Methodology for the Development of
Potential Opioid Service Area Maps

James Cajka MA – RTI International

Rob Lyerla PhD MGIS – Center for Behavioral Health Statistics and Quality, Substance
Abuse and Mental Health Services Administration, U.S. Department of Health and Human
Services

April 2017

Executive Overview

This document summarizes the methodology used by the Center for Behavioral Health Statistics
and Quality (CBHSQ) to create a set of state maps showing specific areas (census tracts) that are
potentially underserved by medication-assisted treatment (MAT) facilities. The tracts take into
consideration the location of existing facilities as derived by two surveys that comprise the Behavioral
Health Treatment Services Locator (BHTSL): the National Survey of Substance Abuse Treatment Services
(N-SSATS) and the National Mental Health Services Survey (N-MHSS). The tracts also take into
consideration the estimated number of illicit drug users, which are derived from the National Survey on
Drug Use and Health (NSDUH).

The need for MAT services was imputed from the percentage of past month illicit drug
userreported by NSDUH by substate region (SSR). SSRs are formed by grouping counties and census
tracts together. Because the total number of people in an SSR is known from simple population counts,
the total number of illicit drug users can be estimated by multiplying the percentage of illicit drug users from NSDUH by the total number of people from the census. This total number of IDUs was disaggregated to the census tract level using the proportion of people living in poverty by census tract, as compared with the entire SSR. In this case, poverty is assumed to relate directly to illicit drug use. The number of people living in poverty is known from the American Community Survey (ACS). This number was summed for each SSR, and therefore the percentage that each tract contributed could be calculated by dividing the number of people in poverty in the tract by the total in the SSR. This percentage was then multiplied by the number of IDUs in the SSR to generate the estimated number of IDUs in each tract.

Each existing MAT facility was plotted according to its latitude and longitude as specified in the downloaded BHTSL files. From these points, a circular catchment of radius $n$ was generated using a formula created at CBHSQ. The formula is a distance decay function that uses the population density of the county that the facility falls within to create a radius between 0.5 miles and 1.0 mile. The radius was used to generate a circular catchment for each facility. The catchments were used to overlay on the census tracts. If the catchment overlapped a given tract, that tract was designated as having access to an MAT facility. The tracts that were not overlapped by a catchment were then processed to identify those without access to an MAT facility that had the greatest need for these services.

Nonoverlapped tracts were divided into quintiles based on the estimated number of IDUs on a county-by-county basis. Those tracts in the top quintile for each county were combined at the state level and then divided into quintiles again. These tracts are displayed on a state map according to their final quintile.
The maps provide a general idea of which tracts in a given state are not currently served by an existing MAT facility and potentially contain a high number of IDUs. Note that the areas identified are specific to this methodology and these data inputs.

Purpose

The purpose of the medication-assisted treatment (MAT) facility maps is to identify areas on a state-by-state basis that may be potentially underserved by existing treatment facilities. The maps are created with a methodology that seeks to include the highest potential need areas from individual counties so that county-level stakeholders are also informed. The maps are meant to be used as a tool for policy makers to determine potentially underserved areas—not as a definitive representation of these areas.

Data Inputs

When performing a facility service area analysis, there are generally two types of data inputs required: supply and demand. The supply in this case was the universe of existing MAT facilities. The demand was the estimated number of illicit drug users that potentially need services from a treatment facility. To determine the number of illicit drug users, two specific datasets were required. The first was the percentage of past month illicit drug users from the National Survey on Drug Use and Health (NSDUH), aggregated at the substate region (SSR) level. The second was the number of people in poverty at the census tract level, taken from the American Community Survey (ACS), which was used to guide the disaggregation of the number of IDUs from the SSR level to the tract level.

Early versions of the mapping algorithm identified potentially underserved areas that were known to be associated with high real estate costs and high land value. A calculated poverty rate was developed to
address access to behavioral health care that could be linked with income levels. Using this poverty variable helped to identify areas where people could least afford treatment and would therefore benefit most from access to a MAT facility. The census tract level of geography was chosen for several reasons including the availability of data, ease of processing, and ability to see the results on a state-level map.

Existing MAT facilities were derived from two sources: the National Survey of Substance Abuse Treatment Services (N-SSATS) and the National Mental Health Services Survey (N-MHSS). These two datasets were downloaded from SAMHSA’s Behavioral Health Treatment Services Locator (BHTSL) website (https://findtreatment.samhsa.gov/locator). Seven types of care were selected and downloaded: 1) Methadone maintenance; 2) Methadone maintenance for predetermined time; 3) Methadone detoxification; 4) Buprenorphine maintenance; 5) Buprenorphine maintenance for predetermined time; 6) Buprenorphine detoxification; and 7) Buprenorphine used in treatment. Once downloaded, they were combined into a single file that included the latitude and longitude for each facility. The latitude and longitude were used to create a spatial data layer for use in a geographic information system (GIS). The facility locations were assumed to be correct—no validation was performed.

Each facility fell within a county boundary. Part of the methodology requires the county population density as an input to a distance-decay function. The map boundary file was obtained from https://www.census.gov/geo/maps-data/data/tiger-line.html. The associated population data were obtained from the 2010 U.S. Decennial Summary File 1, which is a 100 percent count of the population rather than an estimate. The total population table (P1) containing the “D001” variable was downloaded from http://factfinder.census.gov/faces/nav/jsf/pages/download_center.xhtml.

The NSDUH SSR data that contained the IDU data were acquired from http://archive.samhsa.gov/data/NSDUH/Substate2k12/ShapeFile/NSDUHSubstateShapefile2012.htm.
The estimates were based on a small area estimation procedure that combined data from the 2010-2012 NSDUHs with other variables to produce model-based estimates. The variable of interest in the NSDUH data file was “SUMMON,” which provided a percentage of “past month use of illicit drugs (12 and older).” This variable includes past month users of marijuana/hashish, cocaine (including crack), heroin, hallucinogens, inhalants, or prescription psychotherapeutics used nonmedically.

Figure 1. Illicit Drug Use in the Past Month among People Aged 12 or Older, by Substate Region: Percentages, Annual Averages Based on 2012, 2013, and 2014 NSDUHs

The population in poverty at the census tract level was obtained from Table B17001 in the ACS (http://factfinder.census.gov/faces/nav/jsf/pages/download_center.xhtml). The table used a 5-year estimate spanning 2010 to 2014. The ACS assigns people an income-to-poverty rate using federal guidelines on poverty thresholds. The number of people in poverty is reported by age cohort, divided into males and females.

Methodology

Converting Percentage of IDUs to Count of IDUs at the SSR Level

The prevalence of past month illicit drug use is given as a percentage in the NSDUH SSR shapefile. To calculate the total number of illicit drug users within an SSR, a spatial join of census tracts to SSRs was performed using ArcGIS. This associated each census tract with the SSR it fell within. The total number of people aged 12 or older was summed by tract to find the sum of all people within an SSR:

\[ Pop_{12+SSR} = \sum_{1}^{N} Pop_{12+Tract} \]
The number of illicit drug users (IDUs) per SSR was then calculated by simply multiplying the number of people aged 12 or older by the illicit drug use rate:

\[ IDU_{\text{Count}_{SSR}} = Pop_{12\text{plus}_{SSR}} \times IDU_{\text{Rate}_{SSR}} \]

Calculating the Poverty Rate at the Tract Level

Using the poverty counts from ACS Table B17001, the total number of people in poverty between ages 12 and 64 was calculated by adding the number of people in each age cohort. The total number of people aged 12 to 64 within the SSR was calculated similarly to the number of people aged 12 or older:

\[ InPov_{1264}_{SSR} = \sum_{1}^{N} InPov_{1264_{Tract}} \]

Each tract’s poverty weight was calculated as follows:

\[ Poverty\ Weight_{Tract} = \left( \frac{InPov_{1264_{Tract}}}{InPov_{1264_{SSR}}} \right) \]

Allocation of Illicit Drug Users to Tracts

Ultimately, the goal was to estimate the number of IDUs residing in each tract. This is the number that would drive the creation of quintiles and therefore be used to select the areas that were most potentially underserved. To calculate the number of IDUs in an SSR to be distributed to an individual tract, the overall count of IDUs in the SSR was multiplied by the poverty weight for that tract:
\[ IDU_{Count,\text{Tract}} = (IDU_{Count,SSR} \times \text{Poverty Weight}_{\text{Tract}}) \]

This number was stored as an attribute in the census tract geospatial data layer.

**Creation of Catchment Areas**

Catchment areas were estimated using the Potential Attractiveness and Ability to Serve (PAAS) equation. This equation was created by staff at the Center for Behavioral Health Statistics and Quality as a way to create a set of radii that were evaluated for optimal distance. The PAAS equation uses a 0.25-mile minimum radius and a 1.00-mile maximum radius. The PAAS equation can be written symbolically as follows, where \( g \) is geography and \( t \) is time:

\[
PAAS_{gt} = \left[ 1 - \left( \text{Decay}_{gt} \times \text{Adjustment}_{gt} \right) \times \text{Population}_{gt} \right]
\]

Thus, the equation estimates a facility’s plausible catchment area using a particular set of geographies at a particular point in time. The method used tracts, counties, and NSDUH SSR geographies. For computation, the PAAS equation assigns each MAT location’s catchment areas as follows:

\[
1 - \left( [0.1 \times 1.2^{(m-0.01)\times100}] \times \{1.1 + [(m - 0.2) \times -0.77]\} \right) \times \left( \frac{\text{PopDensity}_{gt}}{69,470} \right)
\]

The value of the expression is calculated for values of \( m \), starting with 0.01 miles and proceeding by increments of 0.01 until a value of 1.00 is reached. The PAAS equation computes the radius of a circular catchment area. Radii vary in size as a function of population density only. Thus, all facilities within a given county will have the same catchment radius. The variation in radii size is a function of population density that operates with an exponential distance decay. The value 69,470 represents the highest population density in 2010 (i.e., maximum population density on the county level in 2010). The value of
PAAS is calculated for each increment of $m$. The optimal value is the one closest to 0 without becoming negative. If the optimal value of $m$ is found to be less than 0.25 miles, then $m$ was set to 0.25 miles because this was the minimum radius allowed. If the value of $m$ reached 1.00 mile and the PAAS value was still positive, then $m$ was set to 1.00 because this was the maximum radius allowed.

A buffer was created in ArcGIS for each MAT facility using the radius specified by the optimal PAAS value. All the buffers were created in a single GIS layer (Figure 1).

Figure 1. Census Tracts and MAT Facility Catchments

This buffer layer was then intersected with the census tract layer that contained the estimated number of IDUs as an attribute. All tracts intersecting the buffer layer were removed from the selection set because these tracts were designated as tracts that had access to a MAT facility. The remaining (nonintersecting) tracts were saved to a geospatial data layer (Figure 2).
These tracts were then exported to a text file with each record containing the tract ID, the county ID, the state ID, and the estimated number of IDUs. The text file was then imported into the open-source statistical package R for the creation of quintiles. Tracts were divided into quintiles by county based on the number of IDUs. This ensured that the highest potentially underserved tracts were identified for each county. Counties with less than five eligible tracts were excluded from the quintile assignment because there were insufficient observations. Those tracts that were in the top quintile in their county were extracted, and a second quintile assignment was done at the state level. The final dataset was a set of tracts (approximately 20 percent of all tracts) with a state-level quintile designation of 1 to 5. The state quintile designation was the value depicted on the state maps.

Map Depiction and Symbology

One potential MAT facility service gap area map was created for each state, including the District of Columbia, Alaska, and Hawaii. Each map was designed to fit on 8 ½” x 11” landscape format.
paper. The scale varies from map to map depending on the geographic size of the state. Quintiles are state specific, using state specific population estimates and state specific prevalence estimates, meaning that relative importance within a state is unique to that state and not comparable across states. The projection for all maps is North America Albers Equal Area Conic, with the exceptions of Alaska (NAD 1983 Alaska Albers) and Hawaii (Hawaii Albers Equal Area Conic).

The tracts in the top quintile in their county were symbolized as being in one of quintiles 1 to 5 at the state level. Those tracts in the highest quintile (i.e., had the most IDUs) were symbolized using the darkest red color. Lighter shades of red were used as the state quintile designation moved toward 1. All remaining tracts were symbolized using the lightest shade of red (Figure 4). These are labeled as a “Low Potential Area” in the legend and include tracts that fall into three categories: (1) tracts that are overlapped by an existing treatment facility’s catchment, (2) tracts that did not fall into the top quintile at the county level, and (3) tracts in counties that were not divided into quintiles because the number of observations was less than five.
Figure 3. Example of State-Level Map

Contextual geographic data are also included in the maps. County boundaries are shown in blue, and major roads are shown in green. Several other layers could be shown, such as major cities or the locations of existing MAT treatment facilities. Ultimately, these data were left off the maps because they tended to obscure the tracts that were in the top quintiles, and showing these tracts was the main point of the maps.

Limitations

The potential service area maps were created using a specific set of data inputs, a specific level of geography, and several specific assumptions. Therefore, they reflect one way of identifying potential service area gaps. Initially, when summing the number of illicit drug users’s at the substate level, it is
assumed that the prevalence of illicit drug use is uniform across all tracts within a given SSR. The count of illicit drug users at the tract level represent users of any illicit drugs, not just users of opioids who may benefit from proximity to MAT facilities. In the future, opioid use will be available as a result of a redesign of the NSDUH survey. The universe of facilities offering services is dynamic over time, and thus the list provided in the BHTSL data may not be complete and may contain errors in location (latitude and longitude). Additionally, the universe of service providers does not include individual providers who may have MAT services, but decline to have their practice identified. Further, the analysis could easily be done at a more granular level of census geography (i.e., block groups), which would result in a more exact designation of service/nonservice areas. The drawback of using block groups is that they would be smaller and harder to see on the maps. Finally, the PAAS formula restricts the range of acceptable treatment facility catchment radii to be between 0.25 miles and 1.00 mile. This could be expanded to cover a larger distance or substituted with driving times that would arguably provide a better measure.

Acknowledgments

The Center for Behavioral Health Statistics and Quality would like to thank Dr. Carlos Siordia for his work on the development of the methodology and the creation of the Potential Attractiveness and Ability to Serve (PAAS) measure.