

INTRODUCTION	3
ABOUT THE DATA	3
Data Sources	3
Geographic Units of Analysis	4
Exclusion Criteria	4
Suppression Rules for Client Data	5
METHODS AND ANALYSIS	5
Software	5
MH Client Sample	5
Geocoding	6
Defining Complex Variables	6
LIMITATIONS	7

INTRODUCTION

The purpose of the Community Experience Partnership (CEP) Service Planning Tool (SPT) Technical Report is to summarize the data, methods, and limitations of the SPT. The SPT is a custom application designed by UC San Diego in collaboration with the County of San Diego Behavioral Health Services (BHS). The goals of the tool are to help ensure service provision is informed by data, based in cultural and regional considerations, and focused on communities that may be at greatest risk for unmet behavioral health need. The application uses data to help planners at BHS identify areas in San Diego County where target populations are likely to be highly concentrated. Once areas are identified, users may explore community profiles and download custom reports to summarize the social, economic, and demographic profiles of the selected regions.

ABOUT THE DATA

Data Sources

The SPT includes data for nearly 100 target populations drawn from four data sources (Table 1).

TABLE 1. DATA SOURCES

Data Type	About
Behavioral Health Equity Index (BHEI)	The BHEI is a composite index that allows users to explore differences in the root causes (also known as social determinants) of behavioral health across neighborhoods in San Diego County. A higher BHEI rank means communities may not have access to the resources and services that promote behavioral health. This is a useful tool to identify areas that may serve as priority zones for equity work and service enhancements. To learn more about the BHEI visit “About the BHEI” in the “More Information” dropdown tab in the SPT.
Community Data: Population Overall	Community data are sourced from the U.S. Census Bureau's American Community Survey (ACS) and describe the sociodemographic characteristics of San Diego County communities. These estimates are representative of the population across the entire 60-month data collection period (2018-2022).
Community Data: Population Living in Poverty (below 100% Federal Poverty Level)	Community data for the population living below 100% of the Federal Poverty Level (FPL) are also taken from the ACS. This sample aligns more closely with the population eligible for BHS services (i.e., the BHS target population).
Mental Health (MH) Service Data: All Clients, FY 2022-23	Mental Health (MH) Service data are sourced from Cerner Community Behavioral Health (CCBH) for FY 2022-23.

Data notes for each target population are available in the SPT’s “Learning Resources and Data Notes” section (under the “More Information” dropdown). These notes include technical definitions and sources for each target population.

Geographic Units of Analysis

Data in the SPT are presented at three geographic levels (Table 2).

TABLE 2. GEOGRAPHIC UNITS OF ANALYSIS

Geographies	Description
Health and Human Services Agency (HHSA) Regions	There are six County-approved regions designed to assist with organizing and managing health service delivery.
Subregional Areas (SRAs)	San Diego County has 41 SRAs. These represent cities or communities (e.g., Chula Vista, Escondido). SRAs are aggregated from census tracts and defined by the San Diego Association of Governments (SANDAG).
Zip Codes (ZCTAs)	There are over 100 zip codes represented in the service planning tool. Zip code data are based on Zip Code Tabulation Areas (ZCTAs) defined by the U.S. Census Bureau. Because zip codes are not a reliable geography, the Census Bureau assigns ZCTAs using the most frequently occurring zip code in a small census area. In most cases zip codes are the same as ZCTAs, however, some addresses are assigned to a ZCTA that differs from their zip code. Additionally, zip codes representing very few addresses may not be represented by a ZCTA. This accounts for the holes in the ZCTA maps.

Exclusion Criteria

- **BHEI:** 19 of the 116 San Diego County ZCTAs (16.4%) failed to meet BHEI inclusion criteria and are excluded from the index. More information about the BHEI exclusion criteria, including a list of excluded geographies, is available in the SPT (see: “About the BHEI” under the “More Information” dropdown).
- **Community Data from the U.S. Census Bureau:** Geographies with an estimated total population of zero (no residents) were removed, resulting in the exclusion of zip codes 92096, 92132, and 92182.
- **MH Service Data:** MH clients were excluded from the SPT if any of the following four conditions applied:
 1. Residential address could not be geocoded: Even addresses with errors (e.g. misspelled street names, missing street numbers, incorrect zip codes, etc.) can usually be geocoded by ArcGIS. In some instances, however, addresses contain errors that prevent them from being successfully converted into geographic coordinates. Clients that cannot be geolocated to a census tract in San Diego County are excluded.
 2. Resided outside San Diego County: Clients with a residential address outside of San Diego County are excluded.

3. Resided in an institutional setting: Institutional settings include non-community environments like nursing homes, skilled nursing facilities, residential treatment facilities, jails, and detention centers. Clients residing in institutional settings have different needs and receive different services than clients residing in the community. Including institutional clients in local counts would inflate the service utilization estimate for these areas. For example, including clients from Edgemoor Skilled Nursing Facility in service data for Santee would overestimate service utilization estimates for community residents and could bias decisions about service allocation.
4. Residential address recorded as the BHS Billing Unit (PO Box 129153): Historically service providers were instructed to use "PO Box 129153", the BHS Billing Unit address, for clients experiencing homelessness. More recent guidelines instruct providers to record the zip code where clients experiencing homelessness are most likely to sleep or spend time. While clients with a zip code reflecting the areas they frequent can be geocoded and retained in the data, the P.O. Box address is assumed to be unrelated to a client's actual location. Clients with this address are excluded to avoid misrepresenting their geographic distribution.

Suppression Rules for Client Data

To protect client confidentiality, data are suppressed when the numerator is a non-zero number less than 11 or the denominator is a non-zero number less than 20. When needed, complimentary suppression is applied to meet these minimum suppression thresholds. To improve data presentation and interpretability the following racial/ethnic categories were suppressed for all regions: non-Hispanic (NH) Native American, NH Other, and NH Unknown. San Diego County totals include data that have been suppressed for small client counts but exclude clients that do not meet the inclusion criteria.

METHODS AND ANALYSIS

Software

FY 2022-23 client datasets were compiled from Cerner Community Behavioral Health (CCBH) extracts. Census data were downloaded and prepared using the R Package Tidycensus (v1.3.3; Walker & Herman, 2023). The SPT application was developed using the R package Shiny and is hosted on [ShinyApps.io](https://shinyapps.io). Clients' residential addresses, as recorded in CCBH, were geocoded using ArcGIS Pro (v3.2.1) and Python (v3.12.2). All other data were cleaned and manipulated using R (v2023.06.01).

MH Client Sample

Two client extracts were analyzed: the Children, Youth & Family (CYF; n=11,919) dataset included all clients served by providers in the child/youth BHS system. The Adult Older Adult dataset (AOA; n=43,155) included all clients served by providers in the AOA system.

As defined in the exclusion criteria, clients were removed who could not be geocoded to a census tract in San Diego County (excluded from CYF: n=74; 0.6%; excluded from AOA: n=1,213; 2.8%). Duplicate clients appearing in both the AOA and CYF samples were removed (n=928), and a combined dataset of unique clients was generated (n = 52,859). From this data set, clients were then excluded who were residing in an institutional setting

(n=4,804; 9.1%) or who had an address listed as PO BOX 129153 (n=2,069; 3.9%). This resulted in a final sample size of 45,986.

Geocoding

To protect client confidentiality, geocoding was conducted offline using Street Map Premium, California. Census Tract and ZCTA polygons were obtained through the Census API and converted to the ESPG 4326 coordinate reference system (CRS) to match the geocoded output. Addresses were spatially joined to their respective Census Tract and ZCTA area polygons. Clients were assigned to SRA and HHSA geographies using Census Tract to region crosswalks produced by SANDAG. Clients without a valid street address who had a valid zip code were geocoded to their zip code centroid.

Defining Complex Variables

The MH client data included numerous complex variables that were defined and operationalized by the SPT development team. This included the following client outcomes:

- **Experiencing Homelessness:** The SPT identified clients as "experiencing homelessness" based on two criteria:
 - Clients who identified their current living situation as "homeless".
 - Clients with residential addresses that matched homeless shelters or those containing terms indicating the clients were unhoused (e.g., "homeless," "transient," "unsheltered").
- **Residing in institutional settings:** Client addresses were joined to facility, shelter, BHS Contractor, and BHS provider address lists to identify the name of any facilities where clients resided. Facilities were then coded as institutional versus community settings by the SPT development team. Institutional facilities included nursing homes, skilled nursing facilities, acute psychiatric hospitals, residential treatment programs, jails, and detention centers. Unmatched addresses where fewer than six clients resided were coded as "in the community."
- **Received emergency or crisis treatment:** Clients were categorized as receiving emergency/crisis services if they received any of the following types of service: urgent outpatient, crisis stabilization, crisis residential, Emergency Screening Units, or Psychiatric Emergency Response Team services.
- **Received outpatient treatment:** Clients were categorized as receiving outpatient services if they received any of the following types of service: Assertive Community Treatment, Case Management, Fee for Service, Outpatient, Prevention Services, Juvenile Forensic Services, Wraparound, and Therapeutic Behavioral Services (TBS).
- **Clients with a co-occurring mental health and substance use disorder:** Co-occurring substance use was operationally defined as a dual diagnosis (concomitant substance use diagnosis) and/or involvement with Substance Use Disorder (SUD) services and/or endorsement of substance abuse-related items on the Behavioral Health Assessment form.

Technical definitions for all other variables included in the SPT can be found in the SPT's "Learning Resources and Data Notes" section (under the "More Information" dropdown).

LIMITATIONS

The SPT is a valuable tool for informing service planning, but it should be used in conjunction with other sources of information and informed by local knowledge. Limitations of the tool are listed below:

- The ACS is designed to estimate population characteristics, not to provide exact population counts. These data are survey estimates that are subject to sampling error and may be unreliable, especially for small geographic areas or small population groups. Small differences between groups may not be statistically significant and should be interpreted with caution. Users are encouraged to review sample sizes and reliability notes when working with the tool and to aggregate multiple areas of focus when the numerator or denominators are small to improve reliability. Additionally, a reliability warning is included on the the map labels and data tables when estimates fail to meet the reliability threshold (e.g., the coefficient of variation for the denominator is greater than 40% and the size of the denominator is less than 50). This warning helps to prevent basing decisions about service provision on unstable estimates.
- Geographic aggregation may conceal areas of need. For instance, a zip code may have a relatively low BHEI ranking despite including some neighborhoods at very high risk for behavioral health inequity. This highlights the importance of engaging local stakeholders and community experts when using the tool.
- Populations of focus are independent of one another. This means it would be accurate for a user to say an area has both a high concentration of unemployed residents and a high concentration of Spanish-speaking residents, but not that an area has a high concentration of unemployed residents who are also Spanish-speaking.
- Data are not available for all indicators considered important causes of behavioral health equity in San Diego County (e.g., indicators of systemic racism, crime, etc.).
- When no valid street addresses are available, zip code centroids are used to geolocate clients. This may not accurately reflect where clients spend their time, potentially leading to misclassification within specific geographic units.
- Inaccuracies in data collection or recording can lead to errors and potentially biased results. Examples include outdated or incorrect addresses or miscoding institutional settings as “in the community”.
- MH clients that did not meet the inclusion criteria have been excluded from the SPT. The data are not representative of the entire MH client population.
- Although the SPT uses the most recent data available, time lags do exist. Updated extracts may reveal important changes.