

Protocol for the Implementation Mechanisms of Stepping Up (I.M. Stepping Up) Study, a Longitudinal Natural Experiment Examining Implementation Mechanisms to Improve Mental Health and Substance Use Services for Justice-involved Individuals in 950 Stepping Up and Comparison Counties Across the U.S.

Jennifer E. Johnson (✉ Jennifer.Johnson@hc.msu.edu)

Michigan State University <https://orcid.org/0000-0003-2540-7213>

Jill Viglione

University of Central Florida

Niloofer Ramezani

George Mason University

Alison Cuellar

George Mason University

Maji Hailemariam

Michigan State University

Rochelle Rosen

Miriam Hospital

Alex Breno

George Mason University

Faye Taxman

George Mason University

Study protocol

Keywords: Implementation, jail, mental health, substance use, mechanism, community

Posted Date: January 27th, 2021

DOI: <https://doi.org/10.21203/rs.3.rs-154028/v1>

License:  This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Version of Record: A version of this preprint was published at Implementation Science on March 29th, 2021. See the published version at <https://doi.org/10.1186/s13012-021-01095-2>.

Abstract

Background. The criminal justice system is the largest provider of mental health services in the U.S. Many jurisdictions are interested in reducing the use of the justice system for mental health problems. The national Stepping Up Initiative works with counties to reduce the number of individuals with mental illness in jails and to improve access to mental health services in the community. The Implementation Mechanisms of Stepping Up (I.M. Stepping Up) Study leverages a large natural experiment created by comparing Stepping Up counties to matched comparison counties over time to examine implementation mechanisms and outcomes.

Methods. The study will survey 475 Stepping Up counties and 475 matched comparison counties at three waves: baseline, 18 months, and 36 months. Surveys will be sent to up to four respondents per county including administrators of jail, probation, community mental health services, and community substance use treatment services (3,800 total respondents). Implementation target mechanisms include: (1) use of and capacity for performance monitoring; (2) use and functioning of interagency teams; (3) common goals and mission across agencies; and (4) system integration (i.e., building an integrated system of care rather than adding one program or training). We will examine whether Stepping Up counties show a faster rate of improvement in hypothesized target mechanisms between Wave 1 and subsequent assessments (i.e., Waves 2 and 3) than do comparison counties (primary). We will also examine whether Stepping Up counties show a faster increases in implementation outcomes (number of justice-involved clients receiving behavioral health services, number of behavioral health evidence-based practices and policies [EBPPs] available to justice-involved individuals, and resources for behavioral health EBPP for justice-involved individuals). We will evaluate whether engagement of hypothesized mechanisms explains differences in implementation outcomes. Finally, we will characterize implementation processes and critical incidents using survey responses and qualitative interviews.

Discussion. There are few rigorous, prospective studies examining implementation mechanisms and their relationship with behavioral health implementation outcomes in justice and associated community behavioral health settings. There is also limited understanding of implementation mechanisms that occur across systems with multiple goals. This study will elucidate target mechanisms in multi-goal, multi-agency systems.

Contributions To The Literature

1. This paper describes the methods for the first prospective national study of efforts to help justice systems and community behavioral health providers work together better to keep individuals with mental illness out of jail.
2. This protocol paper provides an example of how to study mechanisms of policy implementation.
3. This protocol paper provides an example of how to study implementation mechanisms when randomization is not possible.

Background

More than 10 million adults are arrested and enter the United States criminal justice (CJ) system, including pretrial detention, jail, probation, and parole, each year.¹ Rates of current mental health (56%) and substance use (66%) disorders are elevated among justice-involved populations,² who disproportionately lack education, experience victimization and homelessness, and have poor employment skills, complicating care and increasing morbidity and mortality.³⁻¹³ The CJ system is a low-resource setting charged with serving people who face complex behavioral and physical health issues, with inadequate access to health care.

Counties (which often operate jails, probation and/or parole services, community mental health, and substance use treatment systems) care for the vast majority of justice-involved individuals. Recognizing that counties are overwhelmed by overuse of incarceration for those with mental health and substance use disorders and by the lack of health services for justice-involved individuals in the community, the National Association of Counties (NACo), the Council of State Governments (CSG) Justice Policy Center, and the American Psychiatric Association Foundation (APAF) developed the Stepping Up Initiative.¹⁴ The goal of Stepping Up is to reduce the number of individuals with serious mental illness in jails. Since it began in 2015, Stepping Up has registered more than 500 counties to address access to mental health services for current or potentially justice-involved individuals. The study team did not design Stepping Up. However, Stepping Up provides an opportune longitudinal natural experiment in which to compare implementation strategies, mechanisms, and outcomes across time in Stepping Up and matched comparison counties.

This study uses a broad definition of mental health and substance use evidence-based practices and policies (EBPPs) including: 1) behavioral and pharmacological treatments; 2) procedures such as assessment and care coordination; 3) practices such as diversion, supported employment, and problem-solving courts; and 4) policies including laws or regulations, judicial decrees, and agency guidelines. Our definition of “justice-involved individuals” is taken from the Sequential Intercept Model,¹⁵ and includes the full range of relevant county and local justice involvement (i.e., 911 calls, local law enforcement contact, pretrial jail detention, court appearances, specialty courts, jail sentences, probation, and parole).

Implementation Mechanisms

Identifying implementation mechanisms can inform future development of more powerful implementation interventions. Lewis et al. define implementation mechanisms as “processes or events through which an implementation strategy operates to effect desired implementation outcomes” (p.3).¹⁶ Mechanisms examined in this study come from the Criminal Justice Interagency Implementation Model (CJ-IIM; Figure 1),¹⁷ which describes the need for cooperation of multiple constituencies for CJ to pursue public health goals (such as improving mental health and substance use services). Implementation mechanisms described in the CJ-IIM address the cross-contextual-layer cooperation needed for behavioral health implementation in CJ systems. This study proposes to examine four of the CJ-IIM

mechanisms (performance monitoring, interagency work groups, goal and mission setting across agencies, and system integration) as mechanisms of action for Stepping Up (Figure 2).

Use of and capacity for performance monitoring. Performance monitoring refers to the use of data to assess population needs, understand how systems work, identify desired outcomes, and monitor progress. Stepping Up recommends that counties develop capacity for ongoing monitoring of four measures: 1) number of people with mental illnesses who are booked into jail, 2) average length of stay in jail, 3) percent linked to community-based treatment after release from jail; and 4) return to jail rate (i.e., recidivism). Doing so often requires infrastructure development since justice and behavioral health agencies lack common identifiers to track clients across systems and few jurisdictions have electronic health record systems. This study will assess what CJ and behavioral health data and performance measures exist and are being used. The study will also evaluate the relationship between performance monitoring (mechanism) and implementation outcomes.

Use and functioning of interagency teams (i.e., mental health, substance use, jail, probation, county government). Moving away from a focus on one agency to the larger system requires interagency teams to exist and to work together. The failure of agencies to function as a team limits their ability to address larger, systemwide issues. The Stepping Up initiative explicitly promotes an interagency stakeholder model. The first step prescribed by the initiative is for the county to convene a team of jail, probation, community mental health, substance use treatment, and elected representatives to work together on the remaining steps. This study will assess whether Stepping Up efforts improve the use and functioning of interagency teams (mechanism), and whether use and functioning of interagency teams improves implementation outcomes.

Having common goals and mission across agencies. EBPPs are more likely to be implemented in systems with clear, visible goals,^{18,19} where the EBPPs are consistent with the agency's mission.^{17,20} CJ agencies such as jails, police, and prosecutors have primary public safety goals, with some secondary public health responsibilities. For behavioral health EBPP implementation, cooperating systems need to incorporate behavioral health EBPPs as a vital ingredient for achieving the primary mission of public safety or agree that public safety can be better achieved through improved behavioral health.¹⁷ For the systems to work together to address common goals, external stakeholders must support this mission shift. To join Stepping Up, counties pass a resolution providing strong, visible support for health and justice agencies to deviate from their siloed missions to create integrated goals. This study will assess whether Stepping Up strategies result in cross-agency goals (mechanism) and will assess how shared goals and mission impact implementation outcomes.

System integration. The CJ-IIM hypothesizes that implementation efforts will be most effective as they broaden engagement and ownership across agencies to develop a county-wide system of care, rather than adding single programs or trainings. Many agencies are involved in CJ behavioral healthcare (e.g., community mental health centers, substance use treatment agencies, police, courts, jail, probation, prison, and parole). Service linkage among them is often inadequate or non-existent.^{21,22} Previous analyses of

CJ reform efforts suggest that the tendency is to implement a new program rather than working toward a cohesive system of care.^{23,24} The study will examine whether Stepping Up counties are more successful at creating integrated systems of care (mechanism), and whether doing so improves implementation outcomes.

Implementation Mechanisms of Stepping Up (I.M. Stepping Up) Study Aims

We will compare Stepping Up counties to matched comparison counties to examine effectiveness of Stepping Up and to identify mechanisms by which Stepping Up counties improve county-level implementation outcomes (see Table 1). We will use surveys and qualitative interviews to compare 475 Stepping Up and 475 paired target counties at 3 waves: study baseline, 18 months, and 36 months. County- and agency-level data will be collected from 4 criminal justice and behavioral health administrators in each county (3,800 total respondents at each wave). Stepping Up and comparison counties will be compared on rates of change in engagement of target mechanisms (primary) and rates of change in implementation outcomes (secondary). Specific aims are to:

1. **Examine target mechanisms.** We will examine whether: (a) Stepping Up counties show a faster rate of improvement in hypothesized target mechanisms (use of/capacity for performance monitoring, use and functioning of interagency teams, common goals across agencies, and system integration) between Wave 1 and subsequent assessments (i.e., Waves 2 and 3) than do comparison counties; and (b) whether engagement of these mechanisms explains any differences found in implementation outcomes.
2. **Examine implementation outcomes.** Examine whether Stepping Up counties show faster rate of improvement in implementation outcomes (number of justice-involved clients receiving behavioral health services, number of behavioral health EBPP available to justice-involved individuals, and resources for behavioral health EBPP for justice involved individuals) between Wave 1 and subsequent and assessments than do comparison counties.
3. **Characterize implementation processes and critical incidents** occurring in Stepping Up and comparison counties. We will (a) use qualitative data to triangulate quantitative findings and enrich an understanding of how the target mechanisms produce outcomes; and (b) explore which naturally occurring implementation strategies used in 950 counties lead to better implementation outcomes.

Innovation

Reform rhetoric is common, but there are few rigorous, prospective studies examining implementation mechanisms and their relationship with behavioral health implementation outcomes in CJ settings. The national Stepping Up Initiative, which seeks to change the interface between behavioral health services and the justice system, provides a large natural experiment and a unique research opportunity.

The proposed project also has novel implications for implementation science. Most implementation change process studies have been conducted within a single organization (e.g., health). There is limited

understanding of implementation mechanisms that occur within or across systems with multiple goals, where some goals may be secondary or foreign.¹⁷ This study will inform other implementation efforts by identifying the potential target mechanisms in complex, multi-agency systems.¹⁶

Methods

County matching and baseline covariate controls increase the rigor for this longitudinal natural experiment, which is described consistent with STROBE guidelines²⁵ (see Supplemental Materials). The study will measure target mechanisms and key outcomes in 475 Stepping Up counties and 475 matched counties at three waves: study baseline, 18 months, and 36 months. Since counties self-select to be in Stepping Up, we cannot randomize counties to the Stepping Up Initiative. Instead, we use a case-matched design using the Centers for Disease Control and Prevention (CDC) peer counties methodology to identify non-Stepping Up comparison counties with similar size, demographics, and health, economic, and justice indicators.²⁶⁻²⁸ Stepping Up began in 2015 and has been registering counties over time. Therefore, at study baseline, the 475 Stepping Up counties had been participating in Stepping Up from one to five years. Wave 1 (i.e., baseline) values and months since each county began Stepping Up will serve as covariates. Therefore, analyses examine differences between groups in within-county rates of change between study waves controlling for Wave 1. Our quantitative survey results will be augmented with qualitative interviews with 60 counties at each wave to enrich our understanding of how the implementation mechanisms work.

The Stepping Up Initiative

The goal of Stepping Up is to reduce the number of individuals with mental illness in jails and to improve access to community mental health services for currently or potentially justice-involved individuals. To join Stepping Up,¹⁴ counties pass a resolution to address behavioral health disorders (i.e., reduce unnecessary use of jail, increase access to behavioral health services) using a broad, locally-adaptable six-step action plan: 1) convene a diverse team of leaders; 2) collect and review data on individuals in the justice system; 3) examine treatment and service capacity; 4) develop a plan with measurable outcomes; 5) implement the plan; and, 6) track ongoing progress with data. Table 1 and Figure 2 show how we mapped Stepping Up's six steps onto hypothesized CJ-IIM mechanisms. Within the six broad steps/strategies outlined by the national Stepping Up Initiative, counties can use their own approaches based on their own priorities.

Stepping Up outlines steps to help counties become more data-driven. Leaders of agencies within each county are asked to agree upon a mutual definition of terms such as "mental illness," "connection to community-based care," and "recidivism." Stepping Up then encourages counties to identify and use a universal, validated mental health screening instrument for new intakes into the jail and other agencies. The screening tool identifies individuals in need of a full clinical assessment. Stepping Up offers toolkits to help counties examine how to capture screening and assessment results electronically and engage in information-sharing agreements. Counties are encouraged to track data on four key measures to assess

impact of their efforts over time: 1) number of people with mental illnesses who are booked into jail, 2) average length of stay, 3) percentage of people connected to community-based treatment after release from jail, and 4) rate of return to jail. With this data infrastructure in place, counties can assess the effects of their efforts to address patient needs (e.g., substance use, mental illness, family discord), continuity of care, and reach of services. As Stepping Up counties work to iteratively improve services, they can track progress, focus county leaders on key outcome measures, and make the budgetary and programmatic case for needed resources. This study will assess counties' progress on these steps and will categorize local approaches.

Survey Methods

Stepping Up counties. When this project was submitted for funding, there were 475 counties designated as Stepping Up counties. Although new counties continue to join Stepping Up, for feasibility, this study will assess the 475 counties originally proposed.

Comparison counties. A peer group of 475 matched comparison counties was created using a county grouping methodology developed by CDC²⁶⁻²⁸ and updated by our team.²⁹ Matching variables were drawn from three primary data sources: (1) Vera Institute's incarceration trends database³⁰ for county pretrial and jail populations; (2) County Health Rankings & Roadmaps³¹ for health, economic, social, and demographic information; and (3) the Uniform Crime Report^{32,33} for crime and policing data.

Counties were nested within states and clustered on health and social indicators. Therefore, a hierarchical matching approach needed to be developed to accommodate state- and county-level covariates. The study principal investigators initially chose 34 of the most potentially relevant variables from the datasets based on expert knowledge. These variables included demographic factors (e.g., median household income, unemployment, total population, high school graduation rate, percent African American, percent Hispanic), inequality indicators (e.g., income inequality, residential segregation), health factors (e.g., poor mental health days, poor physical health days, HIV incidence), healthcare (e.g., mental health providers per capita, primary care physicians per capita, percent of drug treatment paid by Medicaid), crime, and criminal justice (e.g., per capita numbers of police officers, jail population, jail pretrial population, juvenile criminal cases). Based on random forest models and team feedback, these variables were reduced to 29 total variables: 22 predictors and 7 variables reflecting jail populations and mental health providers in the area (factors central to Stepping Up activities). Third, shrinkage based variable selection techniques were applied to select variables that best predicted jail population per capita, pretrial population per capita, and per capita rate of mental health providers, without collinearity. Next, logistic models (which included both predictors and dependent variables in the previous models) were fitted to define variable weights and estimate the likelihood of each county classifying as a Stepping Up or non-Stepping Up county. Using these weights, matching scores were calculated for each county and used in an algorithm to find the best control matched county for each Stepping Up county among potential comparison counties within the same state.

The final variables used for county matching scores included per capita rates of: mental health providers, daily jail population, daily jail pretrial population, primary care providers, police, licensed psychologists, and community mental health centers. Final variables also included average number of physically unhealthy days (of 30), high school graduation and income inequality rates; total county healthcare expenses, percent African American population, percent Hispanic population, percent drug treatment paid by Medicaid, county population, and an indicator reflecting presence of a medical school in the county. In states where the number of Stepping Up counties was higher than the number of potential comparison counties, state location, Medicaid expansion status, and justice/mental health policy were used to pair comparable states and then algorithmically match at the county level. If a county from the comparison group joins Stepping Up during the first year of the study, we find a new matching county. If this occurs after the first year of the study, the pair will be removed from analyses.

Survey respondents. The overall sample is 475 Stepping Up counties and 475 comparison counties. In each county we will survey the administrators of community mental health, jail, probation, and community substance use treatment agencies (i.e., up to 4 respondents per county and ~3,800 total; see Table 2). These respondents were selected because the jail and probation systems have the majority of individuals under justice control in a county, and mental health and substance use treatment administrators are responsible for the provision of behavioral health services for justice-involved individuals in the community.

To compile the respondent list, the research team developed a database of all Stepping Up and matched comparison counties. NACO-CSG-APAF provided a list of county contacts for Stepping Up counties. We contacted these individuals to provide the appropriate contact information for jail, probation, mental health, and substance use administrators in their county. We also conducted web-based searching. For comparison counties, we identified county-level experts through web-based searching. When contact information was not publicly available, we called individual agencies to identify the correct respondents. We also engaged in a snowball technique, in which we contacted experts already identified for assistance in identifying other possible respondents in their county.

Survey administration. The web-based survey is administered using Qualtrics. Using a procedure described by Dillman,³⁴ respondents receive an introductory email that includes a NACO-CSG-APAF endorsement letter of support as well as key information to collect prior to beginning the survey (e.g., budget and staffing data). One week following administration of the introductory e-mail, an invitation to participate in the survey is sent using Qualtrics. The research team sends follow-up e-mails once a week for three weeks following the initial invite with a reminder to participate. If the survey is not completed by the end of week four, research team members make follow-up phone calls. During these calls, the research team provides multiple options for the respondents to complete the survey, including completion of the survey via telephone and receipt of a paper copy. Following the phone call, the research team continues to follow-up with respondents biweekly. Given the current context (i.e., COVID-19), we anticipate encouraging survey participation for six months before closing the survey. We will also provide county-

specific feedback reports on county-level CJ behavioral health indicators as an incentive for study participation.

Survey validation. We use existing, validated measures where possible. When we needed to tailor items to CJ or to mental health, we used Cook³⁵ strategies for item development by testing new items using cognitive interviews.³⁶ Interviews covered question comprehension, decision processes, and response options.

Ten cognitive interviews were conducted in May 2020 with volunteers from Stepping up counties representing jail, probation/parole, community mental health, and community substance use treatment. Interviews were conducted via videoconference. Team members met to iteratively review interview results and revise the survey. At these meetings each interviewer presented the responses and reflections from their interviews. Volunteer comments and interviewers/notetaker feedback, along with expert review by team members, were used to revise the survey. Changes were made to simplify and clarify survey questions and to remove redundancies. The amended survey was again reviewed by all team members.

Measures

All measures will be collected at all 3 time points. All respondents will receive the same assessments. We refer to measures as “agency-level measures” when analyses of these measures will account for nesting within counties, but the primary focus is on agencies. We refer to measures as “agencies nested within counties” when we used nested analyses and our primary focus is on the county level. We refer to measures as “county-level” measures if they produce a single value for the county to be analyzed at the county level.

Descriptors, predictors, and moderators (agency-level). A series of measures will be used to describe the inner context of each agency. Type of agency will be characterized using the National Criminal Justice Treatment Practices (NCJTP) survey About Your Organization scale.³⁷ Staffing, including type, number, and turnover, will be measured using adapted NCJTP Staffing scales.³⁷ Organizational Culture Support for Innovations (a proposed moderator) will be assessed using an adapted version of the NCJTP Assess Your Organizational Culture scale.³⁷

Aim 1: Target mechanisms. Use of and capacity for performance monitoring (agencies nested within counties). We created a Performance Monitoring measure which provides one point for each of the following: (1) whether counties are *able* to measure the 4 Stepping Up core metrics (number of mentally ill people who are booked into jail, average length of jail stay, percent who are connected to community-based treatment upon release from jail, and rate of return to jail) (up to 4 points), (2) each metric they regularly report (up to 4 points), and (3) each metric used for ongoing decision making (up to 4 points) (up to 12 points total). A secondary measure to capture performance monitoring identifies 7 kinds of decisions (e.g., budget preparation, medicine supply), and asks whether they were guided by each of the 4 Stepping Up core metrics (0=no, 1=yes), for up to 28 points possible. This measure was adapted from the

Routine Decision-Making scale of the Performance of Routine Information System Management (PRISM) Toolkit.³⁸

Use and functioning of interagency teams (agencies nested within counties). To examine the activities and functioning of interagency teams, we integrated the NCJTP Relationship Assessment Inventory³⁷ with additional items based on the goals and priorities of Stepping Up. This integrated scale contains 18 items such as “we share general information about populations in need of treatment services” (0=no, 1=yes) with one point assigned for each collaborative activity across the other 3 agencies. The total score (up to 54) reflects joint activities among agencies.

Common goals and mission across agencies (agencies nested within counties). The primary measure (an adapted NCJTP Goals/Mission scale)³⁷ assesses each respondent’s perception of the degree to which their agency goals and overall county goals align. Respondents are given a list of goals (e.g., public safety/protection, provide mental health services) and are asked to rank them according to: (1) their agency’s priorities, and (2) county priorities. A kappa score reflects the degree of consistency between the two lists. The secondary measure, a county-level measure, will be agreement (kappa) among respondents within counties of ratings of the importance of providing mental health treatment services for justice-involved individuals in jail and in the community (separately) on a scale of 1 (unimportant) to 10 (important).

System integration (agencies nested within counties) is a dichotomization of the NCJTP Relationship Assessment Inventory³⁷ total score (i.e., excluding the additional items). Counties with scores of 18 or more are considered to have achieved “system integration.” A secondary (county-level) measure will reflect the degree to which each of 12 listed behavioral health screening and assessment instruments are used by and/or shared among multiple responding agencies within a county. For each of the 12 instruments listed, counties will receive a score (0 = no agencies use the same instrument, 1 = two agencies use the same instrument, 2 = three agencies use the same instrument, and 3 = all four agencies use the same instruments).

Aim 2: Implementation outcomes. Number of justice-involved adult clients receiving behavioral health services (agencies nested within counties; primary). After defining “justice-involved” and asking whether each of the EBPP described below is available in the county, we ask respondents how many justice-involved individuals received any mental health service and how many received any substance use service in their agencies in the past year.

Number of behavioral health EBPP available to justice-involved individuals (county level). Mental health EBPPs were taken from treatment recommendations for justice-involved individuals³⁹⁻⁴⁵ and from community standards for treatment of serious mental illness posttraumatic stress disorder, borderline personality disorder, suicide thoughts or behaviors, anxiety, insomnia, and pain.⁴⁶⁻⁵² Substance use EBPPs were taken from the U.S. National Institute on Drug Abuse’s consensus list.⁵³ Using the EBPP list describe above, we ask whether each EBPP is available to justice-involved individuals in the county. If any

of the respondents answers “yes,” we count that EBPP as being available to justice-involved individuals in the county.

Resources for behavioral health EBPP for justice-involved individuals (agency-level). Respondents will be asked to report whether their agency has experienced an increase (+1), no change (0), or decrease (-1) in funding from the prior year in 13 different areas (e.g., “screening and assessment”). We will cluster these 13 areas using factor analysis and then create total scores for each factor, which will serve as primary outcome/s. Initially, we planned to assess the total dollar amount of resources devoted to behavioral health services for justice-involved individuals, but found that most agencies could not report this number. Secondary measures relate to capacity and training: (1) the proportion of staff in clinical roles, (2) how many staff participated in behavioral health-related training in the past year, and (3) number of staff hired minus the number who left in the prior year. Lastly, we will use the NCJTSP’s Assess Your Resources scale,³⁷ which uses Likert scale items, to measure respondent perceptions of the adequacy of resources available in their agency.

Aim 3: Characterize implementation processes and critical incidents. Qualitative. We will use qualitative data to triangulate quantitative findings, enrich our understanding of how the target mechanisms work and lead to outcomes, and critical incidents to EBPP implementation success or failure. Qualitative data will include interviews with 30 of the 475 county pairs (60 paired counties total). County pairs were randomly selected at Wave 1 (stratified by small, medium, and large county population) and followed longitudinally at Waves 2 and 3. We anticipate 180 qualitative interviews (60 respondents at 3 time points). We will alternate CJ and behavioral health respondents to obtain multiple perspectives on the county’s progress. Respondents will be invited for interviews regardless of their survey status (i.e., completed, not yet completed, declined) for that wave.

Fidelity to Stepping Up/Quantitative characterization of implementation strategies and sub-strategies used to improve mental health or substance use services for justice-involved individuals and/or to reduce the number of people with mental illness in jail (agencies nested within counties). We will use a checklist with strategies and their descriptions constructed from the six main Stepping Up strategies as well as categories conceptualized by Powell⁵⁴ and the CJ-IIM.¹⁷ Respondents will select whether anyone in the county is “planning to address this”, “some progress made”, or “significant progress made” using each strategy.

Power Analyses. The expected sample (475 paired counties, up to 4 respondents per county, resulting sample size of ~3,800) and the response rate of 50% gives an expected sample of 1,900 respondents. Given that anywhere from 0 to 4 respondents may complete the survey in any given county, with a 50% overall response rate, we anticipate that 712 counties will have at least one respondent who completes the survey. We used a conservative (higher than expected) intraclass correlation coefficient of 0.1 for addressing clustering of agencies within counties.

For county-level analysis, an effect size of 0.2, power of 0.8, confidence level of 95%, and statistical significance level of 0.05 were used to calculate minimum sample size. Repeated measures analysis required a minimum total sample size of 304 counties. For logistic regression and other non-linear predictive models, depending on the type and quantity of variables used in the model, the minimum total sample size required varied between 156 and 489 counties. With attrition, our calculations showed a power of 0.9 and more for most county-level analyses.

For agency-level analysis, a conservative agency-level effect size estimate ($d = 0.1$), a power of 0.8, confidence level of 95%, and significance level of 0.05 were used to calculate minimum required sample size. Repeated measures analysis comparing respondents from Stepping Up and comparison counties resulted in a minimum sample size of 524 respondents. When comparing the response measures, measured at the agency level, over time a minimum sample size of 1200 respondents is required. For predictive logistic regression and other non-linear predictive models, the minimum sample size varied between 673 and 1,100 respondents. Given the larger sample size of this study, our calculations showed a power higher than 0.9 for agency-level analyses.

Analysis Plan

General approaches. Primary tests will be 2-sided with $\alpha=0.05$. Analysis approaches accommodate nested and repeated measures data. We will examine predictive associations between Stepping Up membership, hypothesized target mechanisms, and implementation outcomes over time. We will use general linear models and generalized linear mixed models (GLMM) when the dependent variables are continuous and non-continuous, respectively. GEE will be used instead when distributional assumptions are not met. For non-aggregated dependent variables reported at the agency-level (i.e. hierarchical data), a random intercept growth hierarchical linear model (GHLM) will be fitted. All analyses will covary: (1) Wave 1 (baseline) values of dependent variables, (2) months since the county joined Stepping Up, (3) the matching score, (4) an indicator representing whether a the county shares their mental health administrator with other counties, and (5) a similar indicator for shared justice roles across counties.

Missing data. We will review survey completeness and recontact respondents to address quality issues and increase response rates. Logistic regression will be used to determine the type of missingness. Within waves, multiple imputation techniques will be applied. To address issues of missing data across different waves (i.e. over time), we will use generalized estimating equations (GEE) or weighted GEE, depending on the type of missing data.

Aim 1a: Comparison of target mechanisms between Stepping Up and non-Stepping Up counties. Primary. We will test the hypothesis that Stepping Up counties will show faster rate of improvement in use of/ capacity for performance monitoring (i.e., total scores on our Performance Monitoring measure) than comparison counties, using GLMM or GEE. Analyses will test for differences in slopes (rates of change). Separate secondary analysis will compare rates of change in the adapted Routine Decision-Making scale total score between Stepping Up and comparison counties. Secondary. We will test the hypothesis that Stepping Up counties will show faster rate of improvement in the use/functioning of interagency teams

(i.e., total scores on the integrated NCJTP Relationship Assessment Inventory-IOR measure) than control counties using the same statistical techniques. We will test the hypothesis that Stepping Up counties will show faster rate of improvement in common goals and mission across agencies (i.e., agreement between perceived agency and county priorities) than comparison counties using GLMM or GEE. We will conduct similar analyses of agreement among respondents within each county on the importance of mental health treatment for justice-involved individuals. We will test the hypothesis that Stepping Up counties will show faster rate of improvement in system integration (score of 6 or more on the Relationship Assessment Inventory score) using GEE. Separate secondary analysis will compare rates of change in use of the same screening and assessment instruments by multiple agencies in Stepping Up and comparison counties.

Aim 1b: Tests of mediation. Primary. We will test the hypothesis that changes in use of performance measures (i.e., scores on UPMDC and the adapted Routine Decision-Making scale) will mediate any differences found in rates of change in primary measures of justice-involved clients receiving behavioral health services, number of EBPPs, and resources available. These primary mediator analyses will use structural equation models, and path analyses. Secondary. We will conduct a series of analyses examining changes in interagency teams, common goals and missions, and integrated systems of care using scores from respective measures identified above, as mediators of number of justice involved individuals receiving services, number of EBPPs available, and number of resources using appropriate baseline measures or months since joining Stepping Up as controls.

Aim 2: Comparison of implementation outcomes between Stepping Up and non-Stepping Up counties. Primary. We will test the hypothesis that Stepping Up counties will show faster rate of improvement in number of justice-involved clients receiving behavioral health services than will comparison counties, using GLMM and GEE. Analyses will test for differences in slopes (rates of change) between the two sets of counties. Secondary. We will separately test the hypotheses that Stepping Up counties will show faster rate of improvement number of behavioral health EBPPs available to and resources for behavioral health EBPP for justice-involved individuals using GLMM, GEE, and GHLM.

We will examine moderators of the effects of Stepping Up participation on our primary outcome (justice-involved clients receiving behavioral health services) using structural equation models. A priori moderators include months between a county joining Stepping Up and study baseline, levels of implementation outcomes at study baseline, type of agency, organizational culture support for innovations (i.e., score on NCJTP Assess Your Organizational Culture scale³⁷), jails with their own behavioral health staff, yes/no presence of legislative reforms, counties in states that have mental health diversion funding, and counties with divisions that provide cross-system trainings.

Aim 3a: Characterize implementation processes and critical incidents (quantitative). We will examine the relationships between use of implementation strategies identified in the Implementation Strategy Checklist and faster rates of change in implementation outcomes, using GLMM and GEE while controlling for baseline measures and months since joining Stepping Up. We will use Bonferroni correction to control

for multiple comparisons of implementation strategies (using the Checklist) for each of the three implementation outcomes. Fidelity. We will compare Stepping Up and comparison counties on rates of use of Stepping Up strategies as a measure of fidelity to National Stepping Up program, and we will compare counties on rates of use of other strategies to explore whether Stepping Up impacts related strategies.

Aim 3b: Characterize implementation processes and critical incidents (qualitative). Qualitative data will be analyzed in line with study aims and key research questions using a two-stage analysis plan. In Stage 1, after each interview, interviewers will summarize key topics in framework matrix,⁵⁵ which allows key topics to be reviewed quickly. In Stage 2, recordings will be transcribed by a professional transcription service and will be anonymized before coding. *Deductive codes* will be drawn from interview question topics using the CJ-IIM, the 6 Stepping Up main strategies, and critical incidents. *Inductive codes* capturing emergent themes will arise from team-level review of transcripts. Coding team members will independently code transcripts; 20% will be double coded and reviewed for fidelity. Codes will be entered into NVivo¹⁶⁶, using thematic¹⁷¹ analyses; an audit trail will be maintained through code development and analysis. We will compare patterns found in qualitative data to patterns found in our quantitative data; this side-by-side comparison of patterns can identify sign-posts for additional exploration and analyses.

Discussion

This study will accelerate knowledge on how to implement mental health and substance use EBPPs in setting serving justice-involved individuals by evaluating target mechanisms, implementation outcomes, and strategies used in Stepping Up and comparison counties across the United States. This large study of an ongoing natural experiment is an unprecedented opportunity to evaluate multi-system CJ implementation efforts on a national scale. It will elucidate effective policies and strategies for bringing evidence-based practices to large, vulnerable, and underserved populations. Study rigor is improved through use of the CDC matching methodology and covarying baseline values in analyses. Analyses examine differences between groups in within-county rates of change over time, increasing rigor for this naturalistically occurring, non-randomized design.

Abbreviations

APAF: American Psychiatric Association Foundation

CDC: Centers for Disease Control and Prevention

CJ: criminal justice

CJ-IIM: Criminal Justice Interagency Implementation Model

CSG: Council of State Governments

EBPP: evidence-based practice or policy

GEE: generalized estimating equation

GHLM: growth hierarchical linear model

GLMM: generalized linear mixed models

IOR survey: Inter-Organization Relationship survey

NACo: National Association of Counties

NCJTP survey: National Criminal Justice Treatment Practices survey

PRISM Toolkit: Performance of Routine Information System Management Toolkit

UPMDC: Use of Performance Measures and Data Checklist

Declarations

Ethics approval and consent to participate.

The protocol for this study was reviewed by the George Mason University Institutional Review Board (protocol IP#204928), the Michigan State University Biomedical and Health Institutional Review Board (protocol IP# 00056524), and the University of Central Florida Institutional Review Board (protocols IP# 00001497) and declared exempt. Consent for participation in the study will be obtained using an electronic informed consent form.

Consent for publication.

Not applicable.

Availability of data and materials.

Not applicable.

Competing interests.

The authors declare that they have no competing interests.

Funding.

This study was funded by a grant from the National Institute of Mental Health (NIMH; R01 MH118680, PIs Taxman and Johnson). NIMH had no role in the design or implementation of the study, analysis of the data, or writing of the manuscript.

Authors' contributions.

FT and JJ conceptualized the research project. NR designed the statistical analyses and county matching algorithm, with input from JJ and FT. AC designed the economic analyses and provided input on conceptualization of behavioral health services delivery. JJ, RR, and MH designed the qualitative analyses and procedures, with input from FT and the study team. JV led refinement and finalization of the survey instrument and manages the survey project team. AB is assisting NR with data cleaning and analysis for matching and outcomes. JJ drafted the manuscript and all other authors reviewed and edited it.

Acknowledgements.

We would like to thank the National Association of Counties (NACo), the Council of State Governments (CSG) Justice Policy Center, and the American Psychiatric Association Foundation (APAF), who developed the Stepping Up Initiative, for their partnership in this study.

References

1. Kaeble D, Glaze L. *Correctional populations in the United States, 2015*. Bureau of Justice Statistics Special Report 2016.
2. James D, Glaze LE. *Mental Health Problems of Prison and Jail Inmates*. Bureau of Justice Statistics Special Report; 2006.
3. Choucair B. Health care for the homeless in America. *American Family Physician*. 2006;74:1099-1110.
4. Martin C, Bonner A, Brook A, Luscombe C. Factor structure and use of the Hospital Anxiety and Depression Scale in the homeless and socially marginalized. *Psychology, Health, and Medicine*. 2006;11:190-197.
5. Fickenscher A, Lapidus J, Silk-Walker P, Becker T. Women behind bars: Health needs of inmates in a county jail. *Public Health Reports*. 2001;116:191-196.
6. Dooley D, Prause J, Ham-Rowbottom KA. Underemployment and depression: Longitudinal relationships. *Journal of Health and Social Behavior*. 2000;41:421-436.
7. Shinkfield A, Graffam J. The relationship between emotional state and success in community reintegration for ex-prisoners. *International Journal of Offender Therapy and Comparative Criminology*. 2010;54(3):346-360.
8. Lerner D, Adler DA, Chang H, et al. Unemployment, job retention, and productivity loss among employees with depression. *Psychiatr Serv*. 2004;55(12):1371-1378.
9. Marcotte D, Wilcox-Gok V, Redmon PD. Prevalence and patterns of major depressive disorder in the United States labor force. *J Ment Health Policy Econ*. 1999;2(3):123-131.
10. O'Brien CP, Charney DS, Lewis L, Cornish JW, Post RM, Woody GE, Zubieta J, et al. Priority actions to improve the care of persons with co-occurring substance abuse and other mental disorders: A call to

- action. *Biological Psychiatry*. 2004 56:703-713.
11. Freudenberg N. Adverse effects of US jail and prison policies on the health and well-being of women of color. *American Journal of Public Health*. 2002;92(12):1895-1899.
 12. Bureau of Justice Statistics, U.S. Department of Justice. *Women Offenders*. Washington, DC: BJS Clearinghouse; 1999
 13. Browne A, Miller B, Maguin E. Prevalence and severity of lifetime physical and sexual victimization among incarcerated women. *International Journal of Law and Psychiatry* 1999;22:301-322.
 14. American Psychiatric Association Foundation, National Association of Counties, Council of State Governments Justice Policy Center. The Stepping Up Initiative. 2020; <https://stepuptogether.org/>. Accessed 12-14-2020, 2020.
 15. Substance Abuse and Mental Health Services Administration (SAMHSA). The Sequential Intercept Model: Advancing community-based solutions for justice-involved people with mental and substance use disorders. 2020; <https://store.samhsa.gov/sites/default/files/d7/priv/pep19-sim-brochure.pdf>, 10-29-2020.
 16. Lewis CC, Klasnja P, Powell BJ, Lyon AR, Tuzzio L, Jones S, Walsh-Bailey C, Weiner B. From classification to causality: Advancing understanding of mechanism of change in implementation science. *Frontiers in Public Health*. 2018;6:1-6.
 17. Taxman FS, Belenko S. *Implementing Evidence-Based Practices in Community Corrections and Addiction Treatment*. New York: Springer; 2011.
 18. Taxman FS, Cropsey KL, Young DW, Wexler H. Screening, assessment, and referral practices in adult correctional settings: A national perspective. *Criminal Justice and Behavior*. 2007;34(9):1216-1234.
 19. Melnick G, Ulaszek WR, Lin HJ, Wexler HK. When goals diverge: Staff consensus and the organizational climate. *Drug and Alcohol Dependence*. 2009;103:S17-S22.
 20. Friedmann P, Taxman, FS, Henderson CE. Evidence-based treatment practices for drug-involved adults in the criminal justice system. *Journal of Substance Abuse Treatment*. 2007;32:267-277.
 21. Henderson C, Young D, Farrell J, Taxman F. Associations among state and local organizational contexts: Use of evidence-based practices in the criminal justice system. *Drug and Alcohol Dependence*. 2009;103(Supp 1):S23-S32(PMID: 19174321).
 22. Henderson C, Taxman FS, Young DW. A Rasch model analysis of evidence-based treatment practices used in the criminal justice system. *Durg and Alcohol Dependence*. 2008;93:163-175.
 23. Taxman FS, Bouffard JA. The importance of systems in improving offender outcomes: New frontiers in treatment integrity. *Justice Research and Policy*. 2000;2(2):37-58.
 24. Henderson CE, Taxman FS. Competing values among criminal justice administrators: The importance of substance abuse treatment. *Drug and Alcohol Dependence*. 2009;103:S7-S16.
 25. von Elm E, Altman DG, Egger M, Pocock SJ, Gotsche PC, Vandenbroucke JP. Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) Statement: guidelines for reporting observational studies. *Lancet*. 2007;370(9596):1453-1457.

26. Metzler M, Kanarek N, Highsmith K, Straw R, Bialek R, Stanley J, Klein R. Community Health Status Indicators Project: The development of a national approach to community health. *Preventing Chronic Disease*. 2008;5(3):A94.
27. Kanarek N, Bialek R, Stanley J. Use of peer groupings to assess county public health status. *Preventing Chronic Disease*. 2008;5(3):A93.
28. Kanarek N, Tsai HL, Stanley, J. Health ranking of the largest US counties using the Community Health Status Indicators peer strata and database. *J Public Health Manag Pract*. 2011;5:401-405.
29. Ramezani N, Breno A, Mackey B, Cuellar A, Viglione J, Chase A, Johnson JE, Taxman F. Multilevel matching in natural experimental studies: Application to Stepping Up counties. . *American Statistical Association Journal*. 2020;JSM Proceedings, Survey Research Methods section:2408-2419.
30. Vera Institute of Justice. Incarceration Trends [Data set and code book]. 2018; <http://trends.vera.org/>.
31. Givens M, Gennuso K, Joyaag A, Willems Van Dijk J. County Health Rankings & Roadmap. County Report. 2017; <http://www.countyhealthrankings.org/2017-county-health-rankings-key-findingsreport>.
32. U.S. Department of Justice, Federal Bureau of Investigations. Table 78. Full time Law Enforcement Employees: by State by City [data set]. 2011; <https://ucr.fbi.gov/crime-in-the-u.s/2011/crime-in-the-u.s.-2011/police-employee-data>.
33. U.S. Department of Justice Federal Bureau of Investigations. Table 80. Full time Law Enforcement Employees: by State by Metropolitan and Non-Metropolitan Counties [data set]. 2011; <https://ucr.fbi.gov/crime-in-the-u.s/2011/crime-in-the-u.s.-2011/police-employee-data>.
34. Dillman D, Smyth JD, Christian TM. *Internet, Mail, and Mixed-Mode Surveys: The Tailored Design Method*. Hoboken, New Jersey: John Wiley & Sons; 2009.
35. Cook JM, O'Donnell C, Dinnen S, Coyne CJ, Ruzek JI, Schnurr PP. Measurement of a model of implementation for health care: toward a testable theory. *Implementation Science*. 2012;7:59.
36. Willis G. *Cognitive interviewing: A tool for improving questionnaire design*. Sage Publications; 2004.
37. Taxman FS, Young DW, Wiersema B, Rhodes A, Mitchell S. The National Criminal Justice Treatment Practices survey: Multilevel survey methods and procedures. *Journal of Substance Abuse Treatment*. 2007;32:225-238.
38. MEASURE_Evaluation. *Performance of Routine Information System Management (PRISM) Toolkit: PRISM Tools*. . Chapel Hill, NC, USA: University of North Carolina;2019.
39. Substance Abuse and Mental Health Services Administration (SAMHSA). *Principles of Community-based Behavioral Health Services for Justice-involved Individuals: A Research-based Guide*. Rockville, MD: Office of Policy, Planning, and Innovation. Substance Abuse and Mental Health Services Administration;2019.
40. Beck AJ, Maruschak M. *Mental health treatment in state prisons*. Washington DC: United States Department of Justice, Bureau of Justice Statistics;2001.
41. National Institute for Health and Care Excellence (NICE) Guidelines. Mental health of adults in contact with the criminal justice system [NICE Guideline NG66]. 2017;

- <https://www.nice.org.uk/guidance/NG66/chapter/Recommendations#psychological-interventions>. Accessed 1-21-2020, 2020.
42. Wolff N. Community reintegration of prisoners with mental illness: A social investment perspective. *International Journal of Law and Psychiatry*. 2005;28:43-58.
 43. Agency for Healthcare Research and Quality (AHRQ). *Comparative effectiveness review: Interventions for adults with serious mental illness who are involved with the criminal justice system*. 2013.
 44. Human Rights Watch. *Ill-equipped: U.S. prisons and offenders with mental illness*. Washington, DC: Human Rights Watch;2003.
 45. Peters RH, Young MS, Rojas EC, Gorey CM. Evidence-based treatment and supervision practices for co-occurring mental and substance use disorders in the criminal justice system. *Am J Drug Alcohol Abuse*. 2017;43(3):475-488.
 46. Substance Abuse and Mental Health Services Administration (SAMHSA). *MedTEAM: Getting started with evidence-based practices*. 2010.
 47. National Institute for Health and Care Excellence (NICE). Psychosis and schizophrenia in adults: Quality standard. 2015; <https://www.nice.org.uk/guidance/qs80/resources/psychosis-and-schizophrenia-in-adults-pdf-2098901855941>. Accessed 1-23-20, 2020.
 48. Yatahm L, Kennedy SH, Parikh SV, Schaffer A, Bond DJ, Frey BN, Sharma V, et al. Canadian Network for Mood and Anxiety Treatments (CANMAT) and International Society for Bipolar Disorders (ISBD) 2018 guidelines for the management of patients with bipolar disorder. *Bipolar Disorders*. 2018;20:97-190.
 49. CANMAT. Canadian Network for Mood and Anxiety Treatments (CANMAT) 2016 Clinical Guidelines for the management of adults with major depressive disorder. *The Canadian Journal of Psychiatry*. 2016;61(9):504-603.
 50. American Psychological Association Division 12. Psychological diagnoses and other targets of treatment. 2020; <https://div12.org/diagnoses/>. Accessed 4-1-2020, 2020.
 51. American Psychiatric Association. American Psychiatric Association Practice Guidelines. 2020; <https://psychiatryonline.org/guidelines>. Accessed 4-1-2020, 2020.
 52. American Foundation for Suicide Prevention. Treatment. 2020; <https://afsp.org/treatment>. Accessed 4-1-2020, 2020.
 53. National Institute on Drug Abuse. *Principles of drug addiction treatment: A research-based guide*. . National Institutes of Health, U.S. Department of Health and Human Services;2012.
 54. Powell B, Waltz TJ, Chinman MJ, Damschroder LJ, Smith JL, Matthieu MM, Proctor EK, Kirchner JE. A refined compilation of implementation strategies: Results from the Expert Recommendations for Implementing Change (ERIC) project. *Implementation Science*. 2015;10:21.
 55. Gale N, Heath G, Cameron E, Rashid S, Redwood S. Using the framework method for the analysis of qualitative data in multi-disciplinary health research. *BMC medical research methodology*, . 2013;13(1):117.

Tables

Table 1

Study constructs, function, relation to Stepping Up, and relation to the study conceptual model (CJ-IIM)

Construct	Study function	Relation to Stepping Up	CJ-IIM concept
Target mechanisms			
Use of/capacity for performance monitoring	Mechanism	Steps 2, 3, & 6: Collect, measure and track capacity	Performance monitoring
Use and functioning of interagency teams	Mechanism	Step 1: Convene a diverse team of leaders	Interagency workgroups
Common goals and mission across agencies	Mechanism	Steps 4,5: Develop & implement a plan together	Goal, mission setting
System integration (vs. single programs)	Mechanism	Steps 4,5: Develop & implement a plan together	System of care
Outcomes			
Number of justice-involved clients who received behavioral health services	Primary	Outcomes are indicators of county capacity to identify and treat justice-involved individuals with behavioral health problems and link them to or provide appropriate community care.	
Number of behavioral health EBPP available	Secondary		
Dollars spent on behavioral health EBPP	Secondary		

Table 2

Targeted type and number of survey respondents

Respondent type	Estimated number
Community mental health administrators	950
Jail administrators	950
Probation administrators	950
Community substance use treatment administrators	950

Figures

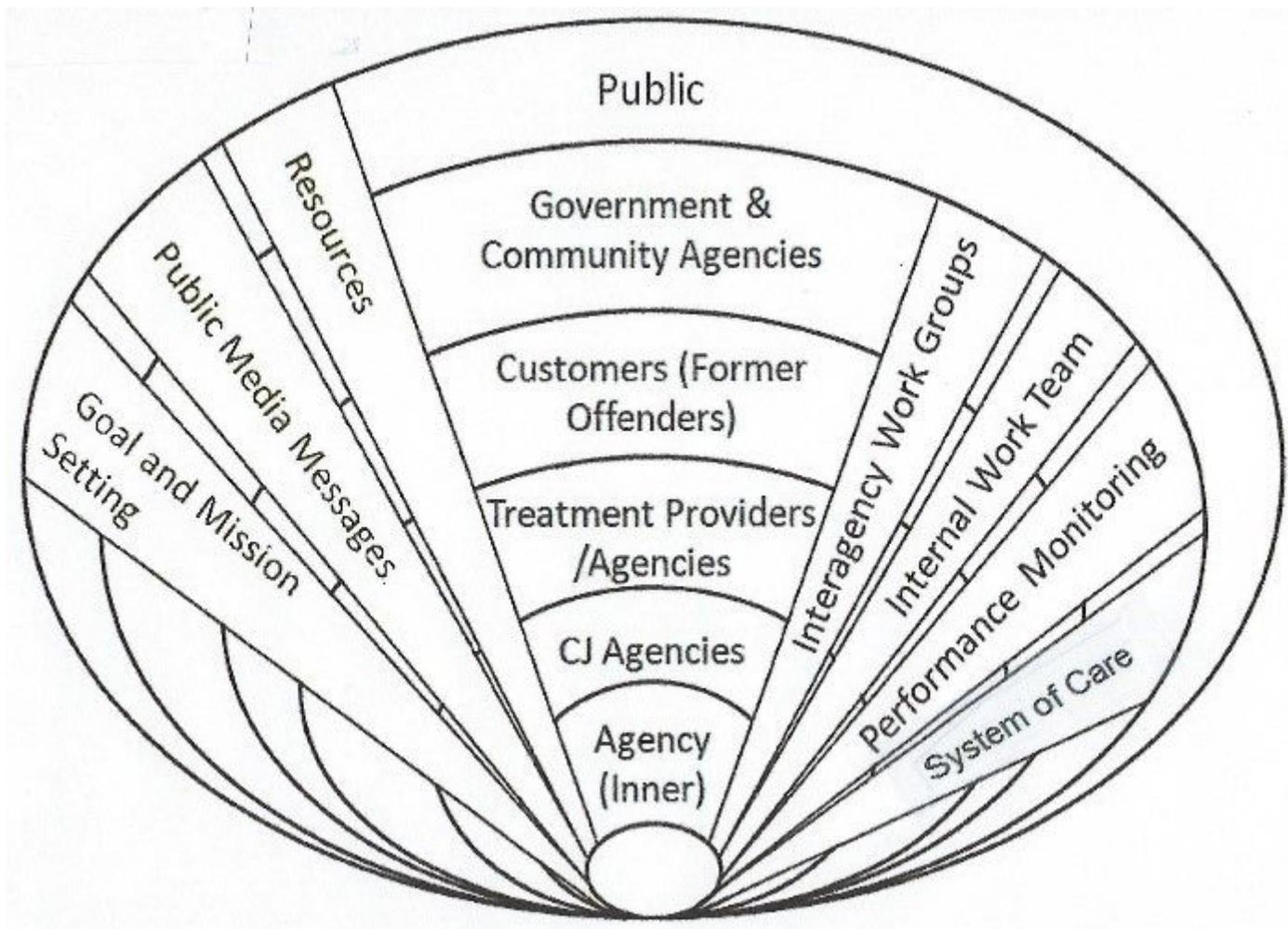


Figure 1

Study implementation mechanisms are taken from the Criminal Justice Interagency Implementation Model (IIM) legend: Figure adapted from Taxman et al.17

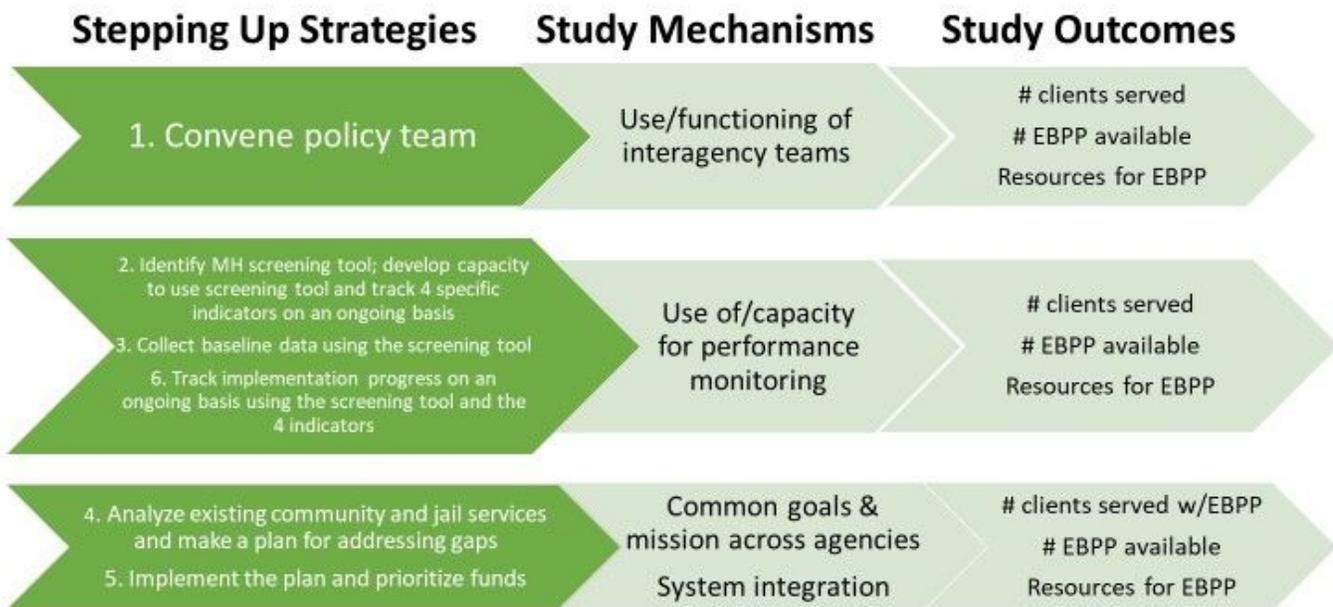


Figure 2

Study conceptual model

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [GMUexemptdetermination.pdf](#)
- [NIHNOA1R01MH11868001A1.pdf](#)
- [STROBEchecklistv4combined1112021.docx](#)