. It is likely the evaluation will begin with electronic gift cards but providing incentives through the incentive management vendor may become possible in later stages of the evaluation.

During a targeted stakeholder Recovery Incentives Program call on 4/12/2022, a large provider confirmed that they would be able to implement baseline surveys this during client enrollment either using iPads, a desktop computer, or asking the client to use their phone to complete the short survey. If needed, they also expressed a willingness to send the link to clients through their approved method of communication.

Once we reach the target N for each provider, they will be asked to stop distribution of the survey link, and new clients from their site will not be allowed to participate (e.g., the survey link may deactivate when a quota for each provider is reached). Based on CalOMS-Tx data, about 5,000 stimulant users participate in publicly funded treatment annually. If only half of these clients participate in the Recovery Incentives Program, as many as 2,500 new clients per month may be admitted. However, previous evaluation findings suggest implementation of new DMC-ODS waiver benefits typically ramp up slowly over time (Urada et al., 2022). We conservatively assume the survey will be initially offered to 600 clients per month. If the response rate is 50% (300/month), it would take only two months to reach the target N of 600.

Data may also be collected from new participants one or more years after implementation has begun to determine whether client responses change after implementation has matured. Methods may mirror those used to collect the initial sample or may involve re-opening the survey for a longer period but only accepting a random sample of respondents to extend the data collection period over a longer period. Methods for these later waves will be based on the degree of success and lessons learned from the initial data collection.

Approximately four weeks after each participant's baseline survey, UCLA will contact the client for a follow-up survey. A 50% response rate from the 600 baseline participants would result in 300 surveys. Participants will also receive incentives for the follow-up survey. This second wave of surveys will include people who may have stopped participating in treatment. Among clients who remained in treatment during the second wave of surveys, a third wave of surveys will occur at a later date, e.g., 13 weeks (estimated N=150), after they have entered the maintenance phase of treatment, resulting in a total of approximately 1,050 client surveys.

Recovery Incentives Program Client Interviews

The study team will conduct semi-structured interviews with approximately 25 clients purposively selected from participants in the baseline Recovery Incentives Program survey who provided permission for UCLA to contact them for an interview during that survey. Participants will be selected to represent a range of perspectives on the Recovery Incentives Program expressed in surveys. Participants will be asked about the program's strengths and ways the program can be improved. Interviews will be recorded, transcribed, and coded using a constructivist grounded theory approach (Charmaz, 2017; Glaser & Strauss, 1967). Recovery Incentives Program client interviews will begin 24 weeks after baseline client survey to allow completion of survey data collection and to allow time for clients to complete or drop out of treatment.

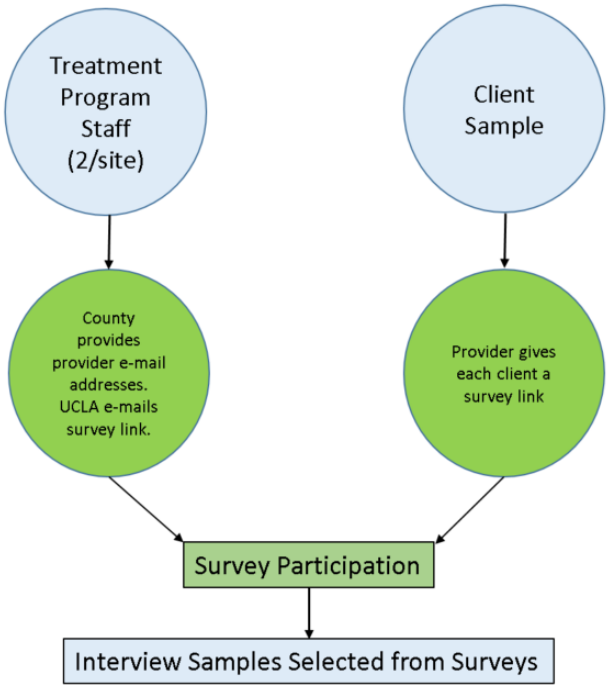
Recovery Incentives Program Provider Surveys

For the Recovery Incentives Program evaluation, provider staff will be surveyed about Recovery Incentives Program implementation, challenges, beliefs, and perceptions and to check for signs of fraud. Counties will be asked to provide an email contact for their participating treatment programs, and evaluators will contact these programs to have online survey invitations sent to the Recovery Incentives Program coordinator and a counselor at each site. The surveys will be conducted online via Qualtrics, early in the implementation process and after the program has participated for approximately six months. A minimum sample of 100 sites will be surveyed and depending on the number of providers sites opting into the Recovery Incentives Program, all sites may be surveyed.

Recovery Incentives Program Provider Interviews

In addition, for the Recovery Incentives Program evaluation the study team will conduct interviews and/or focus groups with a sample of about 25 total provider individuals from agencies that implement the Recovery Incentives Program. Interviews will begin shortly after provider survey data collection has been completed and will end when additional themes cease to emerge from data collection (saturation has been achieved). Interviews and focus groups will focus on identifying the strengths and weaknesses of the Recovery Incentives Program and potential ways to improve the uptake and effectiveness of the program. Interviews and focus groups will be recorded, transcribed, and coded using a constructivist grounded theory approach.

Figure 1. Relationship of Recovery Incentives Program Staff and Client Surveys and Interviews



Fidelity Assessments

California's Recovery Incentives Program training and technical assistance team will collect data on provider knowledge and attitudes during registration for trainings (pre-data), and again after required Recovery Incentives Program trainings have been completed. Following trainings, all participants will receive a link to a post-training test. Providers will also engage in fidelity monitoring sessions twice in the first six months, then every six months thereafter. Tools for these sessions are still in development, but it is anticipated that programs will be rated as high- or low- fidelity through a combination of these fidelity assessments (e.g., trainer's assessments of provider performance on role-playing sessions) and analysis of incentive manager data to measure fidelity to the incentive schedule.

Analytic Methods

Analysis of Quantitative Data

Due to the size of California's population and the associated statistical power available for analysis of statewide databases, comparisons using inferential statistics on many of the datasets used in this report may suggest statistical significance even when these differences are small and not meaningful. Furthermore, inferential statistics are designed to make inferences about a population from a random sample taken from that population. However, many of the datasets used in this evaluation (e.g., DMC claims, CalOMS-Tx, county administrator survey data with near 100% response rates) represent data on essentially the full population of interest rather than a random sample. Therefore, in cases where *p*-values may be inappropriate or misleading, descriptive statistics will be used with percentages, odds ratios, or other methods to convey the size and meaning of differences to readers. However, advanced statistics will also be used to examine multivariate relationships and difference-in-difference analyses as described below.

Event Study (ES) and Difference-in-Difference (DD) designs will be used where appropriate to analyze whether the introduction of the DMC-ODS waiver causally affected certain outcomes of interest. Specifically, we will use these designs when analyzing administrative data (e.g., DMC claims and CalOMS-Tx) for some outcomes. Given the staggered introduction of the DMC-ODS waiver across counties in California over time, exploiting this variation within the ES and DD designs will continue to allow us to estimate a causal effect of the DMC-ODS waiver. These analyses will cover the entirety of the DMC-ODS waiver, including the Medi-Cal 2020 years inclusive of the 2021 extension, and CalAIM. At least 24 months of data (starting in 2015) will also be used for pre-DMC-ODS years.

The canonical difference-in-differences model compares pre-post changes in outcomes in treated units to pre-post changes in outcomes in untreated units, for a single treatment. Given the variation in treatment timing, i.e., the variation in introduction of the DMC-ODS waiver and

programs adopting the Recovery Incentives Program across counties in California over time, exploiting this variation within the ES and DD designs will continue to allow us to estimate a causal effect of the DMC-ODS waiver and the Recovery Incentives Program. This will remain true if new counties opt-in to participate in the DMC-ODS waiver. The widely accepted empirical strategy in this context is the Two-Way Fixed Effect Difference-in-Differences model (2WFE DD) given in the following equation:

$$Y_{it} = \beta_0 + \beta_1 \cdot Treat_{it} + \alpha_t + \theta_i + \epsilon_{it}$$

where *Treat* is a binary variable equal to one when a county or Recovery Incentives Program goes live in the DMC-ODS waiver and equal to zero otherwise; α_t is a time vector containing indicators for the years of data available; and θ_i is a unit vector containing indicators for the 58 counties. Standard errors are clustered by county. The above equation can be modified to include a vector of provider and/or county level time-varying controls. The Average Treatment effect on the Treated (ATT) is given by β_1.

Identification of β_1 comes from within-county variation in DMC-ODS waiver or Recovery Incentives Program implementation during our sample period. The main assumption of DD designs is the parallel trends assumption. This assumption states that in the absence of treatment, the unobserved differences between the treatment and control groups would be similar over time. Although we cannot directly test this assumption, we can assess the assumption in this setting in at least two ways:

1. Include a county-specific linear time trend in the estimating equation. This will control for unmeasured county trends unfolding linearly (e.g., sentiment towards SUD treatments).
2. Perform an event study analysis. This is done by including leads and lags of the DMC-ODS or Recovery Incentives Program indicator variable in the equation above. Ideally, the coefficients on all of the leads of the DMC-ODS or Recovery Incentives Program indicator variable will be statistically insignificant. This will indicate that trends in the main outcomes of interest in the treated and control counties were not trending differently prior to DMC-ODS or Recovery Incentives Program adoption.

We can also modify the above equation to estimate lagged effects and heterogeneous effects of the DMC-ODS waiver or Recovery Incentives Program. Specifically, we can determine if the programs have stronger (or weaker) effects over time and if the effects differ by patient demographics, or by fidelity. For the latter, to determine if the impact of the Recovery Incentives Program differs by high versus low fidelity providers, we can add an interaction term to the above regression, interacting an indicator for high fidelity providers (e.g., high fidelity providers equals one, and zero otherwise) with the Recovery Incentives Program indicator.

Specifically, the DD design will compare the post-treatment (e.g., post-DMC-ODS waiver or Recovery Incentives Program implementation) difference in the outcomes of interest between the

DMC-ODS waiver and State Plan counties (or Recovery Incentives Program and non-Recovery Incentives Program sites/counties) to the pre-treatment (e.g., pre-DMC-ODS implementation/pre-Recovery Incentives Program) difference in the outcomes of interest between DMC-ODS waiver and State Plan counties (or Recovery Incentives Program and non-Recovery Incentives Program sites/counties).

We will do robustness checks to determine if both sets of fixed effects and county/provider controls are needed. Specifically, we will start with a model that only includes time and county fixed effects. We will then estimate another model that includes both sets of FEs plus county and/or provider controls. If the estimates are very similar, we likely do not need to include the controls. However, we will still present both sets of estimates to show how robust they are to strengthen our conclusions about the effect we are seeing. This is standard practice in nearly ever published difference-in-difference paper (including both FEs and time-varying controls). The FEs are only picking up time-invariant county and provider effects. But, if we know things like the poverty rate, unemployment rate, COVID policies, etc. vary across counties and across time, we need to include those in the regression.

The 2WFE DD model captures average treatment effects on the treated but does not allow us to consider time-varying treatment effects. There are several reasons to expect the effects of DMC-ODS waiver to vary over time. To account for potentially time-varying treatment effects, we implement difference-in-differences decomposition (Goodman-Bacon, 2021, Callaway, et. al 2021, Callaway, et. al 2021, Dave, et. al, 2020).

The 2WFE DD estimate is composed of a weighted average of treatment effects estimated from a series of 2x2 treatment/control groups, some of which compare counties treated at the same time to untreated counties, and others compare counties treated at the same time to counties treated at another time (earlier or later).

Comparisons may also be made to always-treated units; however, given that no always-treated counties comprise of only 4% state population and do not form an appropriate comparison group for our treated counties, we cannot pursue this comparison to derive robust average treatment effects. There are 19 timing groups in our data, or groups of counties which experience going live in the DMC-ODS waiver at the same time. There are thus 361 distinct 2x2 treatment/control comparison groups from which the 2WFE DD estimate is constructed: 342 groups in which earlier-treated counties are compared to later-treated counties, or vice versa, and 19 groups in which treated counties are compared to untreated counties. In the presence of time-varying treatment effects, comparisons between earlier and later treated counties may introduce bias into the 2WFE DD estimate. The extent of the bias depends on the share of the 2WFE DD estimate that is derived from these earlier-later comparisons, which in turn depends on group size and the variance of the treatment.

Goodman-Bacon (2021) has developed a method to decompose the 2WFE DD estimate into the 2x2 weighted estimates from which it is derived. Using this difference-in-differences decomposition model, we can uncover the extent to which the 2WFE DD estimate depends on 2x2 DD estimates which compare earlier to later treated counties. The Goodman-Bacon decomposition model is currently only available for strongly balanced panels in which treatment only changes from 0 to 1 over time. To estimate the decomposition model, we define treatment as a binary variable that is equal to one in all years after a county goes live in the DMC-ODS waiver and is equal to zero otherwise. The ES design is similar to the DD design but will allow the effect of the DMC-ODS waiver to vary from a specified number of months prior to introduction of the waiver to a specified number of months after the introduction.

All ES and DD models will continue to use data from either DMC claims or CalOMS-Tx at the county-month-year-level, and control for time-invariant county effects, county-invariant time effects, and the severity of the COVID-19 pandemic, which may be proxied by the county-level COVID-19 case rate per 100,000, and COVID-19 death rate per 100,000 for each month-year cell. All regressions will be weighted by the county population, and standard errors are clustered at the county level (Bertrand, 2004).

The Generalized Synthetic Control (GSC) method introduced by Xu (2017) addresses the case when treatment is imposed at different times for different counties. This approach allows for multiple treated counties and variable treatment periods. This method also has several other advantages. It includes a built-in cross-validation procedure and is easier to implement than other synthetic control methods. The GSC method allows us not only to match counties on pretreatment observables, but also to model unobserved time-varying heterogeneities using interactive fixed effects.

GSC first estimates an Interactive Fixed Effects (IFE) model using only the counties that were never treated and obtains a fixed number of time-varying coefficients (latent factors). It then estimates county-specific intercepts (factor loadings) for each treated county by linearly projecting pretreatment outcomes for treated counties onto the space spanned by the factors. Finally, it generates synthetic control units based on the estimated factors and factor loadings. The method is described as a "bias correction procedure for IFE models when treatment data is heterogeneous across units." (Xu, 2017)

Of note, given that many of DMC-ODS benefits have now been adopted by the state plan,[11] it raises concern regarding the DMC-ODS period under analysis. Since the control group will have similar provisions as DMC-ODS, this falls under spillover effects and violates the assumption of quasi-experimental causal inference methods, called SUTVA (Stable Unit Treatment Value Assumption). This may change the magnitude of estimates. However, since residential treatment will be treated in IMDs in DMC-ODS counties but not in state plan counties, DMC-ODS should

[11] https://www.dhcs.ca.gov/Documents/CA-21-0058-Approval-Package.pdf

maintain an access advantage to residential treatment and we can continue the analysis with the caveat that the average DMC-ODS effect may be reduced after these changes to the state plan.

Analysis of Cost Data

Using the causal inference study designs mentioned above, including but not limited to Difference-in-Difference with staggered implementation, Synthetic Control Methods, or Generalized Synthetic Control Methods, as applicable, UCLA will examine the changes in costs because of DMC-ODS waiver. These costs analyses will focus on total Medicaid and Federal costs as well as cost drivers measured per member per month. Specifically, the analyses will also focus on changes in inpatient costs, non-ED outpatient costs, and ED outpatient costs. The analyses will be based on administrative data provided by DHCS; namely, DMC-ODS claims and Managed Care/FFS claims starting in 2015 (pre-period) and including DMC-ODS implementation from 2017 through CalAIM (including 2021). The DMC-ODS claims data contain all SUD-related claims of Medi-Cal beneficiaries, whereas the Managed Care/FFS claims are all managed care claims of SUD beneficiaries identified in the DMC-ODS claims. This will allow us to identify increased access to residential treatment (the prime goal of DMC-ODS waiver) from DMC-ODS claims data and follow the cost behavior of beneficiaries through the variables and data provided in the Managed Care/FFS claims. A potential hypothesis that UCLA will explore involves cost shift behavior from high-value emergency services (ED costs) to Residential Treatment. However, given that seven Partnership HealthPlan counties joined the DMC-ODS waiver as a regional model on July 1, 2020, it will be difficult to analyze any changes in costs for these counties, given the data lag in sharing Managed Care/FFS claims. Currently, there is a 2–3-year lag, and UCLA is awaiting Managed Care/FFS claims data for 2021. So, the analyses will focus on counties where sufficient post-waiver data is available (if a balanced panel is desired for computation purposes). For Recovery Incentives Program, the evaluation team will also use managed care/fee-for-service claims data to analyze cost-effectiveness, specifically investigating whether emergency department, inpatient hospital utilization, and other medical costs (including any type of physical health problems) are reduced or made more appropriate (e.g., increased primary care costs but reduced emergency department costs) among clients who participated in the Recovery Incentives Program vs. similar clients who did not. Until this data becomes available (projected 2025), UCLA will rely on client self-reports from surveys and interviews, as described above.

Analysis of Recovery Incentives Program Incentive Manager Data

Rates of positive drug tests will be compared to rates from the CM literature using a one-sample t-test or analogous procedure. UCLA reviewed all studies cited in a recent systematic review of CM trials for the treatment of methamphetamine use (Brown & DeFulio, 2020), supplemented by a PubMed search of 2020-2022 articles with the key terms “contingency management” and “stimulant.” Among these sources, three studies (Roll & Shoptaw, 2006; Stitzer et al., 2020;

Strona et al, 2006) reported information sufficient to calculate the percentage of negative results among submitted tests. The average, weighted for study size, was 85.3%.

However, Miguel et al. (2021) determined that the percent of negative urinalysis outcomes out of *all possible* tests showed the most consistent performance, compared to alternative measures e.g., weeks of continuous abstinence. This measure conservatively treats missed tests the same as positive tests. Therefore, a measure similar to this will also be used for the evaluation. Three articles (Carrico et al., 2015; Shoptaw et al, 2006, Miguel et al, 2021) reported sufficient information to calculate the percentage of negative urinalysis results among all possible tests, producing a weighted average of 47.7%.

If the data allows, more advanced techniques (e.g., growth curve modeling) may be used to examine patterns in the drug test data.

Analysis of Quantitative Survey Data

County administrator, provider, and client surveys will include Likert rating scales and binary measures (e.g., yes/no). While the lower Ns for the administrator surveys will mostly limit analyses to descriptive analyses, the provider and client survey data will be analyzed in greater depth.

Descriptive statistics, including mean and standard deviation for continuous outcomes as well as frequency and percentage for binary outcomes, will be estimated for all survey samples. Bivariate comparisons will be made between coordinators and counselors in the case of provider surveys.

The association between pairs of measures in surveys will be estimated using product-moment correlation for continuous measures, point-biserial correlation for the relationship between categorical and continuous measures, and cross-tabulation for categorical measures.

Multiple regression modeling for a continuous outcome (e.g., a 1-5 Likert rating scale) and/or logistic regression modeling for a binary outcome (i.e., yes/no) will be conducted separately. On provider surveys, the staff's role (i.e., coordinators versus counselors) will be a covariate in regression modeling.

For client data, which will consist of multiple waves, descriptive analyses, and trajectory plots in conjunction with the Generalized Growth Curve Model (GGCM) may be applied to examine the change in client responses across the repeated assessments.

All analyses will be conducted at both statewide and county levels, by fidelity level, and by demographic groups to look for differences in access and outcomes by race, ethnicity, gender, and age.

Power analysis

Since statistical significance is a way of evaluating the likelihood that differences found in a sample would be found in the full population, in the case of the main administrative data analyses statistical power will not come into play because we are analyzing the data from essentially the full population. The same is true of surveys of county administrators, since we will be surveying the entire administrator population among counties participating in the Recovery Incentives Program, and we have historically approached a 100% response rate for surveys of California county administrators. For surveys of treatment providers and clients, however, statistical power will become a consideration, since we will be surveying samples of a broader population.

Although the Ns may need to be adjusted based on resource availability, our current estimated Recovery Incentives Program sample sizes of 600 Wave 1 and 300 Wave 2 client surveys will be sufficient to detect a difference in a continuous measure between the waves with a small effect size (d) of 0.20. Our estimated sample size of 300 Wave 2 and 150 client surveys will be sufficient to detect a small effect size of 0.28. Provider surveys from 100 sites will be sufficient to detect a medium effect size of 0.57 when divided into two groups of 50 (e.g., higher and lower fidelity sites). All power analysis computations were computed with a two-sided alpha of .05 and power (beta) of .80.

Analysis of Qualitative Data

Qualitative data will be collected from providers and county administrators through interviews and focus groups. Qualitative data collection will focus on the major themes of the overall evaluation, as well as emerging trends related to SUD and SUD treatment in California. If client perspectives are needed beyond the information they provide through the treatment perception survey, they may also be interviewed, and UCLA will stratify to the extent possible to ensure a representative sample. Qualitative data will be used to contextualize and inform the interpretation of quantitative findings, and identify areas that warrant further inquiry or focus in the evaluation. All interviews and group discussions will be recorded and transcribed, while qualitative data from surveys (e.g., free text responses to open-ended questions) will be extracted and organized into a spreadsheet. Where applicable, the evaluation team will then analyze these data using a systematic and iterative process according to established and accepted procedures for qualitative research.

Recovery Incentives Program evaluation qualitative data analysis

Qualitative data will be collected from different stakeholders, including clients, providers, and county administrators. All interviews and group discussions will be recorded and transcribed, while qualitative data from surveys (e.g., free text responses to open-ended questions) will be

extracted and organized into a spreadsheet. The evaluation team will then analyze these data using a systematic and iterative process according to established and accepted procedures for qualitative research. This process will begin by organizing data into key study domains (King 2004) related to the RE-AIM framework (Reach, Effectiveness, Adoption, Implementation, Maintenance). Within each domain, initial analyses will utilize preliminary codes that are expected to emerge from qualitative data. See Table 2 for a preliminary list of codes that may be used to guide analyses and identify overarching data trends.

Table 2. Preliminary Codes for Recovery Incentives Program Qualitative Data Analysis

RE-AIM DOMAIN	PRELIMINARY CODES
Reach	R1: What determines which StimUD clients receive CM and which do not? R2: What are the barriers and facilitators of Recovery Incentives Program service delivery? R3: Are there disparities in the reach of Recovery Incentives Program services to different treatment populations? What can be done to mitigate these disparities?
Effectiveness	E1: How effective do stakeholders believe the Recovery Incentives Program is in helping clients remain in treatment? Helping them achieve and maintain abstinence from stimulants? E2: Are there aspects of the Recovery Incentives Program (incentives, testing procedures) or other behavioral services and supports delivered in conjunction with CM) that seem to enhance or inhibit the Recovery Incentives Program's effectiveness? E3: What can providers do to enhance the Recovery Incentives Program's effectiveness with the clients they serve? What can administrators and policymakers do to facilitate these changes?
Adoption	A1: What factors do counties consider when deciding whether to participate in the Recovery Incentives Program? What factors do program leaders and individual providers consider? A2: What are the practical barriers to/facilitators of Recovery Incentives Program adoption? A3: What policies and procedures could help promote the effective adoption of the Recovery Incentives Program?

Implementation	I1. What are the barriers to/facilitators of high-fidelity CM implementation? I2. What adaptations are being made to CM as it is being implemented? What impacts do these have on intervention fidelity and effectiveness? I3. What policies and procedures could help promote the effective implementation of the Recovery Incentives Program?
Maintenance	M1. What makes programs and providers decide to participate in the Recovery Incentives Program? What makes them decide to discontinue it? M2. What policies and procedures could help promote the maintenance of the Recovery Incentives program in the future if it becomes a standard Medi-Cal benefit?

After organizing qualitative data with codes, we will use constructivist grounded theory to guide the process of reading transcripts, developing code lists, coding data, and comparing/contrasting emerging patterns and themes using constant comparative methods (Charmaz, 2017; Glaser & Strauss, 1967). Portions of coded transcripts will be randomly and independently coded by two researchers to ensure that the codes are being applied consistently and have acceptable levels of agreement indicating good reliability. The evaluation team will meet regularly to share insights and observations from the interviews and/or focus groups throughout the evaluation and discuss emerging themes. Researchers will review the analytic findings. Qualitative data will be triangulated with survey and other quantitative data to identify areas where the results from the data sets converge, complement one another, and/or expand on one another (Creswell, 2003; Palinkas et al., 2011).

The qualitative data collected from the different stakeholder groups (e.g., county administrators, treatment providers, clients) will be analyzed separately as well as across the different groups, and over time (e.g., early vs. later in the implementation of the project) to identify themes and patterns. Findings will be shared with members of key stakeholder groups (DHCS, county administrators, and program staff) to verify and interpret findings.

Methodological Limitations

The California Administrative data sets used in this evaluation have many of the same shortcomings as other administrative data sets, particularly related to inconsistent reporting and missing data (see, for example, Evans et al., 2010 for a discussion of CalOMS-Tx). Delays in data reporting also limit analyses of recent data. UCLA will analyze CalOMS-Tx and DMC claims using the most recent available complete data, which typically requires disregarding

approximately the most recent 6 months of data due to data reporting lag. This will limit the amount of data that can be used in early reports.

CalOMS-Tx data is partly reliant on self-reported data, particularly with respect to outcome questions (e.g., drug use in the last 30 days). Some terms are also somewhat subjective, like discharge status terms (e.g., completed treatment, satisfactory progress, and unsatisfactory progress). To partly ameliorate this problem, these categories will be combined into "successful" (completed, satisfactory progress) and "unsuccessful" (unsatisfactory progress) discharges.

DMC claims data tend to be more complete than CalOMS-Tx data because providers are more motivated to submit them quickly for payment, but this is not universally true. In some cases, under the DMC-ODS, new billable services (e.g., recovery services) are being delivered but DMC claims are not being submitted, in part due to confusion over what is allowable. While this seems less likely to occur for the relatively well-defined Recovery Incentives Program, UCLA will monitor provider survey and interview responses for signs of billing difficulties that may affect claims data.

While DMC claims data have an advantage over CalOMS in completeness, CalOMS-Tx has advantages in the depth of data. CalOMS includes client background (e.g. demographics, primary and secondary drug, source of treatment referral, number of prior treatment episodes, housing, employment, criminal justice status, number of children), as well as treatment discharge status and a number of outcome measures in the last 30 days, both at admission and discharge (e.g. number of arrests & jail days, family conflicts, social support). These cannot be derived from claims. These datasets are therefore complimentary and can be used together to develop a better understanding of DMC-ODS implementation than either dataset alone.

Interview and survey data are limited by the honesty of respondents and the response rate.

Wherever possible, different types of data will be examined in parallel to converge on underlying constructs being measured and thereby mitigate the limitations of each dataset.

The long time frame since initial implementation of DMC-ODS could introduce challenges in interpreting the data, since external impacts (e.g., COVID-19, changes in the economy and workforce) will affect trends. This will be particularly true if no or few new counties join DMC-ODS and if external impacts have systematically different effects on DMC-ODS and non-DMC-ODS counties. Given that the state's largest counties are all already participating in DMC-ODS, even if new counties do opt-in to DMC-ODS, they would also likely have small beneficiary and treatment client populations and a correspondingly limited impact on statewide analyses. The Recovery Incentives Program evaluation will also depend on implementation of a new program, which has and could continue to experience unforeseen delays or barriers that prevent or limit the planned implementation.

References

Arkansas Center for Health Improvement. (2014). Arkansas Health Care Independence Program ("Private Option"): Proposed evaluation for Section 1115 Demonstration Waiver. Retrieved from http://www.medicaid.gov/Medicaid-CHIP-Program-Information/By-Topics/Waivers/1115/downloads/ar/Health-Care-Independence-Program-Private-Option/ar-private-option-demo-waiver-proposed-eval-02202014.pdf

AshaRani, P. V., Hombali, A., Seow, E., Ong, W. J., Tan, J. H., & Subramaniam, M. (2020). Non-pharmacological interventions for methamphetamine use disorder: a systematic review. *Drug and Alcohol Dependence*, *212*, 108060.

Bass, B., Urada, D., Joshi, V., Padwa, H., Huang, D., Lee, A.B., Gilbert, M., & Antonini, V. (2022). Access to Care. In: Drug Medi-Cal Organized Delivery System FY 2021 Evaluation Report, 35-52. Los Angeles: UCLA Integrated Substance Abuse Programs. Available at: https://www.uclaisap.org/dmc-ods-eval/assets/documents/20220422-DMC-ODS-FY-2021-Evaluation-Report-with-Appendices.pdf

Bass, B., Urada, D., Lee, A.B., Padwa, H., & Tran, E. (2022b). Special Topic: Impact of COVID-19 on SUD Treatment and the Emergence of Telehealth. In: Drug Medi-Cal Organized Delivery System FY 2021 Evaluation Report, 89-102. Prepared for the Department of Health Care Services, California Health and Human Services Agency. Los Angeles: UCLA Integrated Substance Abuse Programs. Available at: https://www.uclaisap.org/dmc-ods-eval/assets/documents/20220422-DMC-ODS-FY-2021-Evaluation-Report-with-Appendices.pdf

Bertrand, M., Duflo, E., & Mullainathan, S. (2004). How much should we trust differences-in-differences estimates?. *The Quarterly journal of economics*, 119(1), 249-275.

Brown, H. D., & DeFulio, A. (2020). Contingency management for the treatment of methamphetamine use disorder: a systematic review. *Drug and Alcohol Dependence*, *216*, 108307.

Callaway, B., Goodman-Bacon, A., & Sant'Anna, P. H. (2021). Difference-in-differences with a continuous treatment. *arXiv preprint arXiv:*2107.02637.

Callaway, B., & Sant'Anna, P. H. (2021). Difference-in-differences with multiple time periods. *Journal of Econometrics*, 225(2), 200-230.

Carrico, A. W., Gómez, W., Siever, M. D., Discepola, M. V., Dilworth, S. E., & Moskowitz, J. T. (2015). Pilot randomized controlled trial of an integrative intervention with methamphetamine-using men who have sex with men. *Archives of Sexual Behavior*, *44*(7), 1861-1867.

Charmaz, K. (2017). The power of constructivist grounded theory for critical inquiry. *Qualitative inquiry*, *23*(1), 34-45.

Creswell, J. W. (2003). *Research design: Qualitative, quantitative, and mixed methods approaches (2nd ed.)*. Thousand Oaks, CA: Sage Publications.

Darfler, K., Santos, A., Gregario, L., Vazquez, E., Bass, B., Joshi, V., Antonini, V., Hall, E., Teruya, C., Sandoval, J., & Urada, D. (2020). California Hub and Spoke MAT Expansion Program: Final Evaluation Report. Prepared for the California Department of Health Care Services, California Health and Human Services Agency. Los Angeles: UCLA Integrated Substance Abuse Programs. Available at: https://www.uclaisap.org/ca-hubandspoke/docs/reports/CA-HSS-Evaluation-Final-Report.pdf

Dave, D., Friedson, A., Matsuzawa, K., Sabia, J. J., & Safford, S. (2020). JUE Insight: Were urban cowboys enough to control COVID-19? Local shelter-in-place orders and coronavirus case growth. *Journal of urban economics*, 103294.

De Crescenzo, F., Ciabattini, M., D'Alò, G. L., De Giorgi, R., Del Giovane, C., Cassar, C., ... & Cipriani, A. (2018). Comparative efficacy and acceptability of psychosocial interventions for individuals with cocaine and amphetamine addiction: A systematic review and network meta-analysis. *PLoS Medicine*, *15*(12), e1002715.

Evans, E., Grella, C. E., Murphy, D. A., & Hser, Y. I. (2010). Using administrative data for longitudinal substance abuse research. *The journal of behavioral health services & research*, 37(2), 252-271

Farrell, M., Martin, N. K., Stockings, E., Bórquez, A., Cepeda, J. A., Degenhardt, L., ... & McKetin, R. (2019). Responding to global stimulant use: challenges and opportunities. *The Lancet*, *394*(10209), 1652-1667.

Fedeli, U., Zorzi, M., Urso, E. D., Gennaro, N., Dei Tos, A. P., & Saugo, M. (2015). Impact of fecal immunochemical test-based screening programs on proximal and distal colorectal cancer surgery rates: A natural multiple-baseline experiment [Epub ahead of print]. *Cancer*. doi:10.1002/cncr.29623

Glaser, B. G. & Strauss, A. L. (1967). The Discovery of Grounded Theory. Strategies for Qualitative Research. Chicago: Aldine. King, N. (2004) Using templates in the thematic analysis of texts. In C. Cassell & G. Symon (eds) *Essential Guide to Qualitative Methods in Organizational Research.* London: Sage Publications. Pp. 256-270.

Glasgow RE, Vogt TM, Boles SM. (1999). Evaluating the public health impact of health promotion interventions: the RE-AIM framework. *Am J Public Health*. 1999;89(9):1322–1327.

Goodman-Bacon, A. (2021). Difference-in-differences with variation in treatment timing. *Journal of Econometrics*, 225(2), 254-277.

Khurana, D., Urada, D., Bass, B., & Joshi, V. (2022). Special Topic: Partnership HealthPlan Cost Study. In: Drug Medi-Cal Organized Delivery System FY 2021 Evaluation Report, 109-118. Los Angeles: UCLA Integrated Substance Abuse Programs. Available at: https://www.uclaisap.org/dmc-ods-eval/assets/documents/20220422-DMC-ODS-FY-2021-Evaluation-Report-with-Appendices.pdf

Lee, A.B., Antonini, V.P., Huang, D., Joshi, V., Urada, D., & Tran, E. (2022). Integration and Coordination of Care. In: Drug Medi-Cal Organized Delivery System FY 2021 Evaluation Report, 70-88. Los Angeles: UCLA Integrated Substance Abuse Programs. Available at: https://www.uclaisap.org/dmc-ods-eval/assets/documents/20220422-DMC-ODS-FY-2021-Evaluation-Report-with-Appendices.pdf

Manja V., Nrusimha A., Gao Y., Sheikh, A., McGovern, M., Heidenreich, P.A., Tarlochan, S.S., & Asch, S. (2023). Methamphetamine-associated heart failure: a systematic review of observational studies, *Heart*, 109:168-177.

Miguel, A. Q., Smith, C. L., Burduli, E., Roll, J. M., & McPherson, S. M. (2021). Validating the clinical relevance of alternative stimulant use treatment outcome measures by examining their association with 3-month follow-up outcomes. *Experimental and Clinical Psychopharmacology*, *29*(3), 288.

Padwa, H., Joshi, V., Huang, D., Khurana, D., Urada, D., Gilbert, M., Loya, C., & Bass, B. (2022). Quality of Care. In: Drug Medi-Cal Organized Delivery System FY 2021 Evaluation Report, 53-69. Los Angeles: UCLA Integrated Substance Abuse Programs. Available at: https://www.uclaisap.org/dmc-ods-eval/assets/documents/20220422-DMC-ODS-FY-2021-Evaluation-Report-with-Appendices.pdf

Palinkas, L. A., Aarons, G. A., Horwitz, S., Chamberlian, P., Hurlburt, M., & Landsverk, J. (2011). *Mixed method designs in implementation research.* Administration and Policy in Mental Health and Mental Health Services Research, *38*, 44-53.

Roll, J. M., & Shoptaw, S. (2006). Contingency management: schedule effects. *Psychiatry Research*, *144*(1), 91-93.

Ronsley, C., Nolan, S., Knight, R., Hayashi, K., Klimas, J., Walley, A., ... & Fairbairn, N. (2020). Treatment of stimulant use disorder: a systematic review of reviews. *PloS One*, *15*(6), e0234809.

Shoptaw, S., Klausner, J. D., Reback, C. J., Tierney, S., Stansell, J., Hare, C. B., ... & Dang, J. (2006). A public health response to the methamphetamine epidemic: The implementation of contingency management to treat methamphetamine dependence. *BMC Public Health*, *6*(1), 1-5.

Stitzer, M. L., Gukasyan, N., Matheson, T., Sorensen, J. L., Feaster, D. J., Duan, R., ... & Metsch, L. R. (2020). Enhancing patient navigation with contingent financial incentives

for substance use abatement in persons with HIV and substance use. *Psychology of Addictive Behaviors*, *34*(1), 23.

Strona, F. V., McCright, J., Hjord, H., Ahrens, K., Tierney, S., Shoptaw, S., & Klausner, J. D. (2006). The acceptability and feasibility of the Positive Reinforcement Opportunity Project, a community-based contingency management methamphetamine treatment program for gay and bisexual men in San Francisco. *Journal of Psychoactive Drugs*, *38*(sup3), 377-383.

Substance Abuse and Mental Health Services Administration. (2021). Key Substance Use and
Mental Health Indicators in the United States: Results from the 2020 National Survey on Drug Use and Health. Available at: https://www.samhsa.gov/data/sites/default/files/reports/rpt35325/NSDUHFFRPDFWHTMLFiles2020/2020NSDUHFFR1PDFW102121.pdf

Teruya C, Joshi V, Urada D, Trabin T, Iturrios-Fourzan I, Huang YC. Development and Measurement of the Treatment Perceptions Survey (TPS) for Clients with Substance Use Disorders. *J Behav Health Serv Res*. 2022 Apr;49(2):190-203. doi: 10.1007/s11414-021-09776-y. Epub 2022 Feb 25. PMID: 35217967; PMCID: PMC8979878.

Urada, D., Antonini, V.P., Teruya, C., Joshi, V., Lee, A.B., Iturrios-Fourzan, I., Tran, E., Huang, D., & Vazquez, M.E. (2019). Drug Medi-Cal Organized Delivery System 2019 Evaluation Report. Los Angeles, CA: UCLA Integrated Substance Abuse Programs. Available at: http://www.uclaisap.org/dmc-ods-eval/assets/documents/DMC-ODS-Year-4-Evaluation-Report-FY-2018-19.pdf

Urada, D., Antonini, V. P., Teruya, C., Tran, E., Huang, D., Padwa, H., Castro-Moino, K., Lee, A.B., & Grossman, J. (2017). California Drug Medi-Cal Organized Delivery System, FY 2016-2017 Evaluation Report. Los Angeles, CA: UCLA Integrated Substance Abuse Programs. https://www.uclaisap.org/dmc-ods-eval/assets/documents/DMC-ODS-Evaluation-Report-FY-2016-2017%20final.pdf

Urada, D., Bass, B., Lee, A.B., Padwa, H., Khurana, D., Antonini, V.P., Joshi, V., Tran, E., Huang, D., Gilbert, M., & Loya, C. (2022). Drug Medi-Cal Organized Delivery System FY 2021 Evaluation Report. Available at: https://www.uclaisap.org/dmc-ods-eval/assets/documents/20220422-DMC-ODS-FY-2021-Evaluation-Report-with-Appendices.pdf

Urada, D., Bass, B., Padwa, H., Huang, D., Joshi, V., & Loya, C. (2022). Special Topic: Residential Length of Stay. In: Drug Medi-Cal Organized Delivery System FY 2021 Evaluation Report, 103-108. Los Angeles: UCLA Integrated Substance Abuse Programs. Available at: https://www.uclaisap.org/dmc-ods-eval/assets/documents/20220422-DMC-ODS-FY-2021-Evaluation-Report-with-Appendices.pdf

Urada, D., Lee, A.B., Bass, B., Teruya, C., Antonini, V.P., Joshi, V., Padwa, H., Tran, E., Huang, D., & Iturrios-Fourzan, I. (2021). Drug Medi-Cal Organized Delivery System 2020 Evaluation Report. Los Angeles, CA: UCLA Integrated Substance Abuse Programs. Available at: https://www.uclaisap.org/dmc-ods-eval/assets/documents/2020-DMC-ODS-Evaluation-Report-with-Appendices_revised_2021-07-09.pdf

Urada, D., Teruya, C., Antonini, V. P., Joshi, V., Padwa, H., Huang, D., Lee, A.B., Castro-Moino, K., & Tran, E. (2018). California Drug Medi-Cal Organized Delivery System, 2018 Evaluation Report. Los Angeles, CA: UCLA Integrated Substance Abuse Programs. https://www.uclaisap.org/dmc-ods-eval/assets/documents/2017-2018%20UCLA%20DMC-ODS%20Evaluation%20Report%2011192018.pdf

Urada, D., Teruya, C., Antonini, V. P., Tran, E., Huang, D., Padwa, H., & Lim, J. (2016). California Drug Medi-Cal Organized Delivery System, FY 2015-2016 Report: Evaluation of California's Section 1115 Demonstration Waiver. Los Angeles, CA: UCLA Integrated Substance Abuse Programs. Available at: https://www.uclaisap.org/dmc-ods-eval/assets/documents/DMC-ODS-Evaluation-Report-FY-2015-2016%20final.pdf

Xu, Y. (2017). Generalized Synthetic Control Method: Causal Inference with Interactive Fixed Effects Models. *Political Analysis*, 25(1), 57-76. doi:10.1017/pan.2016.2

Appendix A: Driver Diagram