

TRIAGE SUMMATIVE EVALUATION PLAN

PREPARED BY:

DENIS HULETT, MS

KALLIE CLARK, PHD, MSW

KAI LEMASSON, PHD, MS

MARTHA CLEMENTE, MPH, CHES

HEATHER BARR, MPH, CPH

HEIKE THIEL DE BOCANEGRA, PHD, MPH

PRINCIPAL INVESTIGATOR

February 2022

TABLE OF CONTENTS

<i>Executive summary</i>	2
<i>Background</i>	2
MHSA	3
Triage Grants	3
SB 82/833 Logic Model	6
<i>The Triage Summative Evaluation</i>	
Data infrastructure	7
Triage Evaluation Questions	9
Outcome Measures	10
Evaluation Design and Analysis	11
Part 1: Descriptive overview	11
Part 2: impact evaluation	16
Sub-Analyses	18
<i>Timeline</i>	19
<i>Report and Dissemination</i>	21
<i>Limitations and Validity Threats</i>	21
CSI Database	21
Record Linking	22
Spillover Effects	23
Heterogeneous County Mental Health Infrastructures	24
Quality of Life	24
Parallel slopes assumption	24
<i>ENDNOTES</i>	26

APPROACH TO STATEWIDE EVALUATION OF TRIAGE PROGRAMS

EXECUTIVE SUMMARY

The Mental Health Services Oversight and Accountability Commission (Commission) leads the statewide evaluation of SB 82/833 Triage Crisis Services (Triage) grants. These grants allow for increased capacity through hiring personnel to provide crisis intervention, diversion from jails and hospitals, and linkages to mental health treatment appropriate in the community.

The first round of Triage grant funded services were implemented between 2014 and 2017. Grantees conducted their own local evaluations, making it difficult to tell a statewide story. A California State Audit report found that “without the statewide metrics, local Mental Health Services Act stakeholders are unable to fully evaluate the effectiveness of the triage grants” and recommended the Commission conduct a statewide evaluation. In response, the Commission took two actions for the second round of Triage grants. First, the Commission contracted with two University of California campuses (University of California, Los Angeles (UCLA) and University of California, Davis (UCD)) to evaluate the implementation of the Triage awards to children and adults, respectively, (formative evaluation). Second, the Commission leveraged its extensive data analysis and linkage infrastructure to assess the impact of crisis services on client outcomes (summative evaluation).

The summative evaluation outlined below seeks to understand the impact of Triage services on post-crisis emergency department use and hospitalization, arrests and recidivism, employment, and educational outcomes for children. The Commission established data-sharing relationships with several California State agencies including the Department of Justice (DOJ), the Department of Public Health (CDPH), the Employment Development Department (EDD), and the Department of Education (DOE).

The Commission also partnered with Triage grantees to receive information about clients who accessed their crisis services (via a safe transfer and storage system in full compliance with the Health Insurance Portability and Accountability Act of 1996 (HIPAA)). This information is to be linked to databases from the state agencies listed above to evaluate the impact of Triage services on various outcomes. Descriptive and quasi-experimental methods will be employed to examine impacts of receiving Triage services.

BACKGROUND

In 2017, UCD and UCLA began the Triage grant formative and summative evaluations in collaboration with the Commission, state governing bodies, and all other engaged stakeholders. Execution of the summative evaluation was passed on to the Commission in 2019 in efforts to resolve regulatory complexities, expand linkages, and meld resources more easily accessible at the state level. In 2020, the Commission carried out HIPAA compliant, written data use agreements (DUAs) with counties statewide to enable Commission researchers to receive the person-level data needed to make the summative evaluation possible. In continued partnership, UCD is conducting client and law enforcement interviews and surveys for the Adult/ transition-age youth (ages 16-24) (TAY) programs, and UCLA is completing interviews and surveys for

Child/Youth programs and School-County collaboratives. The findings from UCD and UCLA's formative evaluations will provide context and deeper understanding of the Commission's summative evaluation. Together, these evaluations will inform the work of stakeholders across the State of California.

MHSA

In 2004, California voters allowed passage of the Mental Health Services Act (MHSA) to augment California's behavioral health system to better approach and fill needs-based gaps. One percent of personal income tax exceeding \$1 million annually is allocated to serve people and their families faced with, or at risk of, serious mental health challenges¹. The MHSA funds, such as SB 82 and SB 833, are imperative in supporting prevention, early intervention, and an array of treatment service, infrastructural, technological, and training components of the system².

The Commission is responsible for conducting formative, process, and outcome/summative evaluations to address the seven negative client outcomes outlined in the MHSA as associated with unmet mental health needs:

- Suicide
- Incarceration
- School failure or dropout
- Unemployment
- Lack of stable housing
- Removal of children from their homes
- Prolonged suffering

TRIAGE GRANTS

Senate Bill (SB) 82, *the Investment in Mental Health Wellness Act of 2013* provides grant funds to improve access to and capacity for local crisis mental health services. The grant funds the hiring of Triage crisis personnel who provide crisis intervention, treatment, and case management services designed to better meet the needs of individuals experiencing a mental health crisis in the least restrictive manner possible.

The following are the SB 82 legislative objectives:

- Expanding crisis treatment services by adding crisis residential treatment beds, crisis stabilization services, Mobile Crisis Support Teams, Triage personnel,
- Improving the client experience, achieving recovery and wellness, and reducing costs,
- Reducing unnecessary hospitalizations and inpatient stays,
- Reducing recidivism and mitigating unnecessary expenditures of law enforcement, and
- Expanding the continuum of services with early intervention and treatment options that are wellness, resiliency, recovery oriented in the least restrictive environment.

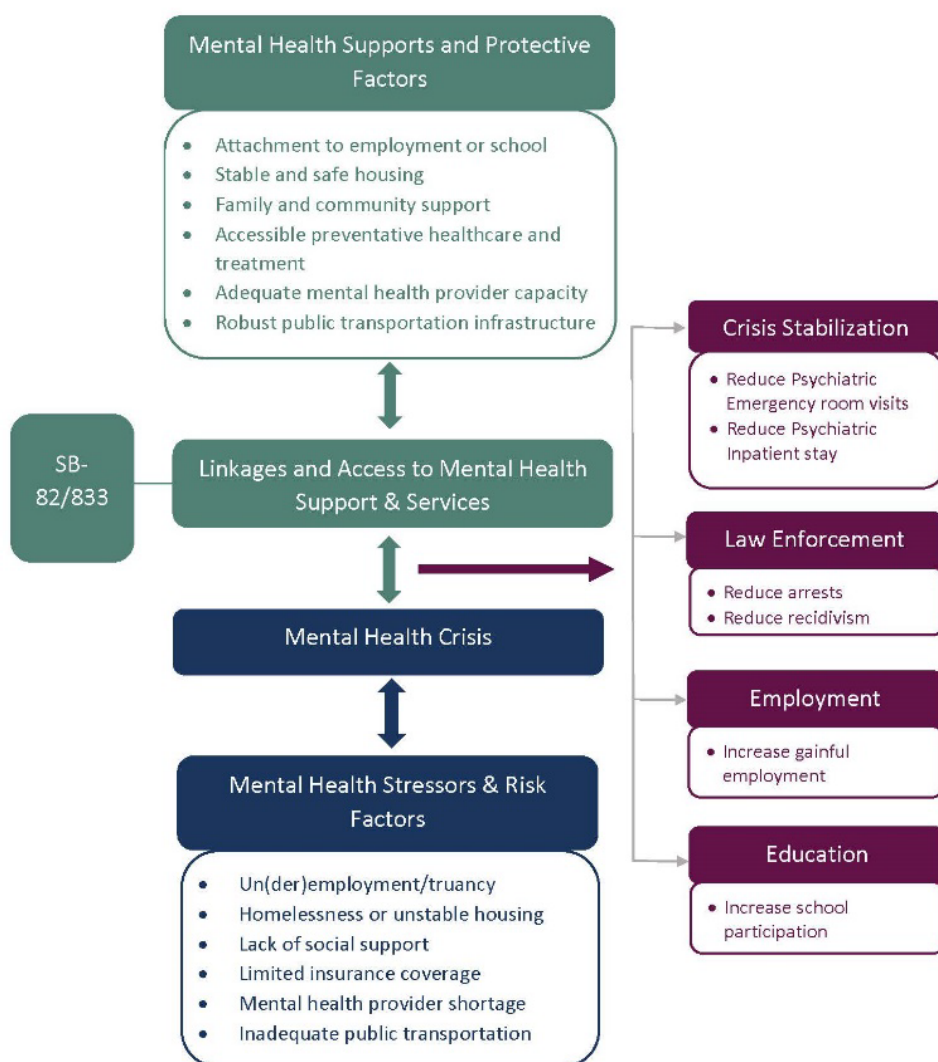
The first round of Triage funding was awarded to 24 counties (2013-17) providing services to adult and TAY clients. Twelve of the original counties applied for and received funding for Round II in addition to 8 new counties providing adult services. To provide funding for youth programs, the

2018-2022 Round II Triage grants were extended with SB-833 to fund 11 awards for children’s crisis services and four awards for school-county collaborative partnerships.

Round II of SB 82/883 (hereby referred to as Triage grant funds) consists of 30 Triage programs operating in 20 counties:

- 15 adult and TAY programs (16-24)
- 11 child and youth programs (under 18), and
- 4 school-county collaboratives (enrolled in K-12)

Figure 1. Socio-ecological Model for the Triage Summative Evaluation



The aim of Triage grant funds is to increase linkages and access to mental health services for individuals with unmet mental health needs. The extent to which these funds can successfully

meet this aim is impacted by client's personal and social context, and the systems clients must navigate to receive services. These factors impact whether, how, and to what extent individuals receive services for their unmet mental health needs. Figure 1 illustrates the social and ecological context in which Triage grant funds are administered, and how those funds interact with an individual's socio-ecological environment, thus impacting the likelihood individuals receive services. It is important to note that the figure is not comprehensive and does not capture the numerous factors outside the scope of this study that may impact client outcomes.

In the figure above, components in green represent socio-ecological factors believed to increase the likelihood that an individual will receive services to meet their unmet mental health needs. Components in blue represent socio-ecological factors believed to decrease the likelihood that an individual will receive services to meet their unmet mental health needs. In the proposed model, Triage grant funds increase linkages and access to mental health care services, thus increasing the likelihood that individuals will receive services for their unmet mental health needs. This in turn impacts client outcomes across a wide spectrum of social systems.

The previous figure depicts the overall context in which Triage funds function, but does not detail how those funds drive improved outcomes for clients. Triage programs were developed by the counties to meet the unique needs of their communities, and therefore Triage interventions vary widely in scope, service location and delivery model. Triage service types may be any combination of the following:

Service Delivery Method

Mobile Crisis: Provides a community with rapid response crises interventions. These vehicles may be staffed with clinical social workers, peer support specialists, or case managers to provide linkage to continued crisis services and deescalate on-site.

Site Based: Provides place-based services, in any capacity or location, including a clinic, school or schools, community center, or social hubs where clients congregate. Site based services are consistently in the same location or series of locations in a predictable manner.

Range of Services

Service linkage: Connecting individuals to internal or external resources to meet crisis related needs, deescalate a crisis situation, or better stabilize a client's physical health, mental health, family wellbeing, or access to safe and stable housing.

Outreach: Outreach is the process of reaching out and engaging individuals experiencing or at high risk of experiencing a crisis. This includes, but is not limited to, responding to referrals and/or embedding services in communities with increased risk of experiencing crisis such as communities comprised of racial, ethnic, gender and sexual minorities, recent immigrants, those on Medi-Cal or uninsured, individuals without safe and stable housing, and high utilizers of emergency crisis services.

Education: Education often accompanies outreach, although they are not synonymous. Educating includes providing information and/or training community members, consumers, families, and

service providers about ways to recognize and respond effectively to early signs of potentially severe and disabling mental illness.

Crisis Line: The federal Substance Abuse and Mental Health Services Administration (SAMHSA) defines a crisis line as a direct service delivered via telephone that provides a person who is experiencing distress with immediate support and/or facilitated referrals. These services help link individuals to on-going services as well as provide de-escalation at the time of the call.

Counseling: Therapeutic mental health services provided by a licensed or certified mental health clinician in an individual, family, or group setting.

Case management: Providing coordination and administrative oversight for client assessment, connection to services, discharge from care, transition planning, and overall client care trajectory.

Service Providers

Peer to peer: Peer support staff have lived experience that reflects the clients they serve. Peer support staff can engage clients in a number of roles.

Clinician: A clinician is an individual who is licensed or certified to provide therapeutic mental health services.

SB 82/833 LOGIC MODEL

The logic model below illustrates the process by which Triage services aim to reduce negative outcomes associated with unmet mental health needs. The legislative goals of SB-82/883 (described on page 3) are the driving force of the logic model, and the impetus of the intervention activities carried out by individual programs across the state. As previously discussed, individual programs can vary widely in the intervention activities they implement and the clients they serve. As such, elements listed under “Intervention Activities” are meant to capture the range of potential activities, not an expectation of activities for individual grantees or programs. The main activities of the summative evaluation are centered around the “Intervention Activities” and “Outcomes” components of the logic model. These categories will serve as the quantitative inputs and outputs of this evaluation.

The outcomes of measurement in this evaluation have been selected based on previous research documenting their relationship with unmet mental health needs. Ample evidence suggests individuals with unmet mental health needs are more likely to die by suicide³, experience incarceration⁴, be under or unemployed⁵, and/or struggle in school⁶.

In addition to examining the previously mentioned outcomes state-wide, a series of sub analyses will be conducted to examine whether overall impacts of Triage services are equitable for subgroups of Californians. Not all Californians have equal access to quality mental health services⁷, and not all Californian’s experience mental health services in similar ways. For instance, clients of color experience higher rates of stigma, and racism in the healthcare system, contributing to reduced likelihood of seeking mental health care services⁸. This reality, while concerning, is not surprising. Previous research clearly indicates disparities in mental health outcomes for vulnerable populations such as individuals living in poverty⁹, youth and older¹⁰

individuals, people of color¹¹, gender and sexual orientation minorities¹², and individuals living in rural communities¹³.

Figure 2. Triage Grant Model, Theory of Change

Legislative Goals	Intervention Activities	Outputs	Outcomes	Long Term Impact
<ul style="list-style-type: none"> ● Provide linkage to mental health services ● Expand the continuum of care for mental health crisis ● Improve consumer wellness ● Reduce unnecessary hospitalizations ● Reduce law enforcement expenditures and recidivism 	<ul style="list-style-type: none"> ● Implement early intervention services ● Staff crisis lines ● Connect clients to continuing mental health services ● Deliver mobile crisis triage services ● Provide post-crisis care (follow-up, case management) 	<ul style="list-style-type: none"> ● Increase access to mental health care providers ● Increase stability of access to mental health care providers ● Increase quality of mental health care services ● Reduce unmet mental health care needs 	<p><i>Equitably:</i></p> <ul style="list-style-type: none"> ● Increase linkage to behavioral health services ● Increase gainful employment participation ● Increase school participation ● Increase academic performance ● Reduce psychiatric ED visits ● Reduce involuntary hospitalizations ● Reduce arrests and recidivism 	<ul style="list-style-type: none"> ● Reduced inequalities in unmet mental health needs ● Greater positive client and family experiences ● Increased connection/attachment to protective systems (preventative mental health, school, employment) ● Reduction in crisis service utilization ● Reduced exposure to systems of risk (law enforcement, psychiatric ED visits)

*For students enrolled in K-12

Note: The logic model presented above has been adapted and reformatted from the formative evaluation conducted by UCD and UCLA¹⁴ to align evaluation frameworks.

THE TRIAGE SUMMATIVE EVALUATION

The following section details the Commission’s plan to evaluate the impact of Round II Triage grant funded services on client outcomes. We present the main summative evaluation questions, describe the outcome measures, evaluation design and analysis, and the limitations inherent to each approach.

DATA INFRASTRUCTURE

The ability to evaluate the impact of Triage services on client outcomes relies on multiple data sources that are typically not housed in a single state agency. Statewide evaluation can be bolstered by linking state databases to measure a wide range of outcomes¹⁵. This section provides an overview of how the Commission's data infrastructure will be harnessed to perform a summative evaluation of these programs.

In recent years, the Commission has procured data from numerous state agencies and linked those data to clients in the public mental health system.¹⁶ At the heart of the Commission's data infrastructure is the Client & Service Information (CSI) System, a client-level data system maintained by the Department of Health Care Services (DHCS). CSI contains client demographic, psychiatric diagnosis, and service utilization data on public mental health clients¹⁷. In addition, the Commission has access to the Data Collection Reporting (DCR) System that houses information about clients participating in Full Service Partnerships (FSP)¹⁸.

The Commission also successfully incorporated data from the DOJ, the CDE, EDD and birth and death records from Vital Statistics into its data infrastructure system. These data linkages allow the Commission to obtain a fuller picture of mental health service clients' experience across time, eventually across a lifespan.

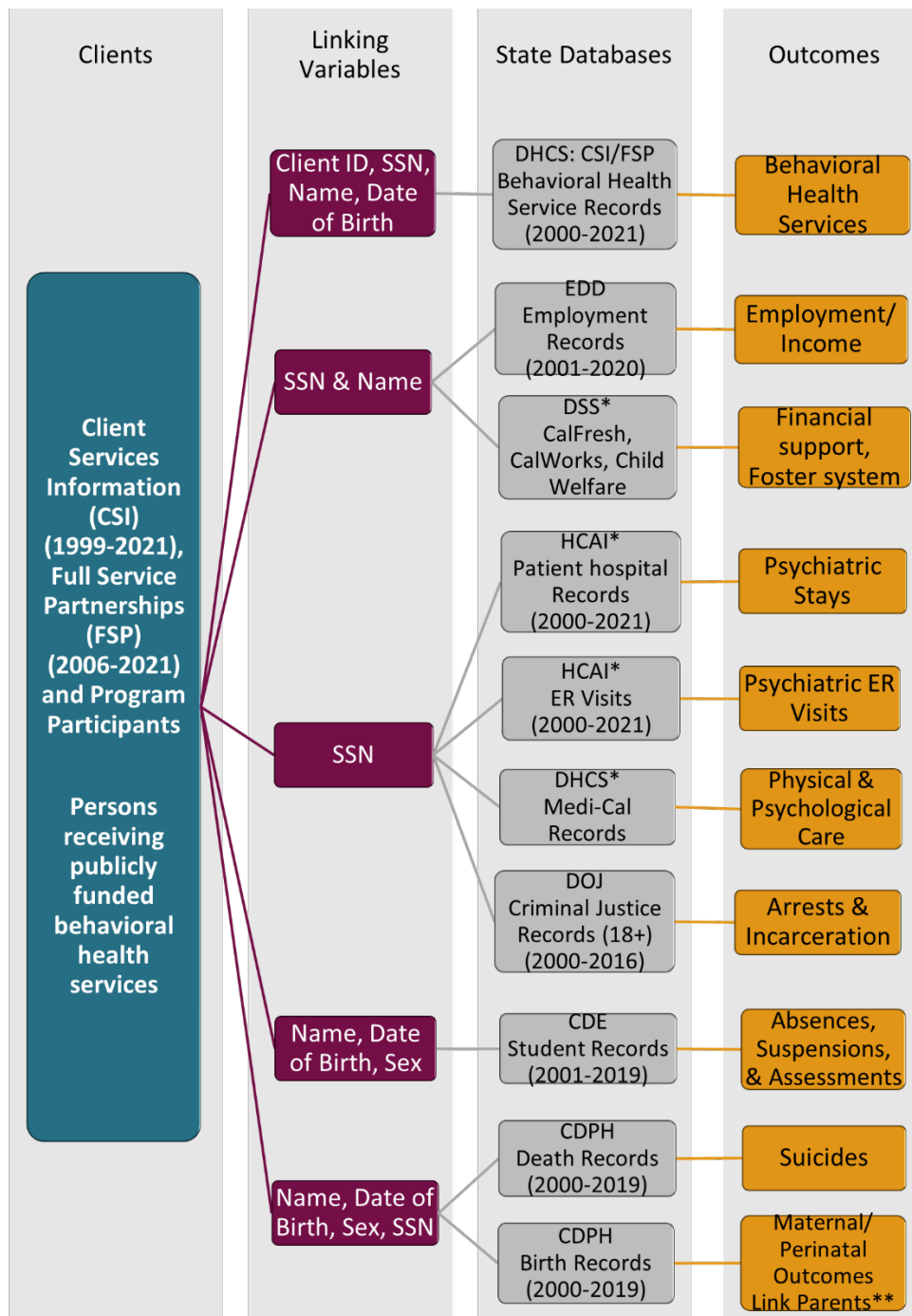
As of December 2021 the Commission is seeking data from the state's Medi-Cal program, the Department of Social Services (DSS), and the California Department of Health Care Access and Information (HCAI) (formerly the Office of Statewide Health Planning and Development). These data support analyses of additional client outcomes including emergency department visits, inpatient stays, medication provision, health outcomes, domestic violence, and out-of-home placement of children.

Figure 3 illustrates the Commission's data infrastructure, including the identifiers used to link the mental health service clients to state agency databases, and some of the outcomes that can be monitored as a result.

For the evaluation of Triage, the Commission has also established relationships with the counties who administer the which allows for the transfer of Protected Health Information/Personally Identifiable Information (PHI/PII). This will include client demographic information and clinical outcomes for Triage clients (some of whom may not be reported to the CSI).

After receiving client information from the Triage grantees, the Commission will link that information to the Client Services Information (CSI) database using a combination of client ID, SSN, name, DOB, sex, and the county in which services were provided. To obtain data for the outcomes listed in Figure 2, the Commission will link all Triage clients to statewide databases.

Figure 3. Evaluation Data Infrastructure and Outcome Sources



*Pending access

**Identifying who is a parent allows for investigations into maternal depression and other effects of childbearing upon behavioral health.

TRIAGE EVALUATION QUESTIONS

The legislative objectives of SB-82/833 address aspects of both implementation and outcomes. The diverse experiences in implementation of the Triage programs will be assessed by the formative evaluation conducted by UCLA and UCD. From the legislative objectives that address program outcomes the following evaluation questions were formulated to guide the summative evaluation.

- 1) Do Triage programs for adults, TAYs, and children:
 - a) Reduce psychiatric hospitalizations?
 - b) Reduce the rate of mental health emergency department encounters?
 - c) Reduce arrests and recidivism?
 - d) Increase participation in gainful employment?
 - e) Provide linkages to other behavioral health services and increase provision of those services?
- 2) Among behavioral health clients under the age of 16, do Triage programs:
 - a) Increase school participation?
 - b) Increase academic performance?

OUTCOME MEASURES

Table 2 describes the outcome measures that have been selected for the statewide evaluation of the Triage programs and maps them to the evaluation question they address.

Table 2. Triage Evaluation Outcome Measures

Outcome Measures	Evaluation Questions Addressed	Description
Psychiatric inpatient stays	1a.	Reduction in frequency of future inpatient psychiatric hospital stays.
Psychiatric ED visits	1b.	Reduction in frequency of future emergency department visits.
Arrests	1c.	Reduction in frequency of future arrests.
Recidivism (Arrests/Convictions)	1c.	Reduction in arrest or conviction after a previous conviction.
Employment	1d.	Increased participation in the workforce.
Behavioral Health Services	1e.	Increased service linkage and receipt of post-crisis services (e.g., crisis stabilization and crisis residential services, case management, and outpatient behavioral health services).
CSI Service Categories	1e.	Increase in frequency, intensity and/or duration of mental health services that clients receive post initial triage contact.

School participation	2a.	Reduce absenteeism, suspension, and expulsion post-receipt of crisis services.
Academic performance	2b.	Increase K-12 standardized test scores post-receipt of crisis services.

EVALUATION DESIGN AND ANALYSIS

The formative evaluation has provided a valuable typology of Triage grant funded programs, and rich insight into the heterogeneity of such programs. However, little is known about the composition of programs by county, including the range of services provided within each county, variations in populations served within each county, and the extent to which Triage clients reflect the overall CSI client population state-wide and by county. Given the structure of the data (clients nested within programs, nested within counties), it is critical to understand not just the heterogeneity of program type, but the heterogeneity of clients served before conducting an evaluation of Triage service impact.

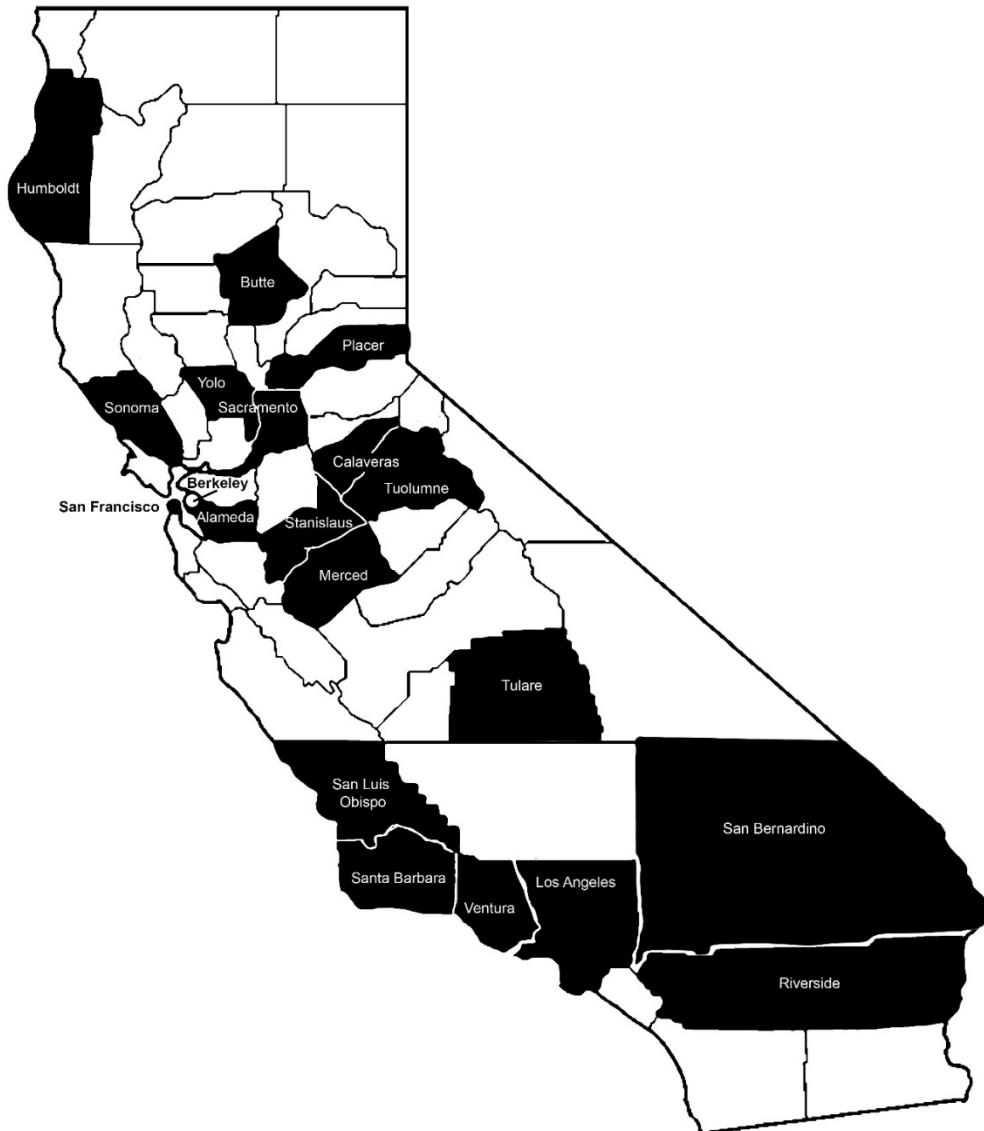
For instance, it is possible that Triage programs in County A target a specific population of clients, such as individuals in transitional housing or without safe and secure housing, whereas Triage programs in County B may not target specific populations. In this case, triage clients in County A would not be reflective of the larger CSI population in County A, nor would they be reflective of triage clients in County B. Conducting an overview of populations served within each program, before conducting an evaluation of program impacts, improves the robustness of the evaluation and ensures methodological choices are conscious of bias introduced by implementation realities and constraints.

As such, this plan outlines three main components of the summative evaluation. The first being a descriptive overview of triage programs and the clients they serve at the state-level, by county, and by program type. The second being an evaluation of the impact of triage programs overall, by county, and by program type. The third being an evaluation of the equity of triage programs for the outcomes listed above. The following sections outline the three components of the summative evaluation in greater detail.

PART 1: DESCRIPTIVE OVERVIEW

The aim of the first component of the summative evaluation is to document the scope of Triage funded grant programs across the State of California. Figure 4 maps the counties receiving triage grant funds.

Figure 4: Summative Evaluation, Participating Counties



Note: The City of Berkeley is represented in this map in addition to Alameda County as Berkeley is a separate grantee from Alameda County.

As Figure 4 illustrates, counties with Triage funded grant programs are spread across the state and vary substantially in geography, size and urbanicity. The population makeup of counties also varies for several reasons including employment opportunities, housing costs, and historical and current immigration patterns. As such, it would be reasonable to assume variance in demographic characteristics, such as race and ethnicity, age, socio-economic status, and access to mental health services across counties. A list of potential demographic variables of interest for the descriptive overview are included below.

Table 3. Variables of Interest for State, County and Program Level Descriptive Overview

Variable	Level of data collection	Data source	Definition
Size (population)	County / City	US Census	Total civilian noninstitutional population
Urbanicity	County / City	US Census	Density of development (residential and commercial)(rural vs urban)
Race / ethnicity	County	US Census	Racial and ethnic composition of county
Gender	County	US Census	Gender composition as reported by US Census
Percent living below poverty line	County	US Census	Income threshold by family size as set by US Census
High school dropout rate	County	US Census	Cohort specific high school dropout rate
Employment rate	County	Department of Labor	Total workforce (employed or seeking) divided by the working age population (civilian, noninstitutionalized)
Annual earnings	County	Bureau of Labor Statistics	Average annual pay from all sources (pre-tax)
Homelessness	County	US Census	Percent of individuals living without stable housing
Licensed psychiatric unit beds	County	California Health and Human Services	Number of psychiatric beds available county wide
Psychiatric bed occupancy rate	County	California Health and Human Services	Percent of licensed psychiatric beds currently occupied
Transportation accessibility	County / City	California Department of Transportation	An index created by MHSOAC researchers using principal component analysis
Populations served	Program level	Program data	Any restrictions to client eligibility or access based on client characteristics such as (age, diagnosis, source of initial contact etc.).
Funding Increase	Program level	Program data	Increase in funding as a percentage of previous funding
Funding per capita	Program level	Program data	Total funding per total clients served

Program type	Program level	Program data	Non-exclusionary categories of services listed on page 5
Services provided	Program level	Program data	Specified by CSI data dictionary
Race and Ethnicity	Client level	Program / CSI data	Self-reported categories from Triage program data
Age	Client level	Program / CSI data	Program demographic data / CSI data
Gender	Client level	Program / CSI data	Male, female, non-binary or other, unknown/ not reported
Transgender	Client level	Program / CSI data	Yes, no, missing/not reported
Source of contact	Client level	Program data	Program encounter data
Diagnosis	Client level	CSI data	Primary diagnosis according to CSI

The variables in Table 3 will be examined (as appropriate) at the state, county or city, and program level. Statewide means will only be reported for subgroups with 100 or more clients. County, city and program means will only be reported for subgroups with 10 or more clients.

Baseline Equivalence

In randomized control trials, individuals have an equal chance of assignment to either treatment or control. If assignment is truly random, then any observed differences between the control and treatment groups prior to treatment is random in nature, meaning it is due to chance. These types of differences can be accounted for in the statistical models used to estimate treatment impacts. There are numerous reasons why random assignment to treatment is not employed. In some cases random assignment is simply not possible, such is the case when studying the impacts of birth order on education attainment. It is not possible to randomly assign birth order. Other times, random assignment to treatment may be possible but is not ethical, or raises ethical concerns. Such would be the case when studying phenomena that inflict physical or mental harm, such as a chronic disease. Randomly assigning individuals to contract a chronic disease would violate both ethics and contemporary research standards, regardless of how valuable the potential knowledge may be to the medical community or general population. When random assignment to treatment is not possible or ethical, quasi-experimental methods are employed.

In quasi-experimental studies, two groups of individuals are compared, one that receives or experiences a particular intervention or phenomenon (treatment) and one that does not (control). Fundamental to quasi-experimental research is the creation of comparable groups. It is essential to establish the degree to which the “control”, in this case the non-triage clients, and the “treatment”, in this case the triage clients, are similar and different. If both groups are similar on key characteristics, then one can assume the two samples yield from the same population, and therefore any difference between the two groups post-intervention can be attributed to the intervention. It is unknown the extent to which triage clients will reflect the overall CSI population

across the state and within counties. It is possible, depending on how triage clients were selected, that in some counties triage clients may reflect the overall CSI population and in other counties they may not. The extent of these differences is essential to understand before identifying the appropriate methods to use for impact analysis. To assess the extent to which triage clients reflect the larger pool of CSI clients, both at the state and local levels, a baseline equivalence will be conducted.

Assessing baseline equivalence involves comparing mean difference in key client characteristics between triage clients and non-triage CSI clients. For this evaluation, differences between group means will be compared in effect size, using Hedge’s G for continuous variables and Cox’s index for dichotomous variables. If baseline equivalence is established, then methodologies for estimating program impacts can be conducted without threat of bias. However, if baseline equivalence is not established, then additional statistical procedures (propensity scores or other matching techniques) must be undertaken to reduced potential bias in program impact estimates.

Demographic variables to be used to establish baseline equivalence between triage clients and non-triage CSI clients are presented in Table 4.

Table 4. Triage Client Demographic Variables for Baseline Equivalence

Variable	Data source	Definition
Age	Program data / CSI data	Chronological age
Race and ethnicity	Program data / CSI data	Self-reported race and ethnicity categories obtained from Triage program data
Diagnosis	CSI	Primary diagnosis according to CSI (based on DSM V)
Multiple diagnoses	CSI	Indicator for whether the client has multiple diagnoses (yes/no)
Previous ER visit	CSI	Client had a previous psychiatric ER in the last 2 years
Previous conviction	Department of Justice	Client experienced a previous conviction in the last 7 years*
Health insurance	CSI	Public or private
Gender	Program / CSI data	Male, female, missing/unknown
Transgender	Program / CSI data	Yes, no, missing/not reported

Notes: Seven-year time frame for previous conviction aligns with the Fair Credit Reporting Act and California limitations on background checks.

In addition to reporting effect size difference in client characteristics by state and county, baseline equivalence will also include an examination of the correlation between client characteristics. Such analysis is important in establishing the extent to which overall variation in client characteristics may be related to a single or small group of key client characteristics.

PART 2: IMPACT EVALUATION

There are several methodologies that can be employed to infer a causal impact when random assignment is not possible. Difference-in-differences and propensity score matching/weighting are common methodologies used in studies using a quasi-causal methodology. Determining when to use either, or even both, methods depends on the constraints and potential sources of bias in the given study. The following section outlines potential sources of bias in the proposed summative evaluation and methodological approaches which can be employed to mitigate these sources of bias.

Evaluation study designs of health interventions and programs commonly use longitudinal interrupted time series (ITS) designs to compare outcomes pre- vs. post-intervention¹⁹. However, external factors other than program participation may have an impact on outcomes. One salient example is how the COVID-19 pandemic affected employment, housing, food security, and behavioral health. Mandated lockdowns, school closures, and social distancing requirements have taken a toll on Californians' mental health and required drastic modifications of mental health services²⁰. These and other unobserved or poorly measured statewide factors can introduce bias in the evaluation findings²¹. A potential methodology to control for unobserved differences over time, when random assignment is not possible, is the Difference in Differences (DID) study design. The basis of the DID design is the assumption that outcomes for separate groups of individuals can be compared before and after an intervention time point when certain criteria are met. The first criteria is that the comparison group must be similar enough to the intervention group to reasonably infer they would react similarly to the intervention at hand. The second criteria is that the comparison and intervention groups must exist in comparable environments, so that differences in outcome can be attributed to the intervention and not external factors unique to one group. If the comparison and intervention groups are similar enough and exist in comparable environments, then DID may be an appropriate quasi-experimental approach.

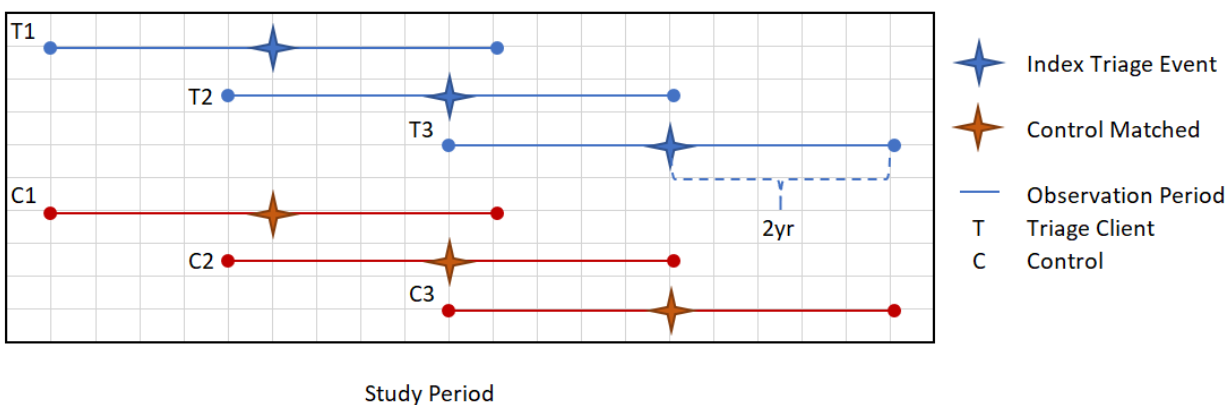
As previously mentioned, the success of the DID design relies heavily upon the existence of a comparison group that is like the intervention group²². In some cases, the comparison group and intervention group are naturally alike. Such might be the case if two clinics are in the same metropolitan area and served similar client populations. If clients for these two centers are drawn from the same larger population, then a comparison group is naturally occurring. The baseline comparison analysis discussed in the previous section assesses the extent to which Triage and non-triage clients resemble one another. If Triage and non-Triage clients do not resemble one another on key client characteristics, then a comparison group for Triage clients is not naturally occurring and will need to be created from the larger pool of non-Triage CSI clients.

To increase the likelihood that clients are experiencing similar time-related factors, such as impacts of the COVID-19 pandemic, comparison individuals will be selected from a subset of Triage eligible, CSI clients with a similar age, sex, residence, diagnoses, and date and type of initial service as their matched Triage client. Even after matching on these characteristics, it is possible that Triage and matched non-Triage clients will not be similar enough to compare without additional statistical adjustments. If differences between Triage and non-Triage clients are such that it is reasonable to infer they did not have similar likelihoods of receiving treatment, then propensity for treatment would be calculated for each Triage and non-Triage client. Triage clients

would then be matched to non-Triage clients based on three criteria: similar date of service, type of service, and propensity for treatment. *See Appendix B for methods of propensity for treatment.*

Figure 5 provides a graphical example of three Triage clients and matched to non-Triage clients. The Triage clients are labeled T1-T3 while the comparison clients are labeled C1-C3. Note that each Triage client has an initial Triage encounter date that falls somewhere within the study/evaluation period. Comparison clients will have corresponding events such as an arrest or ED visit proximate to the date of Triage. This will serve to identify clients at similar points in their mental health trajectory. Where possible, client experience will be observed for two years before and after the index date as the available data allows.

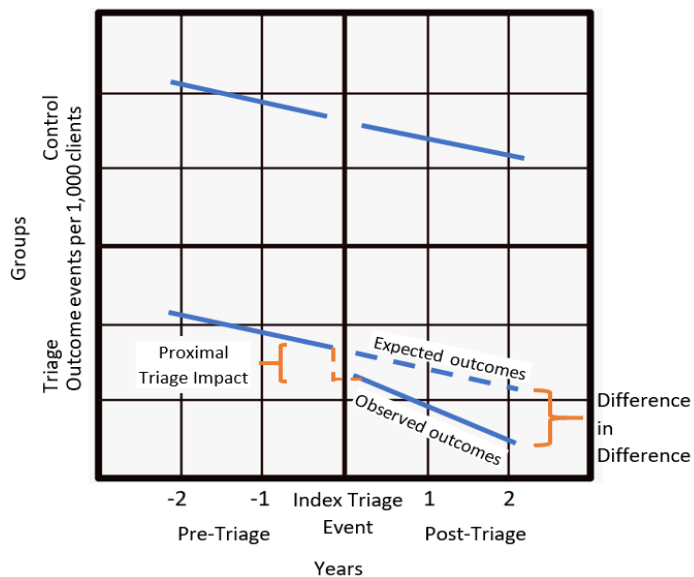
Figure 5. Case Control Longitudinal Framework for Difference-in-Different (DID) Study



The DID model explores whether the rate of change (trend) among Triage clients is significantly different from the change observed among non-Triage clients. It is important to establish that both groups have similar pre intervention trends in the outcomes of interest. For instance, to make a comparison about psychiatric emergency room admittance, both the Triage and the comparison group would need to have similar trends in psychiatric emergency room admittance prior to the onset of Triage services. This pre-intervention stage is key to establishing that groups are comparable. If pre-intervention trajectories are similar than any differences between groups in trajectory post intervention can be attributed to the intervention.

Figure 6 is a graphical representation of the DID design. The figure shows an example where the average rate of outcome decreases over time for all individuals in the study regardless of participation in the Triage program. Such might be the case if the outcome of interest were arrests, and due to the COVID-19 pandemic there was a reduction in arrests overall. However, the rate of reduction post intervention is decreasing faster among the Triage clients than among the non-Triage group. There is also a sharp reduction observed immediately after the intervention labeled proximal Triage impact. Therefore, in this example, it would be inferred that Triage services had an immediate impact on reducing arrests (the break in regression line for Triage clients), as well as a long-term impact on reducing arrests (steeper slope after intervention for Triage clients).

Figure 6. Difference-in-Difference (DiD) Statistical Analysis



To model these differences in arrests, ED visits, and inpatient stays, a multi-level regression model will be employed. Multi-level statistical methods recognize that the relationship between a dependent variable and independent variables may differ depending on a larger contextual group membership such as county. For example, the relationship between ethnicity and an outcome may differ between urban and rural regions in the state.

Differences pre and post Triage, and between the Triage and non-Triage groups will be modeled and tested for statistical significance. However, there are two tests that are of principal interest to evaluators:

- 1) The interaction between the pre-/post differences and the Triage/non-Triage differences. This interaction term tests the differences in regression line slopes between the Triage and control groups post intervention;
- 2) Differences in the actual vs. predicted post-Triage regression intercepts revealing the immediate impact of the Triage programs proximal to intervention²³.

See Appendix A. Multi-Level Difference-in-Differences Study Design for an expanded description of these statistical methods.

SUB-ANALYSES

To better understand whether groups of individuals benefit differently from Triage services, sub-analyses will be performed. Potential sub-analyses to be conducted include grouping by program type and or client population. Analyses by race and ethnicity, age, diagnoses, and other demographic characteristics will be conducted where possible to assess equity for Triage services rendered and outcomes achieved. The extent to which sub-analyses can be conducted will be determined by the prevalence of outcomes and sample size. Given the relatively complex nature of the nested data (clients within programs within counties), there may not be large enough

numbers of clients in certain demographic subgroups or program types to conduct sub-analyses on all potential outcomes. No subgroup analysis will be conducted on demographic groups containing less than 100 individuals statewide, or program groups containing less than two program sites.

Sub-analyses for Triage currently under consideration are:

1) Client age group: Whether differences in program effectiveness or equity of program outcomes differs by client age group, specifically programs serving children versus programs serving adult and TAY clients.

2) Geographic level: Whether differences in program effectiveness or equity of program outcomes differs between counties or region. Examples of geographic variables for sub-analysis include:

- Individual county
- Geographic region

3) Program level: Whether difference in program effectiveness or equity of program outcomes differs by program type, objective, or populations served. Example of program level variables for sub-analysis include:

- Service delivery method
- Service type
- Service provider type

4) Individual level: Whether differences in program effectiveness differ overall for clients of certain demographic groups. Examples of individual level variables for sub-analysis include:

- Race/ethnicity
- Age
- Diagnosis
- Whether the client has more than one primary diagnosis group
- Trauma exposure
- Place of birth
- Primary language

TIMELINE

Figure 7 provides a brief overview of the anticipated timeline for data acquisition, analysis, feedback and dissemination. Data acquisition agreements are anticipated to be received in the first half of 2021, with data collection occurring thereafter and through 2024, with the bulk of the data collected by the end of 2023. Data sets will be linked, and analysis will commence in spring of 2022. Stakeholder feedback will take place iteratively, in tandem with ongoing analysis, until the evaluation is complete. Preparation of the final summative report will commence in early 2023 and the report is scheduled for public dissemination by June 2024.

Figure 7: Triage Summative Evaluation Timeline

Triage Summative Evaluation Timeline	2021	2022	2023	2024
Data Requests				
Round I: Demographics	July- Sept			
Round II: Encounters & Diagnoses	Oct- Dec			
Round III: Final data requests		Jan- March		
School-county data requests		April - Sept		
Grant End Dates				
SB-82 & 833 grants			Jan- March	
Awards w/ no cost extension			Oct- Dec	
Los Angeles grant			Oct- Dec	
School-county collaboration grants			Oct- Dec	
Completeness of BAAs (18 counties)				
Review Completeness	July- Sept			
Obtain signed BAAs	Jan-Sept		April - June	
Adult/TAY & Child Programs (26 awards)				
Obtain Data Dictionaries	July- Sept		April - June	
Round I data due	Oct- Dec			
Round II data due	Oct- Dec	Jan- March		
Round III final data due			Jan- June	June
Check data for accuracy & completeness	July- Dec	Jan- June	July- Sept	
Link data statewide	Oct- Dec	Jan- Sept	July- Sept	
Aggregate all data for analysis	Oct-Dec	Jan- Sept	July- Sept	
Data Linkages				
Check accuracy: Adult/TAY, EDD, HCAI, & DOJ	Oct- Dec	Jan-Sept	July- Sept	
Check accuracy: Child, CDD, & HCAI	Oct- Dec	Jan- Sept	July- Sept	
Document strategies to for cross population		April- June	July- Sept	
School County Summative Evaluation (4 awards)				
Coordinate receipt of data	July- Sept	July- Sept		
Statistical Analysis				
Categorize of service codes	July- Sept			
Review statistical plans w counties	July- Dec			
Construct outcome measures		April - June		
Identify additional data sources		Jan- March		
Test model assumptions		April - June		
Construct comparison groups		April - June		
Conduct statistical analysis				Jan-June

Dissemination				
Interim Report on data transfer process		April - June		
Develop dashboard		April - June	April - June	
Prepare Round I summary reports	Oct- Dec			
Prepare reports by program (Adult/TAY, Child, School)		April- Dec	Jan- March	
Prepare briefs			Jan- Sept	
Prepare final report			Oct- Dec	Jan- June
Review feedback from stakeholders		April - June	Jan- Dec	April- June

REPORT AND DISSEMINATION

Presently, the Triage programs are presented on the MHSOAC website with descriptions of the SB-82/833 grant and the Adult/TAY, Children’s, and School-County Collaborative programs, aiming to capture the diverse range of services and clients served with Triage funds across the state.

The Triage Summative Evaluation team will yield the following reports detailing the outcomes and efficacy of the Triage programs supported by the SB 82/833 grant:

- Statewide triage summative evaluation report
- Sub-analysis and special reports (as appropriate)

Additionally, a document will be available to the public to explain how we elicited stakeholder engagement and incorporated feedback to the Triage Summative Evaluation Plan. The main deliverable from the Triage Summative Evaluation Plan is the Triage Summative Evaluation report, including the executive summary, full report, and appendices. The full report and appendices are scheduled to be posted on the MHSOAC website in 2024.

LIMITATIONS AND VALIDITY THREATS

CSI DATABASE

Utilization of services reported in the CSI is a centerpiece of this analysis and is key to creating non-Triage comparison groups. It is also the source of historical service provision for Triage and non-Triage clients alike. MHSOAC ensures all PHI/PII provided on Triage clients meet HIPAA compliance.

While the MHSOAC requires counties to document all public mental health services and clients, services that are not billable to Medi-Cal may not be fully documented. Given California’s complex system of mental health care funding this might result in under-reporting of services provided to persons not eligible for Medi-Cal, services not covered by Medi-Cal, or services provided at a site not eligible for Medi-Cal among others²⁴. Furthermore, MHSOAC analysis of the CSI data found that the completeness and timeliness of service reporting varies widely from county to county. In 2020, the Commission published a dashboard with information about CSI clients reported by

counties. It is the Commission's hope that such reporting will not only serve to better inform the public about our state's provision of mental health services, but also encourage counties to reduce CSI data reporting lags and improve turnaround time. In addition, Commission researchers reached out to county representatives to encourage reporting if there was a substantial lag. Improved reporting to CSI is essential to the validity of this evaluation.

One additional challenge of using the CSI database is the standardization of data that are transmitted by counties for Triage clients. The Research and Evaluation Division is developing metrics and variable definitions that will be used uniformly statewide in future evaluation of programs. To aid in data standardization and data collection timelines, Triage evaluators recommend establishing data requirements with counties prior to awarding funding.

RECORD LINKING

Evaluation of the impact of Triage services relies heavily on linking clients to their records in other state databases. Each database has its own set of linking criteria and contains some level of missing or incorrect data. In cases where name, sex and date of birth are the primary linking variables, common names may prevent a valid link. Linking errors come in two forms, false positives where two records are incorrectly linked and false negatives where two records that should be linked are not. Depending on the information available, strategies may be employed to estimate these errors. However, for the purposes of analyses such as program evaluation, knowing the error rates is not nearly as important as assessing the degree to which the errors are not randomly distributed²⁵.

Studies have shown that it is common for linking errors to be unevenly distributed across populations and thus introduce bias within analysis. For example, individuals who change their name upon marriage may have more linking errors than individuals who do not change their name²⁶. Younger adults are more likely to change locations than older adults. Individuals from varied ethnic groups may have names that are difficult for others to spell, or they may inconstantly use a Western standardization of their name²⁷. A comprehensive review of the literature²⁸ identified differences in linking error rates by age, sex, ethnicity, geography, socioeconomic status and health status. These and other factors introduce linking errors unevenly across sub-populations with a given study and between linking processes²⁹.

To assess the potential bias introduced by linking errors, the distributions of sub-populations among the matched records should be compared to the distributions among the records that did not match. If discrepancies are found evaluators must assess whether these discrepancies are likely due to non-random distributions of linking errors or other plausible explanations. For example, research indicates that African American and Latino individuals are disproportionately arrested as compared to their proportions of the California population³⁰. As such, more matches are expected when linking to a criminal justice database. While the impact of linking errors upon sub-populations may be difficult to estimate, the discrepancies should be reported. To mitigate threats to the validity of the analyses, weighting results by sub-population may be employed to adjust for estimated linking errors. Perhaps more importantly sensitivity analysis should be conducted to consider the effects at the extremes of these estimates.

SPILOVER EFFECTS

Spillover effects are the phenomenon of a program's impact reaching beyond the direct program participants³¹. In mental health, a common spillover effect occurs among family members, where the mental health of one family member impacts the mental health of other family members³². Further, research indicates that among college students who accessed mental health services there is a positive impact upon the propensity of their peers to access mental health services³³.

Angelucci and DiMarco³⁴ state the importance of understanding spillover effects and suggest ways to account for them in program evaluation. In the analysis of outcomes attributed to a program such as Triage, where individuals receiving treatment may share an ecological space with individuals not receiving Triage services, spillover effects are possible. For instance, if an individual receiving Triage services shared a social group with individuals not receiving Triage services, and subsequently connected those individuals to services they received through Triage staff, then the impact of Triage services would have “spilled over” to non-Triage clients. If these individuals were both in the present study, this spillover might pose a threat to the validity of the analysis. The threat arises in cases where the program clients and controls are drawn from the same location or community. For this evaluation, Triage clients and comparison groups will be residing in the same county when possible.

One way to remove this threat is to choose controls from counties that do not have a Triage program. However, there is great benefit to the face validity of the evaluation by choosing controls from the same county as the Triage clients. The clients and controls will share the ability to access the same mental health infrastructure and transportation system. They are more likely to share similar cultural and social norms. They are more likely to encounter the same law enforcement agencies and reside in communities that share similar relationships with police. Environmental and economic milieus are more likely to be similar.

For this evaluation, we assume that the benefits of choosing controls from the same county as the Triage clients outweigh the threat of spillover effects. Moreover, any potential spillover of the intervention effect would potentially underestimate the program's impact - which is a bias preferred to an unknown bias. However, given the low probability that the Triage client and comparison clients will be in the same family, have the same peers, or even live in the same community, the effect is not expected to be substantial if it exists.

While we prefer that comparison clients be from the same county as the Triage clients, this may not always be possible given the criteria for matching. It may be necessary to pool comparison clients from similar counties, especially for counties with a smaller number of clients. In these cases, the impact estimates for between same-county and out-of-county comparison clients can be compared to assess potential spillover. If there is a high likelihood of spillover effects that would impact outcomes, Triage impact estimates would be systematically different for these groups. If this is the case, sensitivity analysis could be performed adjusting for the estimated extremes of the effect.

Lastly, interconnectivity of mental health programs may introduce bias into the estimates. A county may implement other MHSA programs concurrently to Triage programs, or Triage personnel grants may be used to enhance staffing of existing programs. This opens the concern of

threats to validity from competing interventions that are not affecting all participants evenly. We will monitor whether this threat needs to be addressed by examining systematic differences in allocation of resources that do not align with the study design.

HETEROGENEOUS COUNTY MENTAL HEALTH INFRASTRUCTURES

How a Triage program is implemented and the mental health infrastructure within a county will certainly affect the outcomes being measured. To the extent possible, Triage program type and county mental health infrastructure (e.g., per capita child and adult psychiatric beds available, mental health provider shortage areas) will be accounted for in the multilevel statistical models. Thus, the between county effects will be taken into account. Moreover, since Triage clients and their comparison clients are, when at all possible, from the same county, the effects of mental health infrastructure upon the Triage clients and their comparisons within a given county are similar.

QUALITY OF LIFE

One of the legislative objectives of the SB 82/833 Triage Crisis Services grants reads: Expanding the continuum of services with early intervention and treatment options that are wellness, resiliency, recovery oriented in the least restrictive environment. While the evaluation of services and the experience of clients is clearly in the realm of the formative evaluation, it is arguable that quality of life (QoL) measures are an activity appropriate for a summative evaluation.

QoL is an important dimension of mental health beyond symptom reduction. Many questionnaire assessments such as the Mental Health Quality of Life questionnaire (MHQoL) have been developed to assess the quality of life (QoL)³⁵. After a systematic review of relevant literature and interviews persons experiencing a broad range of mental health conditions and severities, researchers identified seven domains important to quality of life³⁶: well-being and ill-being; relationships and a sense of belonging; activity; self-perception; autonomy, hope and hopelessness; and physical health. The MHQoL was developed to provide a quick and simple assessment in these domains. It is common for QoL assessments to come in the form of pre- and post-treatment questionnaires.

The Triage grants did not require a standardized QoL assessment to be administered to Triage clients. If such a requirement were to be included, it is not clear how and when it would be administered. It would not seem appropriate to burden many of the clients to answer such questions during a crisis. Furthermore, since many of the Triage programs link clients to other programs and services rather than manage a client's care over time, it is not clear when a post-treatment assessment would be administered and by whom. Given these obstacles, this summative evaluation will not attempt to assess improvements of QoL resulting from participation in a Triage program.

PARALLEL SLOPES ASSUMPTION

An assumption of the difference-in-difference model proposed for this evaluation is that the pre-intervention trends of the outcome measures among Triage clients and the comparison clients will be parallel. A failure to meet this assumption is a threat to valid interpretations of the results.

The most rudimentary way to assess parallel trends is a visual inspection, meaning the pre-intervention trends appear similar between intervention and comparison groups. There are statistical methods to check for parallel trends if a visual inspection does not produce a clear interpretation. However, these methods must also be interpreted with caution. Bilinski and Hatfield suggest an approach for interpreting parallel trends³⁷. Statistical differences in pre-intervention slopes are influenced by calculations of the probability of the result being random error and the statistical power. These calculations are comprised of effect size and sample size. Large sample sizes may cause statisticians to detect significant, but non-meaningful differences. Low statistical power may cause statisticians to fail to detect meaningful differences. It is also possible that there is a true difference in the pre-intervention slopes.

Bilinski and Hatfield introduce the non-inferiority model assumptions tests to assess the effects of non-parallel slopes. Rather than constraining the difference to be zero, they calculate a threshold for the difference by testing whether predictions made by more complex models fall within a range predicted by the simplest model. If a visual inspection of pre-intervention trends yields inconclusive results, the threshold method suggested by Bilinski and Hatfield will be considered.

ENDNOTES

¹ *The Mental Health Services Act*. (n.d.) Mental Health Services Oversight and Accountability Commission. Retrieved January 3, 2022 from <https://mhsoac.ca.gov/the-act-mhsa/>

² *Mental Health Services Act*. (2021, August 30, 2021). Department of Health Care Services. Retrieved January 3, 2022 from https://www.dhcs.ca.gov/services/MH/Pages/MH_Prop63.aspx

³ Kochanek, K. D., Murphy, S. L., Xu, J. Q., & Arias, E. (2017). Mortality in the United States, 2016. NCHS Data Brief, no 293. Hyattsville, MD: National Center for Health Statistics; Hedegaard, H, & Curtin, S.C., Warner M. (2018) Suicide mortality in the United States, 1999–2017. NCHS Data Brief, no 330. Hyattsville, MD: National Center for Health Statistics. Retrieved from <https://www.cdc.gov/nchs/products/databriefs/db330.htm>

⁴ King, A., & Sider, S. (2018). Incarceration, relationships, and belonging: Insights into the experiences of two male youth recently released from custody facilities. *Canadian Journal of Family and Youth*, 10(1), 235-258; Kondrat, D. C., Linhorst, D. M., Dirks-Linhorst, P. A., & Horning, E. (2018). An analysis of readmissions to a mental health court. *Social Work Research*, 42(3), 237-250. DOI: <http://doi.org/10.1093/swr/svy014>; Pyle, N., Flower, A., Fall, A. M., & Williams, J. (2015). Individual-level risk factors of incarcerated youth. *Remedial and Special Education*, 37(3), 172-186. DOI: <http://doi.org/10.1177/0741932515593383>; Stokes, M. L., Mccoy, K. P., Abram, K. M., Gayle R. Byck, & Teplin, L. A. (2015). Suicidal ideation and behavior in youth in the juvenile justice system: A review of the literature. *Journal of Correctional Health Care*, 21(3), 222-242. DOI: <http://doi.org/10.1177/1078345815587001>; Wylie, L.E. & Rufino, K.A. (2018). The impact of victimization and mental health symptoms on recidivism for early system-involved juvenile offenders. *Law and Human Behavior*, 42(6), 558-569. doi: 10.1037/lhb0000311.

⁵ Landsbergis, P.A., Grzywacs, J.G., & LaMontagne, A.D. (2014). Work organization, job insecurity, and occupational health disparities. *American Journal of Industrial Medicine*, 57, 495-515. <https://doi.org/10.1002/ajim.22126>; Hendriks, S. M., Spijker, J., Licht, C. M. M., F. H., De Graaf, R., Batelaan, N. M., . . . Beekman, A. T. F. (2015). Long-term work disability and absenteeism in anxiety and depressive disorders. *Journal of Affective Disorders*, 178(1), 121-130. <https://doi.org/10.1016/j.jad.2015.03.004>

⁶ Freeman, J., & Simonsen, B. (2015). Examining the impact of policy and practice interventions on high school dropout and school completion rates: A systematic review of the literature. *Review of Educational Research*, 85(2), 205–248 <https://doi.org/10.3102/0034654314554431>; Jozefowicz-Simbeni, D.M.H. (2008). An ecological and developmental perspective on dropout risk factors in early adolescence: Role of school social workers in dropout prevention efforts. *Children & Schools*, 30(1), 49–62. <https://doi.org/10.1093/cs/30.1.49>; Christle, C. A., Jolivet, K., & Nelson, C. M. (2007). School characteristics related to high school dropout rates. *Remedial and Special Education*, 28(6), 325–339.

<https://doi.org/10.1177/07419325070280060201>; Rumberger, R. W. (2011). *Dropping out: Why students drop out of high school and what can be done about it.*

<https://doi.org/10.4159/harvard.9780674063167>

⁷ Eberhart, N.K., Burnam, M.A., Seelam, R., Bogdan, O., & Breslau, J. (2019). Monitoring Californians' mental health: Population surveillance reveals gender, racial/ethnic, age, and regional disparities. *Rand Health Quarterly*, 8(3), pmiD: 31205805; Brown, P.M., Gonzalez, M.E., Sandhu, R., Conroy, S.M., Wirtz, S. Peck, C., Nunez de Ybarra, J.M. (2015). California Department of Public Health. *Economic Burden of Chronic Disease in California 2015*. Sacramento, CA; Cook, B. L., Trinh, N. H., Li, Z., Hou, S. S. Y., & Progovac, A. M. (2017). Trends in racial-ethnic disparities in access to mental health care, 2004–2012. *Psychiatric Services*, 68(1), 9-16.

⁸ Alang, S. M. (2019). Mental health care among blacks in America: Confronting racism and constructing solutions. *Health services research*, 54(2), 346-355; Misra, S., Jackson, V. W., Chong, J., Choe, K., Tay, C., Wong, J., & Yang, L. H. (2021). Systematic review of cultural aspects of stigma and mental illness among racial and ethnic minority groups in the United States: implications for interventions. *American journal of community psychology*, 68(3-4), 486-512.

⁹ Substance Abuse and Mental Health Services Administration. (2019). *Behavioral Health Barometer: United States, Volume 5: Indicators as measured through the 2017 National Survey on Drug Use and Health and the National Survey of Substance Abuse Treatment Services*. HHS Publication No. SMA–19–Baro-17-US. Rockville, MD: Substance Abuse and Mental Health Services Administration; Burns, J.K. (2015). Poverty, inequality, and a political economy of mental health. *Epidemiology and Psychiatric Sciences*, 24(2), 107-113. <https://doi.org/10.1017/S2045796015000086>; Burns, J.K., Tomita, M.A., & Kapadia, A.S. (2014). Income inequality and schizophrenia: increased schizophrenia incidence in countries with high levels of income inequality. *International Journal of Social Psychiatry*, 60, 185-196; Weich, S., Lewis, G., & Jenkins, S. P. (2001). Income inequality and the prevalence of common mental disorders in Britain. *The British Journal of Psychiatry*, 178(3), 222-227.

¹⁰ Substance Abuse and Mental Health Services Administration (2018a). *Key substance use and mental health indicators in the United States: Results from the 2017 National Survey on Drug Use and Health* (HHS Publication No. SMA 18-5068, NSUDH Series H-53). Rockville, MD: Center for Behavioral Health Statistics and Quality, Substance Abuse and Mental Health Services Administration. Retrieved from <https://www.samhsa.gov/data/sites/default/files/cbhsqreports/NSDUHFFR2017/NSDUHFFR2017.pdf>; Substance Abuse and Mental Health Services Administration. (2018). *Reports and Detailed Tables From the 2018 National Survey on Drug Use and Health (NSDUH)*. Retrieved from: <https://www.samhsa.gov/data/nsduh/reports-detailed-tables-2018-NSDUH>; Hodgkinson, S., Godoy, L., Beers, L.S., & Lewin, A. (2017). Improving mental health access for low-income children and families in the primary care setting. *Pediatrics*, 139(1). DOI: 10.1542/peds.2015-1175; Frank, J. C., Kietzman, K. G., & Palmaru, A. (2019). California's Behavioral Health Services Workforce is Inadequate for Older Adults. *Policy Brief (UCLA Center for Health Policy Research)*, 2019(1), 1-8; Solway, E., Estes, C. L., Goldberg, S., & Berry, J. (2010). Access barriers to mental health services for older adults from diverse

populations: Perspectives of leaders in mental health and aging. *Journal of aging & social policy*, 22(4), 360-378.

¹¹ California Health Care Foundation. *For Too Many, Care Not There*. March 2018. Retrieved from: <https://www.chcf.org/wp-content/uploads/2018/03/MentalHealthCalifornia2018.pdf>; California Department of Health Care Services. (2019). February 2019 Statewide Aggregate Report. Retrieved from: <https://www.dhcs.ca.gov/provgovpart/pos/Pages/Performance-Outcomes-System-Reports-and-Measures-Catalog.aspx>; Karvonen, K. L., BLEBU, B., CALTHORPE, L., FUCHS, J. D., JELLIFFE-PAWLOWSKI, L., KARASEK, D., ... & PANTELL, M. S. (2021). Racial Disparities in Emergency Mental Health Care Utilization Among Birthing People with Preterm Infants. *American journal of obstetrics & gynecology MFM*, 100546.

¹² Platt, L.F., Wolf, J.K., Scheitle, C.P. (2018). Patterns of mental health care utilization among sexual orientation minority groups. *Journal of Homosexuality*, 65(2), 135-153. DOI: [10.1080/00918369.2017.1311552](https://doi.org/10.1080/00918369.2017.1311552); Russell, S.T., & Fish, J.N. (2016). Mental health in lesbian, gay, bisexual, and transgender (LGBT) youth. *Annu Rev Clin Psychol*, 12, 465-487. Pmid: 26772206/; Martin-Storey, A. (2016). Gender, sexuality, and gender nonconformity: Understanding variation in functioning. *Child Dev Perspect*, 10(4), 257-262; Yuan, N., Duran, B., Walters, K., Pearson, C., & Evans-Campbell, T. (2014). Alcohol misuse and associations with childhood maltreatment and out-of-home placement among urban two-spirit American Indian and Alaska Native people. *International Journal of Environmental Research and Public Health*, 11(10), 10461-10479; Frazer, S.M., & Pruden, H. (2010). Reclaiming our voices: Two Spirit health and human service needs in New York State. NYS DOH AIDS Institute: Albany, NY.

¹³ Ivey-Stephenson, A.Z., Crosby, A.E., Jack, S.P., Haileyesus, T., Kresnow-Sedacca, M. (2017). Suicide trends among and within urbanization levels by sex, race/ethnicity, age group, and mechanism of death — United States, 2001–2015. *MMWR Surveill Summ*, 66(No. SS-18), 1–16. DOI: <http://dx.doi.org/10.15585/mmwr.ss6618a1external icon>; Frank, J.C., Kietzman, K.G., & Palimaru, A. (2019 Jan). *California's behavioral health services workforce is inadequate for older adults*. UCLA Center for Health Policy Research. <http://healthpolicy.ucla.edu/publications/Documents/PDF/2019/workforce-policybrief-jan2019.pdf>; Kietzman, K.G., Dupuy, D., Damron-Rodriguez, J., Palimaru, A., del Pino, H.E., & Frank, J.C. (2018). Older Californians and the Mental Health Services Act: Is an older adult system of care supported?. UCLA Center for Health Policy Research. <http://healthpolicy.ucla.edu/publications/search/pages/detail.aspx?PubID=1708>

¹⁴ Padovani, A. et al. (2020, February 12). Deliverable 3: Draft Summative Evaluation Plan. In *Mental Health Services Oversight and Accountability Commission*. Retrieved from https://mhsoac.ca.gov/wp-content/uploads/UC-Davis-Triage-Adult-TAY-Draft-Summative-Evaluation-Plan_Full-Document.pdf; Miranda, J. (2020, May 15). Deliverable 3: Draft Summative Evaluation Plan. In *Mental Health Services Oversight and Accountability Commission*. Retrieved from

https://mhsoac.ca.gov/sites/default/files/2020-05/UCLA%20Triage%20Child%20and%20Schools%20Draft%20Summative%20Evaluation%20Plan_Full%20Document_1.pdf

¹⁵ Hser, Y. I., & Evans, E. (2008). Cross-system data linkage for treatment outcome evaluation: Lessons learned from the California Treatment Outcome Project. *Evaluation and Program Planning* 31(2), 125-135. <https://doi.org/10.1016/j.evalprogplan.2008.02.003>

¹⁶ Kanowitz, S. (2021, June 9, 2021). *California connects datasets to show how mental health services can reduce arrests*. Retrieved December 1, 2021 from <https://gcn.com/articles/2021/06/09/california-mental-health-arrest-data.aspx>

¹⁷ *Updates to the Client and Services Information system*. (2016). Health and Human Services Agency Retrieved from https://www.dhcs.ca.gov/services/MH/Documents/MHSUDS_InformationNotice_16-014.pdf

¹⁸ *MHSA Transparency Glossary*. (n.d.). Mental Health Services Oversight & Accountability Commission. Retrieved December 8, 2021 from <https://mhsoac.ca.gov/transparency-suite/mhsoa-transparency-glossary/>

¹⁹ Dimitrov, D. M., & Rumrill, P. D., Jr. (2003). Pretest-posttest designs and measurement of change. *Work*, 20(2), 159-165. <https://www.ncbi.nlm.nih.gov/pubmed/12671209>

²⁰ Byrne, A., Barber, R., & Lim, C. H. (2021). Impact of the COVID-19 pandemic: A mental health service perspective. *Progress in Neurology and Psychiatry*, 25(2), 27-33b. <https://wchh.onlinelibrary.wiley.com/doi/10.1002/pnp.708>

²¹ Naci, H., & Soumerai, S. B. (2016). History bias, study design, and the unfulfilled promise of Pay-for-Performance policies in health care. *Preventing chronic disease*, 13, E82. <https://doi.org/10.5888/pcd13.160133>

²² Dimitrov, D. M., & Rumrill Jr, P. D. (2003). Pretest-posttest designs and measurement of change. *Work*, 20(2), 159-165.

²³ Mascha, E. J., & Sessler, D. I. (2019). Segmented regression and Difference-in-Difference methods: Assessing the impact of systemic changes in health care. *Anesthesia & Analgesia*, 129(2), 618-633. <https://doi.org/10.1213/ANE.0000000000004153>; Zhou, H., Taber, C., Arcona, S., & Li, Y. (2016). Difference-in-Differences method in comparative effectiveness research: Utility with unbalanced groups. *Applied health economics and health policy*, 14(4), 419-429. <https://doi.org/10.1007/s40258-016-0249-y>

²⁴ Arnquist, S., & Harbage, P. (2013). A complex case: Public mental health delivery and financing in California. *California HealthCare Foundation*.

²⁵ Randall, S., Brown, A., Boyd, J., Schnell, R., Borgs, C., & Ferrante, A. (2018). Sociodemographic differences in linkage error: An examination of four large-scale datasets. *BMC Health Services Research*

18(1), 678. <https://doi.org/10.1186/s12913-018-3495-x>

²⁶ Randall, S., Brown, A., Boyd, J., Schnell, R., Borgs, C., & Ferrante, A. (2018). Sociodemographic differences in linkage error: An examination of four large-scale datasets. *BMC Health Services Research* 18(1), 678. <https://doi.org/10.1186/s12913-018-3495-x>; Boyd, J. H., Ferrante, A. M., Irvine, K., Smith, M., Moore, E., Brown, A., & Randall, S. M. (2017). Understanding the origins of record linkage errors and how they affect research outcomes. *Australian and New Zealand journal of public health*, 41(2), 215-215. <https://doi.org/10.1111/1753-6405.12597>

²⁷ Miller, E. A., McCarty, F. A., & Parker, J. D. (2017). Racial and ethnic differences in a linkage with the National Death Index. *Ethnicity & disease*, 27(2), 77-84. <https://doi.org/10.18865/ed.27.2.77>

²⁸ Bohensky, M. A., Jolley, D., Sundararajan, V., Evans, S., Pilcher, D. V., Scott, I., & Brand, C. A. (2010). Data linkage: A powerful research tool with potential problems. *BMC health services research*, 10(1). <https://doi.org/10.1186/1472-6963-10-346>

²⁹ Randall, S., Brown, A., Boyd, J., Schnell, R., Borgs, C., & Ferrante, A. (2018). Sociodemographic differences in linkage error: An examination of four large-scale datasets. *BMC Health Services Research* 18(1), 678. <https://doi.org/10.1186/s12913-018-3495-x>

³⁰ Lofstrom, M., Goss, J., Hayes, J., & Martin, B. (2019). *Racial disparities in California arrests*. Public Policy Institute of California. Retrieved December 2, 2021 from <https://www.ppic.org/publication/racial-disparities-in-california-arrests/>

³¹ Angelucci, M., & Di Maro, V. (2010). Program evaluation and spillover effects: Impact-evaluation guidelines [Technical notes]. 41. Retrieved December 4, 2021, from <https://publications.iadb.org/publications/english/document/Program-Evaluation-and-Spillover-Effects.pdf>

³² Fletcher, J. (2009). All in the family: Mental health spillover effects between working spouses. *The BE journal of economic analysis & policy*, 9(1). <https://doi.org/10.2202/1935-1682.1967>

³³ Golberstein, E., Eisenberg, D., & Downs, M. F. (2016). Spillover effects in health service use: Evidence from mental health care using first-year college housing assignments. *Health economics*, 25(1), 40-55. <https://doi.org/10.1002/hec.3120>

³⁴ Angelucci, M., & Di Maro, V. (2016). Programme evaluation and spillover effects. *Journal of Development Effectiveness*, 8(1), 22-43.

³⁵ van Krugten, F. C. W., Busschbach, J. J. V., Versteegh, M. M., Hakkaart-van Roijen, L., & Brouwer, W. B. F. (2021). The Mental Health Quality of Life Questionnaire (MHQoL): Development and first psychometric evaluation of a new measure to assess quality of life in people with mental health problems. *Quality of Life Research*, 1-11. <https://doi.org/10.1007/s11136-021-02935-w>

³⁶ Connell, J., O'Cathain, A., & Brazier, J. (2014). Measuring quality of life in mental health: Are we asking the right questions? *Social science & medicine*, 120, 12-20.

<https://doi.org/10.1016/j.socscimed.2014.08.026>

³⁷ Bilinski, A., & Hatfield, L. A. (2019). Nothing to see here? Non-inferiority approaches to parallel trends and other model assumptions. *Cornell University* <https://arxiv.org/abs/1805.03273>