#  **Hot-Spotting High Risk/High Cost Behavioral Health Patients to Improve Outcomes and Reduce Acute Care Utilization**

 Sharon Touw, MPH, Sarah Jalbert, Ph.D, Colleen O’Brien, Psy.D.

## Introduction

 Emergency department visits and inpatient stays (for both medical and behavioral health reasons) made by high-need patients are a known driver of healthcare costs (Blumenthal and Abrams, 2016; Freeman et al, 2014). Reducing these types of encounters has been an objective of health policy and program interventions for over two decades. Complex care management (CCM) interventions are a common approach used to reduce the utilization of these complex, high-cost patients, who often have multiple chronic health conditions. They use strategies such as care coordination, team-based care, and data-driven identifications of high-need, high-utilizing patients via claims data (Bodenheimer and Berry-Millett, 2009). Many CCM programs identify patients based on actuarial findings of their service utilization, grouping medically diverse patients together in medical care management programs (Quinton et al, 2021).

Systems that have experimented with grouping patients by diagnosis type have done so without systematic, evidence-based approaches to identifying patient needs (O’Malley et al, 2019). Additionally, programs based on medical models of care may not be appropriate for the needs of patients whose health needs and utilization patterns stem from their behavioral health diagnoses and are highly intertwined with social determinants of health, including housing, food insecurity, and other economic and social needs. Programs that merely connect patients with existing social services might not provide enough needed additional support to change outcomes at either the individual or population level (Fitchenberg et al, 2020). Yet few CCM programs were designed explicitly for behavioral health patients with these additional support needs, and some, like the Camden, NJ Coalition of Healthcare Providers “Hotspot”, explicitly exclude patients with behavioral health diagnoses who have no co-occurring chronic medical conditions (Finkelstein et al, 2020).

Meeting the behavioral health and social needs of a high needs population with significant behavioral health disorders could have consequences for accountable care organizations and multiple other systems charged with service delivery and public safety. They a distinct sub-population among high-cost, high-utilizing patient groups (Quinton, et al, 2021), and behavioral health diagnoses are an independent predictor of high utilization and high healthcare costs, particularly among Medicaid populations (Sterling, et al, 2018; Freeman et al, 2014). Further, high-utilizing behavioral health patients are also frequently high-utilizers of other public services, like law enforcement and homeless services (Treglia et al, 2017, Vickery et al, 2018). Yet often the burden of accessing and coordinating behavioral health resources and resources to address social needs falls on the patient, or the patient’s family.

The Behavioral Health Intensive Clinical Advisor (BHICA) program was designed to address the complex needs of this sub-population. In 2020, the Cambridge Health Alliance Accountable Care Organization (CHA ACO) engaged with the Institute for Community Health (ICH) to understand the impact of the BHICA program on a group of high cost patients with severe mental illness and/or substance use disorders. CHA’s preliminary analyses suggested that the intervention significantly reduced behavioral health acute utilization. To confirm that this finding was not due to regression to the mean, CHA requested an independent review. ICH developed a quasi-experimental evaluation plan that compares the claims data from BHICA patients with a matched sample of Medicaid patients drawn from other Tufts Health Partnership Plans. Confirming the program’s impact was also important because recent research suggested that complex care management programs that focused on high medical utilizers did not reduce cost or utilization, (Finkelstein et al, 2020).

Additionally, although initially promising, some of these intensive programs to manage patient care, such as the Camden, NJ Coalition of Healthcare Providers “Hotspot” program, have failed to demonstrate their success in reducing health care costs for patients with complex medical needs (Finkelstein et al, 2020), suggesting that focusing on the highest utilizing patients may be reactive and ineffective, and supporting a shift away from *high risk* to *rising risk* when allocating limited healthcare management resources. But identifying eligible rising risk patients and targeting services according to their needs presents additional challenges to programs.

Confirming the program’s impact was significant because recent research had suggested that programs that focused on high medical utilizers were not able to control cost or utilization, (Finklestein et al, 2020). Meeting the behavioral health and social needs of this population could have consequences for accountable care organizations and multiple other systems charged with service delivery and public safety. Patients with behavioral health diagnoses are a distinct population among high-cost, high-utilizing patient groups (Quinton, et al, 2021), and behavioral health diagnoses are an independent predictor of high utilization and high healthcare costs, particularly among Medicaid populations (Sterling, et al, 2018; Freeman et al, 2014). Further, high-utilizing behavioral health patients are also frequently high-utilizers of other public services, like law enforcement and homeless services (Treglia et al, 2017, Vickery et al, 2018). Yet often the burden of accessing both behavioral health resources and resources to address social needs falls on the patient, or the patient’s family. Emergency department visits and inpatient stays (for both medical and behavioral health reasons) made by high-need patients are a known driver of healthcare costs (Blumenthal and Abrams, 2016; Freeman et al, 2014), and reducing these types of encounters has been an objective of health policy and program interventions for over two decades. Complex care management programs are medical interventions that have been widely implemented to better help patients with multiple chronic health conditions, using care coordination, team-based care, and data-driven identifications of high-need, high-utilizing patients via claims data (Bodenheimer and Berry-Millett, 2009). Although initially promising, some of these intensive programs to manage patient care, such as the Camden, NJ Coalition of Healthcare Providers “Hotspot” program, have failed to demonstrate their success in reducing health care costs for patients with complex medical needs (Finklestein et al, 2020), suggesting that focusing on the highest utilizing patients may be reactive and ineffective, and supporting a shift away from *high risk* to *rising risk* when allocating limited healthcare management resources. But identifying eligible rising risk patients and targeting services according to their needs presents additional challenges to programs.

## The BHICA program

**About the Intervention**

 There are few evidence-based practices, models, or frameworks for care transitions for patients after mental health hospitalization. (Viggiano, et al, 2021) In considering strategies to manage the quality, cost, and utilization of behavioral health services, the BHICA intervention targeted patients with primary behavioral health conditions and inpatient psychiatric readmissions or high levels of ED utilization. It aims to improve patient outcomes through proactive treatment planning and well-coordinated access to community-based services and supports, which will reduce costs associated with avoidable emergency department visits and inpatient readmissions. Patients with a behavioral health inpatient readmission or five or more ED visits in the previous six months, (at CHA or other hospitals), are eligible for enrollment. This study focused on patients who are members of a MassHealth Medicaid plan in Cambridge Health Alliance’s (CHA’s) Accountable Care Organization (ACO). The ACO provides claims data to identify eligible patients and the program also accepts direct referrals.



The Intensive Clinical Advisors (ICAs) are experienced, masters-level social workers and licensed mental health counselors who are trained in and understand the complex and fragmented behavioral health delivery system. They view high utilization as a failure of the healthcare system to serve high-need patients, and act to better integrate care without burden to patients. First, the ICAs conduct a root cause analysis on the drivers of acute utilization for each patient. Next, they coordinate care behind the scenes, focusing treatment plans, identifying missing supports, and integrating care to improve communication between internal and external members of the care team (see Figure 1). Internal members include providers or case workers from inpatient, outpatient, or emergency departments. External members of the team may include non-CHA behavioral health providers, guardians, family members, Behavioral Health Community Providers (a MassHealth care management program), and the Department of Mental Health (DMH) or other state-funded service providers. ICAs ensure any missing services for which the patient is eligible are put in place. Patients with Addiction and Mental Health conditions report that information sharing between health providers and between programs and services was not consistent or timely, and that having continuity of care to help with transitioning into the community after being discharged was very important to long term recovery. (Liu, et al, 2018) The ICAs focus on this barrier, which appears to be more significant in the behavioral health care delivery system than it is in other medical care.

The ICA model also includes a CHA Community Health Worker (CHW) who is a key bridge to outpatient support, resources and community based care management programs. For example, the CHW may visit the patient if they are readmitted and obtain the patient’s perspective on barriers to successful discharge. The CHW may offer to accompany the patient home from the hospital or to the discharge follow-up outpatient behavioral health appointment. The CHW will call the patient discharged from inpatient care to confirm that they have a seven-day follow-up outpatient appointment, and assist them in obtaining access to this appointment or arranging transportation if necessary. The CHW will also confirm that the patient has a safe place to stay, with access to food and prescription medications. If not, the CHW will provide short-term care management to help negotiate shelter, housing applications, and timely access to food and medication.

The ICAs strategically anticipate each patient’s potential use of avoidable acute services, and develop a plan to support alternate interventions in community-based settings. For example, the ICA might work with the ED to develop a care plan in advance of the patient's next ED visit that includes ED diversion plans. Patients experiencing homelessness who visit the ED multiple times for services that could be provided in less intensive settings might be referred to Healthcare for the Homeless for proactive outreach; a patient who is visiting the ED for non-medical issues might be referred to the patient assistance program team who will provide proactive home visits. Patients’ tenure in BHICA is dependent on the severity of their needs, their engagement in community-based treatment and support services, and their acute (inpatient and ED) behavioral healthcare utilization.

## Data and Methods

A previous non-experimental analysis of BHICA patients’ outcomes showed a significant reduction in emergency department and inpatient utilization following enrollment in BHICA. However, the literature suggests that regression to the mean following a spike in patient utilization is common among patients with complex needs (Johnson et al, 2015), meaning that utilization might have decreased for these patients anyway without the BHICA intervention. To understand whether BHICA reduces utilization and improves outcomes for patients who otherwise would continue to generate high utilization costs, an evaluation comparing BHICA patients to an untreated comparison group is necessary.

While resources and program design prevented us from implementing a randomized controlled trial (RCT), the authors sought to test the impact of the program using the most rigorous quasi-experimental methods possible. The study team used a matched-comparison difference in difference evaluation design to estimate the effect of the BHICA program on emergency department and inpatient utilization on enrolled patients compared to similar patients enrolled in a managed care organization. Using claims data for inpatient and emergency department utilization, the authors estimated the impact of the BHICA program on post-program utilization compared to utilization for comparable patients in the same time period.

*Obtaining data and defining the sample*

ICH worked with analysts from CHA to first identify BHICA patients and then obtain data claims data about this set of patients. Patients were considered eligible for the intervention group if they enrolled in BHICA for the first time for at least 30 days between 9/1/2018 and 9/30/2019. September 1, 2018 was chosen as the start date to allow for a 6-month pre-period. In February 2021, the study team received their demographic, utilization and health plan enrollment data for the period of March 1, 2018 (launch of CHA’s ACO) and October 31, 2020, allowing for an 8-month enrollment period and a 3-months claims lag.

The pull included patients who met the eligibility criteria for BHICA (5+ ED visits or 2+ inpatient stays for behavioral health) and those who passed through BHICA because they were having issues connecting to behavioral health in the community (Behavioral Health Community Partners (BHCP) program). Therefore, the BHICA eligibility criteria was applied to the BHICA pull to remove the BHCP patients.

For the comparison group, ICH worked with the analyst from Tufts Health Plan (THP) to define a data pull of the entire THP Managed Care population enrolled between March 1, 2018 and October 31, 2020. Data privacy restrictions limited the team’s access to identified member records in the THP administrative data. Instead, the authors requested de-identified demographic, plan enrollment and utilization data from THP, and assembled a comparable dataset provided by the CHA. BHICA eligibility criteria were then applied to the population to identify a comparison group for the same time period. Their enrollment date was considered the day that they were discharged from either their fifth ED visit or second inpatient stay in a six-month period.

A manufactured treatment period of 8 months (the average length of enrollment in BHICA) was applied to both groups. This was done to set a standard treatment length in order to apply it to the comparison group. Their 6-month pre-period was assigned as the 6 months prior to their enrollment and their six-month post period began at month 9. See example below.

|  |  |  |  |
| --- | --- | --- | --- |
| Enrolled or met eligibility criteria on: 11/1/2018 | Pre-period: 5/1/2018 to 10/31/2018 | Treatment period: 11/1/2018-6/30/2019 | Post-period: 7/1/2019-12/1/2019 |

After defining the two groups and their pre, treatment and post-periods, additional exclusion criteria were applied. The study team removed anyone for whom death was recorded as their disposition. From the BHICA pull, members under the age of 21 were excluded since the THP pull was limited to adults. This resulted in a sample of 687 THP members and 116 BHICA members. The study team used this sample to analyze the outcome of member engagement in the health care plan.

Member enrollment in MassHealth can be volatile based on members’ engagement and eligibility for the program each month. Patient churn, or frequent changes to MassHealth enrollment requirements, has been a persistent problem among health care organizations that serve Medicaid populations in Massachusetts, and often leads to interruptions and disruptions in care (Seifert, 2013). Patient churn can be caused by missed paperwork deadlines, processing delays, change in job status, or incarceration. Members who are dis-enrolled in MassHealth do not have observable utilization claims for that period of time, influencing the outcomes the study team is able to measure. To analyze the outcomes of emergency department and inpatient utilization, the sample was further restricted to those who had at least ⅓ of a month of enrollment in their health care plan during the pre, treatment and post-periods. This resulted in a reduced sample size of 462 THP members and 91 BHICA members eligible for the propensity score weighting process.

The authors applied a propensity score weighting process to the utilization sample, choosing to use weighting instead of matching to preserve the sample size and to increase precision when estimating treatment effects (Desai, 2019; SAS, 2016).

First, the authors specified a logistic regression model that creates the propensity score for each observation, that is to say the probability that the member would be assigned to BHICA. Analysts adjusted for gender, age category, rating category, and homelessness and substance use disorder recorded as a diagnosis at an encounter in the 9 months prior to enrollment, as these were covariates that differed between the groups and were potentially related to treatment assignment. To generate similar ranges of propensity scores between groups, the matching statement was modified to use the common support region instead of all observations. The common support region is the largest interval that contains overlapping propensity score values for subjects from both treatment conditions. Only two observations from the comparison group were excluded and the option resulted in a better match. The authors applied inverse probability of treatment weighting to estimate the average treatment effect (ATE) and output these weights to be used in the analysis. To assess balance between the treatment comparison group, the authors looked at the frequencies and standardized mean differences (SMD) before and after weighting. A SMD of less than .10 is indicative of well-balanced covariates between groups (Zhang et al, 2019).

The definitions of utilization outcomes of the previous evaluation of the BHICA program were adopted for this analysis. Behavioral health inpatient stays were defined as those admissions with an inpatient psychiatric room/board revenue code. Medical/surgical stays were defined as those without an inpatient psychiatric room/board revenue code. All stays with an OB/GYN or residential treatment room and board code were excluded.

Emergency room visits were defined as those with an emergency department revenue code, place of service listed as the emergency department, or a procedural code associated with the emergency department. Only emergency room visits that did not result in an inpatient stay were included. The team defined three types, primary behavioral health, secondary behavioral health and non-behavioral health based on the up to 25 ICD-10 codes associated with the visit. The list of ICD-10 codes for behavioral health was provided by the CHA IT department. If the first ICD-10 code was on the list, it was considered a primary behavioral health visit. If the first ICD-10 code was not on the list, but at least one of the other ICD-10 codes was, then it was a secondary behavioral health visit. If all of the ICD-10 codes at the visit were not on the behavioral health list, it was considered it a non-behavioral health ED visit.

For each member, the authors also calculated the length of enrollment time in the health plan for each time period.

Applying the propensity score weights, we described the average number of in-patient stays and in-patient days (total length of stay), relative to the number of months that the member was enrolled in the health plan, in the pre-period and post-period, as well as the difference between the two periods (post-pre). We did the same for the number of the three types of emergency visit outcomes.

We then used generalized linear models to look at the difference variables for each outcome, controlling for language of care category (English, not English, missing data) as this was not included in the weighting process. We were unable to include race/ethnicity category because of the high level of missingness in this data, only for the comparison group.

##

## Findings

## *Demographic and clinical characteristics of the groups*

We looked at the demographic and clinical characteristics of our two groups. Although the sample was slightly larger for the engagement outcome, the trends that we saw were fairly consistent across the two groups. Before the weighting process, BHICA members were more likely to be in the older age group, >54 years old, than the THP comparison group (24% vs 15%). They were also more likely to speak a language of care other than English (13% vs 2 %) and more likely to be male (65% vs 56%). BHICA patients were also most likely to come from the Greater Boston region, while the THP patients were more likely to come from the Central region of the state. BHICA patients were less likely to have an SUD diagnosis recorded than the THP patients (93% vs. 98%) and more likely to have a diagnosis of homelessness in the 9 months prior to their enrollment (46% vs. 39%). Nearly all BHICA patients and THP patients were very likely to have a Serious Mental Illness (SMI). Finally, BHICA patients were more likely to be in the rating category of disabled adults than the THP comparison group (46% vs 31%). After applying weights, many of these differences were no longer noticeable with the exception of race/ethnicity. We were unable to include race/ethnicity in the propensity score model because it was missing for 26% of the comparison group. Generally, BHICA patients were more likely to identify as Black or Hispanic/Latinx than the comparison group. For the complete demographic data, see appendix tables A1, A2, and A3 for complete demographic information.

## *Utilization*

Table 2 shows estimates of the treatment effect of BHICA on inpatient stays and inpatient length of stay (i.e. number of days across all stays), relative to the number of member months enrolled. There were statistically significant differences in the number of BH stays between the pre-periods and the change (post to pre) between the two groups. There were also statistically significant differences between the groups for behavioral health days during the pre- and post- periods and on the change (post to pre). The asterisk (\*) denotes places where we saw statistically significant differences.

*Table 2. Inpatient stays and days*

|  |  |  |
| --- | --- | --- |
|  | **BHICA (n=91)** | **THP (n=460)** |
| Measure | Pre | Post | Difference: Post-Pre | Pre  | Post | Difference |
| **Behavioral health stays** |  |  |  |  |  |  |
| Mean (SD) | .54 (1.56)\* | .15 (1.15) | -0.39 (1.93)\* | .29 (.42)\* | .13 (.28) | -0.16 (.49)\* |
| 95% Confidence Interval | .41, 68 | .06, .25 | -.56, -.22 | .26, .33 | .11, .15 | -.16, -.12 |
| Median | .33 | 0 | -0.33 | .2 | 0 | -0.16 |
| **Med-surg stays** |  |  |  |  |  |  |
| Mean (SD) | .15 (.65) | .11 (.58) | -0.04 (.61) | 0.13 (0.4\*) | .09 (.55)\* | -0.04 (.33) |
| 95% Confidence Interval | .10, .21 | .06, .16 | -10, ,01 | .11,.16 | .07,.12 | -0.07,-0.01 |
| Median | 0 | 0 | 0 | 0 | 0 | 0 |
| **Behavioral health days** |  |  |  |  |  |  |
| Mean (SD) | 3.59 (11.28)\* | 0.46 (3.09)\* | -3.13 (11.5)\* | 2.26 (4.06) | 1.05 (2.98) | -1.21 (4.53)\* |
| 95% Confidence Interval | 2.61, 4.57 | .20, .73 | -4.14, -2.11 | .1.92, 2.6 | .80, 1.3 | -1.6, -0.82 |
| Median | 1.84 | 0 | 0 | 1.17 | 0 | -0.37 |
| **Med-surg days** |  |  |  |  |  |  |
| Mean (SD) | .57 (3.02) | .22 (1.88) | -0.35 (3.21)\* | .43 (1.17) | .37 (1.73) | -0.06 (1.8)\* |
| 95% Confidence Interval | .29, .86 | .08, .36 | -0.64, -0.07 | .33, .53 | .21, .52 | -22, .10 |
| Median | 0 | 0 | 0 | 0 | 0 | 0 |

 We also looked at emergency department visits across the categories in the three groups, but there were no statistically significant findings. See Table A4 in the appendix.

Table 3 shows estimates of the treatment effect of BHICA on health care utilization outcomes in a model adjusting for language of care category (English, language other than English or unknown). Consistent with the program’s stated goals, BHICA decreased the number of hospital stays for behavioral health care by .193 per member month when compared to the THP group. This difference is statistically significant (p=<.05). Consistent with this finding, BHICA members had 1.6 fewer behavioral health days per member month when compared with the THP group. Similarly, BHICA members had .3 (p=.02) fewer med-surg stays compared with the THP group, and .17 (p=.02) fewer emergency department visits with a secondary BH diagnosis compared with the THP group, both significant findings. Other emergency department outcomes had no significant differences.

*Table 3. Impact of BHICA on Utilization outcomes*

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Outcome | Model | Estimate BHICA vs THP | p-value | Interpretation |
| Difference in BH stays (post-pre) | Linear model adjusted for propensity score weights | -0.193 | **0.006** | Adjusted for language of care, BHICA patients, on average, decreased their number of stays (post to pre) by .193 more per member month than the THP population. This difference was significant. |
| Difference in med-surg stays (post-pre) | Linear model adjusted for propensity score weights | -0.013 | 0.6026 | There is no significant association in the difference in number of med surg stays (post - pre) between the two groups. |
| Difference in BH days (post-pre) | Linear model adjusted for propensity score weights | -1.619 | **<0.0001** | Adjusted for language of care, BHICA patients, on average, decreased their number of days in a BH stay (post to pre) by 1.6 more per member month than the THP population. This difference was significant. |
| Difference in med-surg days (post-pre) | Linear model adjusted for propensity score weights | -0.305 | 0.0213 | Adjusted for language of care, BHICA patients, on average, decreased their number of med-surg days by .3 per member month compared to the THP population. The difference was significant.  |
| Difference in ED visits with a primary BH diagnosis (post-pre) | Linear model adjusted for propensity score weights | 0.035 | 0.6428 | There is no significant association in the difference in number of ED visits for a primary BH diagnosis (post - pre) between the two groups. |
| Difference in ED visits with a secondary BH diagnosis (post-pre) | Linear model adjusted for propensity score weights | -0.172 | 0.0238 | Adjusted for language of care, BHICA patients, on average, decreased their number of ED visits with a secondary BH diagnosis by .17 per member month compared to the THP population.  |
| Difference in ED visits with a non BH diagnosis (post-pre) | Linear model adjusted for propensity score weights | 0.003 | 0.9469 | There is no significant association in the difference in number of ED visits with no BH diagnoses (post - pre) between the two groups. |

##   *Engagement in MassHealth*

As our study team assembled the final sample for the impact evaluation, the authors encountered a significant difference between BHICA member engagement in MassHealth as compared to the eligible THP members. Although member engagement was not an initial outcome for the study, the authors made a decision to include engagement in MassHealth as an exploratory outcome for this study. The BHICA program is not intended to impact engagement in MassHealth directly, and the study team did not set out to test hypotheses related to patient engagement in this evaluation, so results of this exploratory analysis should be interpreted with caution.

Enrollment in BHICA was significantly associated with the total number of months patients were engaged in MassHealth in the post-treatment period (Table 1). BHICA members were engaged in MassHealth, on average, .82 months more than THP members (p=.0003). When exclusion criteria were applied to THP and BHICA member samples, a significantly greater proportion of THP patients (33% vs 22%) were excluded because they did not have enough observable months engaged in MassHealth (p=.016) These results suggest that enrollment in BHICA is correlated with better patient engagement in MassHealth.

*Table 1. Patient Engagement in MassHealth- Months*

|  |  |  |  |
| --- | --- | --- | --- |
|   | BHICA (n=116) | THP (n=687) | p-value |
| Measure | Mean | SD | CI | Median | Mean | SD | CI | Median |   |
|  enrollment during the pre period | 5.16 | 1.61 | 4.86 to 5.45 | 5.98 | 5.32 | 1.34 | 5.22 to 5.42 | 5.97 | 0.1453 |
| enrollment during the post period | 4.4 | 2.51 | 3.94 to 4.87 | 5.98 | 3.58 | 2.76 | 3.38 to 3.79 | 5.83 | **0.0003** |
| enrollment during the treatment period | 6.9 | 2.05 | 6.53 to 7.28 | 7.98 | 6.23 | 2.48 | 6.05 to 6.42 | 8 | 0.0665 |

## Limitations

 BHICA is a novel approach to reducing utilization costs without burdening the patient with additional provider visits, intake procedures, or responsibilities for engagement. As a novel program, the evaluation team was challenged in identifying an appropriate rigorous methodology for testing its effectiveness. Several key issues presented barriers to rigorous impact evaluation, including lack of a straightforward comparison group with a known counterfactual. We were able to build a comparison group of similar patients served in other managed care programs, but we know nothing about the programming they may have been enrolled in during the evaluation period. The study team also had to receive the data from two different sources, so although the authors worked to align the data pulls across the two sources, there could be some unknown differences in how the pulls were executed.

 The study team was challenged by poor documentation of the BHICA patient population and the use of nuance in defining eligibility for BHICA. Early in the ICA program, BHICA patients were labeled in similar ways to patients who received brief ICA support for transition to enrollment in MassHealth’s Behavioral Health Community Partner (BHCP) program. The lack of a clear indicator distinguishing ongoing BHICA enrollment from short-term, transitional support may have caused us to include patients who did not receive ongoing support from the BHICA program. To mitigate this challenge, we applied the BHICA claims-based eligibility criteria. This excluded BHCP patients who did not also meet the BHICA criteria. Finally, clinical advisors used appropriate nuance in determining enrollment. For example, if a patient went to the emergency room for the consequences of not taking their anti-seizure medication due to behavioral health issues, they would be deemed eligible for BHICA. However, in looking at the claims data, this patient would not be eligible unless the ED provider coded some sort of behavioral health diagnosis.

 There were also challenges with missing data, especially with the comparison group. 26% of the comparison group was missing their race/ethnicity, so the authors were unable to account for this important covariate in our analysis.

 Small sample sizes in the treatment group also presented challenges. As a program targeted to the high-utilizing patients, the program is naturally limited to a subset of high-need patients. Given the size of our sample, the authors computed the average treatment effect (ATE) across patients enrolled in BHICA. These patients are both high utilizers and super-utilizers, and the super-utilizers likely influenced the ATE. This study cannot determine whether BHICA would be effective without the inclusion of super-utilizers in the sample.

 Readers may also be concerned about endogeneity issues presented by significant differences in patient engagement in MassHealth, and thus in the observable patient utilization data. Since claims data were only available for this study for months when patients were engaged in MassHealth, the study team was very concerned that engagement in BHICA influenced the availability of outcomes data. We attempted to mitigate endogeneity concerns, as described above, by examining outcomes relative to the availability of patient claims data in each month of the post period. However, this is a serious limitation of this study.

Additional research and evaluation of the BHICA program will need to take changes to healthcare delivery into account, and may be similarly challenged by sample sizes and comparison time periods due to changes in healthcare brought on by COVID. Future studies that utilize a similar retrospective difference in difference estimator will be challenged by comparing time periods before and after the initiation of the COVID pandemic, which impacted availability and accessibility of in-person behavioral health care services (Pierce et al, 2021). Some COVID-related changes to behavioral health care delivery may be beneficial to community mental health care delivery (Kopelovich et al, 2021), but comparing pre-COVID and post-COVID patient outcomes in future evaluations would not, in the authors’ opinion, be a valid evaluation strategy.

## Discussion

##

 Despite the limitations described above, there is reason to believe that BHICA offers value to the CHA ACO by increasing patient engagement in MassHealth and decreasing the number of behavioral health-related hospital stays and length of behavioral-health related hospital stays. Furthermore, these outcomes are achieved by the program without adding additional burdens on patients.

 Consistent with the BHICA program’s focus on behavioral health care improvement, patients enrolled in the program experienced fewer hospital stays and reduced length of stays for behavioral health-related care in the post-treatment period. These reductions were not observed for non-behavioral health related medical-surgical stays, although the number of post-program medical-surgical days associated with a stay were reduced and the p-value approached significance at p <.07).

The BHICA program presents minimal burden to patients, streamlines their connections to care with community partners, is associated with greater engagement in MassHealth, and appears to reduce the incidence and length of future behavioral health-related hospital stays.

For those who have co-occurring behavioral health and medical disorders, medical case management programs that focus on reducing costs and utilization may not be appropriate interventions, and may be focused on the wrong metrics for success. Integrated care management programs for high-risk, high-need patients are attempting to address both health and social needs by measuring outcomes based on health care utilization only, failing to acknowledge the social and financial benefits these programs may be realizing across multiple systems of care. For behavioral health patients in particular, the availability of services that meet their treatment and recovery needs may lead to long term community tenure (Hamilton et al, 2016), improved quality of life and experience with care, and other social benefits (e.g. reduced jail and prison utilization, reduced overdose) that are not identifiable in hospital utilization records but are clearly important positive outcomes.

Implications for Behavioral Health

 The ICA experience suggests that programs that risk-stratify behavioral health acute care utilization (inpatient admissions and ED visits) can have an impact on the cost of care by decreasing avoidable readmissions and strengthening community-based supports to reduce the patient’s need for ED visits. Examining behavioral health outliers can help ACOs and focus on issues where the fragmentation of the overall behavioral healthcare delivery system has left patients vulnerable to readmissions or avoidable ED visits. This fragmentation, combined with inequities and social determinants of health, (for example, racism and homelessness, lack of access to healthy food, prescribed medication, and outpatient mental health services) increase the risk that patients who could be stabilized in community settings will instead present in the ED.

Perhaps uncontrolled, avoidable acute care utilization is more of a risk for patients with primary or comorbid behavioral health conditions than it is patients with medical conditions alone, making hotspotting high risk/high cost patients a more effective intervention for this population that was seen in the Camden study of patients without mental illness or addictions. Of note, some of the patients with substance use disorders in the ICA program were repeatedly brought to the ED by the police, (often due to sleeping in public and the risk of lethal overdose). Providing residential care and recovery services dramatically impacted the ED use in this group, as even impacting a small number of patients disrupted patterns generating more than 80 visits per person each year. The ICA experience strongly suggests that focusing on the most challenging patients can yield a return, particularly if multiple readmissions and avoidable ED visits are seen as delivery system failures that can be remedied, and not unavoidable consequences of addiction and chronic mental illness. Investments in access to intermediate and outpatient levels of care, and community-based programming, may bend the cost curve by resulting in acute care savings. Over time, the impact of the ICA intervention on physical health may also be realized in a reduction in the cost of care of patients suffering the physical consequences of addiction or disengagement from healthcare in general.

## Recommendations for future research

 Future research on BHICA and programs like it should be conducted prospectively, and include qualitative measures of program implementation and fidelity to the model. While we believe that randomization would prove extremely challenging both ethically and logistically, a randomized controlled trial (RCT) would allow for more confidence in the treatment effect for BHICA. Absent an RCT, a next-best option for evaluating BHICA would be a prospective regression-discontinuity (RD) study that employs strict, continuous eligibility criteria for BHICA patients.

 A potential prospective RD study could take advantage of existing eligibility criteria for BHICA (a threshold for BH acute care utilization) to study the impact of programming for patients closest to the cutoff point. Above, we describe the limitation of using the ATE to estimate the impact of a program like BHICA. An RD approach can help to mitigate this limitation by testing the impact of BHICA in a way that approximates randomization without the logistical and ethical challenges of randomly assigning patients to different treatment conditions. Furthermore, an RD approach would allow the CHA ACO to test the impact of variations to the eligibility threshold, assuming there is time to allow for sufficient sample sizes to accumulate. However, an RD approach would require strict adherence to the enrollment criteria, which would not allow for nuance in assigning patients to BHICA (i.e. taking referrals for patients who do not strictly meet utilization criteria but might benefit from the program).

 Future evaluations may also benefit from a benefits-cost analysis of BHICA programming to determine both what the optimal threshold for BHICA engagement might be, and to determine the return on investment to the CHA ACO. An evaluation that includes qualitative measures of program implementation might also identify other social returns on investment of the BHICA program, including avoided costs in the criminal justice system, emergency services, and improved quality of life measures.

# Appendix A. Demographic Tables

*Table A1. Patient Demographics (Engagement in MassHealth Outcome)*

|  |  |  |  |
| --- | --- | --- | --- |
|  | **BHICA (n=116)** | **THP (n=687)** |  |
|  | **n** | **%** | **n** | **%** | **p-value** |
| ***Age range*** |   |   |   |   | **0.082** |
| 21 to 54 years old | 92 | 79.31% | 588 | 85.59% |   |
| >54 years old | 24 | 20.69% | 99 | 14.41% |   |
| ***Language of care***  |  |  |  |  | **<.0001** |
| English | 103 | 88.79% | 636 | 92.58% |   |
| Non-English\* | 13 | 11.21% | 13 | 1.89% |   |
| Missing data | 0 | 0.00% | 38 | 5.53% |   |
| ***Combined race ethnicity*** |   |   |   |   | **<.0001** |
| African American/Black (non-Hispanic) | 25 | 21.55% | 25 | 3.64% |   |
| Hispanic/Latino Ethnicity | 40 | 34.48% | 39 | 5.68% |   |
| White (non-Hispanic) | 39 | 33.62% | 415 | 60.41% |   |
| Another race/ethnicity | 12 | 10.34% | 22 | 3.20% |   |
| Unknown/Missing data | 0 | 0.00% | 186 | 27.07% |   |
| ***Gender*** |   |   |   |   | **0.105** |
| Female | 40 | 34.48% | 292 | 42.50% |   |
| Male | 76 | 65.52% | 395 | 57.50% |   |
| ***Region*** |  |  |  |  | **<.0001** |
| Central | 4 | 3.45% | 273 | 39.74% |   |
| Greater Boston | 70 | 60.34% | 122 | 17.76% |   |
| Northern | 41 | 35.34% | 251 | 36.54% |   |
| Southern | 1 | 0.86% | 12 | 1.75% |   |
| Western | 0 | 0.00% | 29 | 4.22% |   |
| ***BH Diagnoses*** |  |  |  |  |  |
| SUD | 110 | 94.83% | 678 | 98.69% | 1.32% |
| SMI | 116 | 100.00% | 685 | 99.71% | 1.000 |
| MBHP | 115 | 99.14% | 672 | 97.82% | 0.491 |
| ***Indicator of Homelessness*** |  |  |  |  |  |
| Ever homeless | 73 | 62.93% | 374 | 54.44% | 0.089 |
| Homeless in the 9 mos prior to enrollment | 57 | 49.14% | 281 | 40.90% | 0.097 |
| ***Rating Category*** | ***BHICA*** | ***THP*** | ***0.040*** |
| RC2A: Disabled - adult | 54 | 46.55% | 226 | 32.90% |   |
| RCIA: Mother and child - adult | 14 | 12.07% | 117 | 17.03% |   |
| RCIX: Medicaid Expansion | 46 | 39.66% | 330 | 48.03% |   |
| RCX: Medicaid Expansion - disabled  | 2 | 1.72% | 14 | 2.04% |   |

 *Table A2. Patient (Unweighted) (Patient Utilization Outcomes)*

|  |  |  |  |
| --- | --- | --- | --- |
|  | **BHICA (n=91)** | **THP (n=462)** | **Standardized difference** |
|  | **n** | **%** | **N** | **%** |
|  |  |  |  |  |
| ***Age range*** |  |  |  |  | 0.24 |
| 21 to 54 years old | 69 | 75.82% | 394 | 85.28% |   |
| >54 years old | 22 | 24.18% | 68 | 14.72% |   |
| ***Language of care***  |  |  |  |  |  0.56 |
| English | 79 | 86.81% | 428 | 92.64% |   |
| Non-English\* | 12 | 13.19% | 8 | 1.73% |   |
| Missing data | 0 | 0.00% | 26 | 5.63% |   |
| ***Combined race ethnicity*** |   |   |   |   |  1.59 |
| African American/Black (non-Hispanic) | 22 | 24.18% | 15 | 3.25% |   |
| Hispanic/Latino Ethnicity | 33 | 36.26% | 26 | 5.63% |   |
| White (non-Hispanic) | 29 | 31.87% | 286 | 61.90% |   |
| Another race/ethnicity | 7 | 7.69% | 12 | 2.60% |   |
| Unknown/Missing data | 0 | 0.00% | 123 | 26.62% |   |
| ***Gender*** |   |   |   |   | 0.18 |
| Female | 32 | 35.16% | 202 | 43.72% |   |
| Male | 59 | 64.84% | 260 | 56.28% |   |
| ***Region*** |  |  |  |  |  1.28 |
| Central | 3 | 3.30% | 179 | 38.74% |   |
| Greater Boston | 54 | 59.34% | 76 | 16.45% |   |
| Northern | 33 | 36.26% | 185 | 40.04% |   |
| Southern | 1 | 1.10% | 6 | 1.30% |   |
| Western | 0 | 0.00% | 16 | 3.46% |   |
| ***BH Diagnoses*** |  |  |  |  |   |
| SUD | 85 | 93.41% | 455 | 98.48% | 0.26 |
| SMI | 91 | 100.00% | 460 | 99.57% | 0.09 |
| ***Indicator of Homelessness*** |  |  |  |  |   |
| Homeless in the 9 mos prior to enrollment | 42 | 46.15% | 179 | 38.74% | 0.15 |
| ***Rating Category*** | ***BHICA*** |  | ***THP*** |  |  0.27 |
| RC2A: Disabled - adult | 40 | 43.96% | 145 | 31.39% |   |
| RCIA: Mother and child - adult | 12 | 13.19% | 77 | 16.67% |   |
| RCIX: Medicaid Expansion | 37 | 40.66% | 231 | 50.00% |   |
| RCX: Medicaid Expansion - disabled  | 2 | 2.20% | 9 | 1.95% |   |

*Table A3. Patient Demographics (Weighted with propensity score) (Patient Utilization Outcomes)*

|  |  |  |  |
| --- | --- | --- | --- |
|  | **Exposed (n=91)** | **Unexposed (n=460)** | **Standardized difference** |
|  | N | Weighted | % | n | n weighted | Percent |   |
|  |   |   |   |   |   |   |   |
| ***Age range*** |   |   |   |   |   |   | 0.04 |
| 21 to 54 years old | 67 | 452.45 | 82.8% | 394 | 461.26 | 84.3% |   |
| >54 years old | 24 | 94.17 | 17.2% | 66 | 85.94 | 15.7% |   |
| ***Language of care***  |   |   |   |   |   |   |  0.58 |
| English | 79 | 471.07 | 86.2% | 426 | 506.70 | 92.6% |   |
| Non-English\* | 12 | 75.56 | 13.8% | 8 | 9.27 | 1.7% |   |
| Missing data | 0 | 0.00 | 0.0% | 26 | 31.23 | 5.7% |   |
| ***Combined race ethnicity*** |   |   |   |   |   |   |  1.51 |
| African American/Black (non-Hispanic) | 22 | 131.46 | 24.0% | 15 | 18.20 | 3.3% |   |
| Hispanic/Latino Ethnicity | 33 | 179.30 | 32.8% | 26 | 31.27 | 5.7% |   |
| White (non-Hispanic) | 29 | 187.48 | 34.3% | 284 | 336.86 | 61.6% |   |
| Another race/ethnicity | 7 | 48.39 | 8.9% | 12 | 14.07 | 2.6% |   |
| Unknown/Missing data | 0 | 0.00 | 0.0% | 123 | 146.80 | 26.8% |   |
| ***Gender*** |   |   |   |   |   |   | 0.05 |
| Female | 32 | 220.23 | 40.3% | 202 | 232.90 | 42.6% |   |
| Male | 59 | 326.40 | 59.7% | 258 | 314.29 | 57.4% |   |
| ***Region*** |   |   |   |   |   |   |  1.23 |
| Central | 3 | 18.04 | 3.3% | 179 | 213.68 | 39.0% |   |
| Greater Boston | 54 | 301.18 | 55.1% | 75 | 89.85 | 16.4% |   |
| Northern | 33 | 221.93 | 40.6% | 184 | 217.04 | 39.7% |   |
| Southern | 1 | 5.47 | 1.0% | 6 | 7.26 | 1.3% |   |
| Western | 0 | . | 0.0% | 16 | 19.36 | 3.5% |   |
| ***BH Diagnoses*** |   |   |   |   |   |   |   |
| SUD | 85 | 532.63 | 97.4% | 455 | 539.76 | 98.6% | 0.09 |
| SMI | 91 | 546.63 | 100.0% | 458 | 544.81 | 99.6% | 0.09 |
| MBHP | 90 | 542.54 | 99.3% | 453 | 539.08 | 98.5% | 0.07 |
| ***Indicator of Homelessness*** |   |   |   |   |   |   |   |
| Homeless in the 9 mos prior to enrollment | 42 | 225.27 | 41.2% | 177 | 217.13 | 39.7% | 0.03 |
| ***Rating Category*** |   |   |   |   |   |   |  0.02 |
| RC2A: Disabled – adult | 40 | 181.60 | 33.2% | 143 | 180.21 | 32.9% |   |
| RCIA: Mother and child – adult | 12 | 84.97 | 15.5% | 77 | 88.30 | 16.1% |   |
| RCIX: Medicaid Expansion | 37 | 270.29 | 49.4% | 231 | 267.77 | 48.9% |   |
| RCX: Medicaid Expansion - disabled  | 2 | 9.77 | 1.8% | 9 | 10.91 | 2.0% |   |

Propensity score adjusted for homeless\_9mos gender rating\_cat age\_cat sud

*Table A3. Emergency utilization outcome*

|  |  |  |
| --- | --- | --- |
|  | **BHICA** | **Comparison group** |
|  | Pre | Post | Difference (Post-Pre) | Pre | Post | Difference (Post-Pre) |
| **Primary BH encounters per member month** |
| Mean (SD) | 0.41 (1.7) | 0.3 (1.8) | -0.11 (2.26) | 0.41 (0.63) | 0.27 (0.73) | -0.14 (0.84) |
| Median | 0.17 | 0 | 0 | 0 | 0 | -0.17 |
| 95% CI | 0.26-0.56 | 0.17-0.44 | -0.30-.08 | 0.35- 0.46 | 0.21-0.33 | -0.16 - -0.08 |
| **Secondary BH encounters per member month** |
| Mean (SD) | **0.67 (0.33)\*** | **0.36 (1.6)\*** | -0.31 (2.3) | **0.29 (0.44)\*** | **0.18 (0.37)\*** | -0.11 (0.51) |
| Median | .33 | 0 | -0.17 | 0.17 | 0 | 0 |
| 95% CI | 0.45-0.88 | 0.23- 0.49 | -0.51- -0.1 | 0.26- 0.33 | 0.14-0.21 | -0.16 - -0.08 |
| **Non BH encounters per member month** |
| Mean (SD) | 0.17 (1) | 0.08 (0.25) | -0.09 (0.83) | 0.16 (0.6) | 0.07 (0.2) | -0.09 (0.56) |
| Median | 0 | 0 | 0 | 0 | 0 | 0 |
| 95% CI | 0.09-0.25 | 0.04-0.11 | -0.16- -0.03 | 0.11- 0.21 | 0.05-0.08 | -0.14 - -0.04 |

# Appendix B: References

Blumenthal D, Abrams MK. Tailoring Complex Care Management for High-Need, High-Cost Patients. JAMA. 2016;316(16):1657–1658. doi:10.1001/jama.2016.12388

Bodenheimer, T., & Berry-Millett, R. (2009). Care management of patients with complex health care needs. Policy, 1(6).

Desai R J, Franklin J M. Alternative approaches for confounding adjustment in observational studies using weighting based on the propensity score: a primer for practitioners BMJ 2019; 367 :l5657 doi:10.1136/bmj.l5657

Finkelstein, A., Zhou, A., Taubman, S., & Doyle, J. (2020). Healthcare hotspotting—a randomized, controlled trial. New England Journal of Medicine, 382(2), 152-162.

Fichtenberg, C., Delva, J., Minyard, K., & Gottlieb, L. M. (2020). Health and Human Services Integration: Generating Sustained Health and Equity Improvements: An overview of collaborations, partnerships, and other integration efforts between health care and social services organizations. Health Affairs, 39(4), 567-573.

Freeman, E., McGuire, C. A., Thomas, J. W., & Thayer, D. A. (2014). Factors affecting costs in Medicaid populations with behavioral health disorders. Medical care, S60-S66.

Hamilton, J. E., Desai, P. V., Hoot, N. R., Gearing, R. E., Jeong, S., Meyer, T. D., ... & Begley, C. E. (2016). Factors associated with the likelihood of hospitalization following emergency department visits for behavioral health conditions. Academic Emergency Medicine, 23(11), 1257-1266.

Johnson, T.L., Rinehart, D.J., Durfee, J., Brewer, D., Batal, H., Blum, J., Oronce, C.I., Melinkovich, P. and Gabow, P., 2015. For many patients who use large amounts of health care services, the need is intense yet temporary. Health Affairs, 34(8), pp.1312-1319.

Kopelovich, S. L., Monroe-DeVita, M., Buck, B. E., Brenner, C., Moser, L., Jarskog, L. F., ... & Chwastiak, L. A. (2021). Community mental health care delivery during the COVID-19 pandemic: practical strategies for improving care for people with serious mental illness. Community Mental Health Journal, 57(3), 405-415.

Liu P, Currie S, Adamyk-Simpson J. What are the most important dimensions of quality for addiction and mental health services from the perspective of its users? *Patient Experience Journal*. 2018; 5(1):106-114. doi: 10.35680/2372-0247.1244.

O'Malley AS, Rich EC, Sarwar R, Schultz E, Warren WC, Shah T, Abrams MK. How Accountable Care Organizations Use Population Segmentation to Care for High-Need, High-Cost Patients. Issue Brief (Commonw Fund). 2019 Jan 1;2019:1-17. PMID: 30645057.

Pierce, B. S., Perrin, P. B., Tyler, C. M., McKee, G. B., & Watson, J. D. (2021). The COVID-19 telepsychology revolution: A national study of pandemic-based changes in US mental health care delivery. American Psychologist, 76(1), 14.

Quinton, J.K., Duru, O.K., Jackson, N. et al. High-cost high-need patients in Medicaid: segmenting the population eligible for a national complex case management program. BMC Health Serv Res 21, 1143 (2021). https://doi.org/10.1186/s12913-021-07116-6

SAS Institute Inc. 2016. SAS/STAT® 14.2 User's Guide. Cary, NC: SAS Institute Inc.

Sterling, S., Chi, F., Weisner, C., Grant, R., Pruzansky, A., Bui, S., & Pearl, R. (2018). Association of behavioral health factors and social determinants of health with high and persistently high healthcare costs. Preventive medicine reports, 11, 154-159.

Treglia, D., Johns, E. L., Schretzman, M., Berman, J., Culhane, D. P., Lee, D. C., & Doran, K. M. (2019). When crises converge: hospital visits before and after shelter use among homeless New Yorkers. Health Affairs, 38(9), 1458-1467.

Viggiano, T., Pincus, HA, and Stephen Crystal. “Care Transition Interventions in Mental Health.” Current Opinion in Psychiatry. 2012; 25(6):551-8. https://doi.org/[10.1097/YCO.0b013e328358df75](https://doi.org/10.1097/yco.0b013e328358df75)

Vickery, K. D., Bodurtha, P., Winkelman, T. N., Hougham, C., Owen, R., Legler, M. S., ... & Davis, M. M. (2018). Cross-sector service use among high health care utilizers in Minnesota after Medicaid expansion. Health Affairs, 37(1), 62-69.

Zhang, Z., Kim, H. J., Lonjon, G., Zhu, Y., & written on behalf of AME Big-Data Clinical Trial Collaborative Group (2019). Balance diagnostics after propensity score matching. Annals of translational medicine, 7(1), 16. https://doi.org/10.21037/atm.2018.12.10