MATHEMATICA Policy Research

REPORT

THIRD ANNUAL REPORT

Evaluating the HCIA - Behavioral Health/Substance Abuse Awards: Third Annual Report

January 23, 2017

Henry Ireys
Tricia Higgins
Ellen Bouchery
Jonathan Brown
Crystal Blyler
Liz Babalola
Michael Barna
Dan Friend
Matt Kehn
Lareina La Flair
Jennifer Lyons
Rachel Miller
Brenda Natzke
Laura Ruttner
Allison Siegwarth
Michaela Vine
Aparajita Zutshi

Submitted to:
Center for Medicare & Medicaid Innovation
Social Science Research Analyst Rapid Cycle Evaluation Group
WB 18-13 7205 Windsor Blvd
Baltimore, MD 21244

Project Officer: Vetisha McClair, Ph.D.
Contract Number: HHSM-500-2010-00025I

Submitted by:
Mathematica Policy Research
1100 1st Street, NE, 12th Floor
Washington, DC 20002-4221
Telephone: (202) 484-9220
Facsimile: (202) 863-1763

Project Director: Henry T. Ireys, Ph.D.
Reference Number: 40271.E20

This page has been left blank for double-sided copying.

CONTENTS

TABLES

This page has been left blank for double-sided copying.

FIGURES

This page has been left blank for double-sided copying.

EXECUTIVE SUMMARY

The Affordable Care Act authorized the Center for Medicare & Medicaid Innovation (CMMI) in the Centers for Medicare & Medicaid Services (CMS) to test innovative health care payment and service delivery models with the potential to lower spending on Medicare, Medicaid, and the Children's Health Improvement Program (CHIP) while maintaining or improving beneficiaries' health and the quality of care they receive. The 107 awardees in the first round of the Health Care Innovation Awards (HCIA) used a broad range of service delivery models. Innovations that succeed in meeting their objectives may lend themselves to implementation on a broad scale. Consequently, rigorous evaluation of the interventions is critical to achieving HCIA's goals.

In the first round of the HCIA initiative, 10 awardees implemented programs that focused primarily on individuals with mental health or substance use disorders (Table ES.1). The three-year award period began in July 2012. For three awardees (Felton, ICSI, and KMHS), the project period ended on June 30, 2015. Four awardees (Feinstein, HLN, MMC, and Vinfen) received three-month, no-cost extensions to close out their awards. Two awardees (FPHNY and ValueOptions) received no-cost extensions of six months, and one (CHCS) received a 12-month, no-cost extension. These extensions gave the awardees time to complete their own evaluations and transition their projects to more sustainable sources of funding.

The 10 projects in this group had some common goals—for example, training staff to coordinate care and using health information technology (IT) to monitor care—but the approaches to achieving them varied widely. They also focused on different subgroups within the broad priority population, such as individuals with schizophrenia or with serious mental illness and a chronic physical condition. The awardees implemented their programs in settings that ranged from primary care practices and mental health clinics to a campus serving the homeless population. The number of participants enrolled in these projects also varied widely, depending on the awardees' specific objectives and recruitment strategies. Half the awardees met their enrollment goals; for three, enrollment levels were substantially less than originally planned.

Findings from rigorous, multifaceted evaluations of these programs should help policymakers and program administrators identify promising approaches to delivering care that could be replicated, expanded, or studied in more depth. Understanding the implementation and impacts of these interventions is important; individuals with mental health and substance use disorders are among the most vulnerable Medicare, Medicaid, and CHIP beneficiaries, and their care is often expensive. They must be part of any enduring solution to improving health care and lowering costs.

Table ES.1. Awardees in the field of behavioral health and substance abuse

Awardee[a]	Overview of intervention	Intervention population	Dollars awarded	Enrollment goal (percent achieved)
Center for Health Care Services (CHCS)	Integrated primary care clinic into behavioral health service setting	Adults in San Antonio, Texas, who are homeless	$4,557,969	260[b] (100)
The Felton Institute (Felton)	Implemented an integrated model of early intervention for psychosis	Patients (ages 14–29) with symptoms of schizophrenia, schizoaffective disorder, or schizophreniform disorder	$4,703,817	140 (100)
Feinstein Institute for Medical Research (Feinstein)	Improved treatment for schizophrenia through training, care management, and new technology	Patients with schizophrenia recently discharged from the hospital and receiving community treatment in one of eight states	$9,380,855	770 (66)
Fund for Public Health in New York (FPHNY)	Provided crisis intervention services to facilitate early engagement and continuity of care, combining community-based care, access to primary care, and peer support	Individuals in Manhattan, Brooklyn, the Bronx, and Queens who have been diagnosed with psychosis or severe mental illness	$17,608,085	2,232 (63)
HealthLinkNow (HLN)	Provided behavioral care services via telehealth to individuals in rural areas that lack access to these services	Patients with behavioral health needs in rural areas with shortages of behavioral health clinicians (Montana, Washington, and Wyoming)	$7,718,636	1,534 (88)
Institute for Clinical Systems Improvement (ICSI)	Implemented collaborative care management for patients with depression and diabetes or cardiovascular disease	High-risk adult patients with Medicare or Medicaid coverage in one of eight states who have active depression and uncontrolled diabetes or cardiovascular disease or both	$17,999,635	2,704 (100)
Kitsap Mental Health Services (KMHS)	Integrated primary care and care for co-occurring physical disorders with mental health services	Adults with severe mental illness and one comorbidity; children with severe emotional disturbance and one physical comorbidity; Kitsap County, Washington	$1,858,437	Not applicable[c]
Maimonides Medical Center (MMC)	Coordinated mental and physical health care through advanced health IT	Adults with serious mental illness living in southwest Brooklyn	$14,842,826	500[d] (100)
ValueOptions (ValueOptions)	Provided support for recovery through reinforcement-based treatment model	Plan members in Massachusetts with two or more detoxification admissions	$2,760,737	1,492[b] (82)
Vinfen Corporation (Vinfen)	Integrated health care services into existing behavioral health outreach teams in community	Individuals in the Boston area with serious mental illness	$2,942,962	400 (54)

Source: Enrollment targets are awardees' self-reported enrollment goals as specified in their applications or quarterly reports to CMMI's technical assistance contractor (the Lewin Group). We obtained award amounts in February 2015 from http://innovation.cms.gov/Files/x/HCIA-Project-Profiles.pdf.

[a] In this report, we use the acronym or name abbreviations indicated in parentheses to designate the awardees.

[b] Intervention group participants only.

[c] KMHS did not specify enrollment goals. Instead, it identified cohorts of individuals within its service population for whom it provided quantitative data on outcome measures.

[d] Direct participants only. MMC's project also included 7,000 Medicaid-enrolled indirect participants.

Evaluation goals and methods

In September 2013, CMMI contracted with Mathematica Policy Research to evaluate the 10 projects in Table ES.1. Our evaluation, which concludes in September 2017, has three broad and interrelated goals.

First, we have responded to a series of specific evaluation questions that CMMI asked us to address. Previous reports—the first and second annual report—provide the results of early analyses conducted to answer these questions.[1] In these reports, we paid particular attention to CMMI's four core quantitative measures of program effectiveness:

- Total Medicare and Medicaid expenditures
- Hospitalization rates
- Hospital readmission rates
- Rates of emergency department (ED) use

The second goal is to identify general lessons learned about successful projects by synthesizing findings for the different awardees. For example, in our previous reports, we discussed common challenges that awardees faced in implementing their projects and the solutions they developed to address them.

In this third annual report, we focus on the third goal: "telling the story" of each awardee by describing its program objectives, implementation experiences, and participants' outcomes, using CMMI's four core measures to the extent possible. To accomplish this goal, we used a mixed-methods approach, combining quantitative and qualitative analyses. Because of differences among awardees, we essentially conducted a separate evaluation for each of them. This report's 10 awardee-specific chapters present the key findings from our evaluation of each program.

We drew on five types of data for our evaluation:

- Enrollment data obtained from CMMI's technical assistance contractor (the Lewin Group)
- Medicare and Medicaid claims data obtained either through CMS or states
- Survey and administrative data obtained directly from the awardees
- Data from our workforce survey, conducted in 2014 and 2015

[1] These reports are available at https://innovation.cms.gov/Files/reports/HCIA-BHSA-FirstEvalRpt.pdf and https://innovation.cms.gov/Files/reports/hcia-bhsa-secondevalrpt.pdf.

- Qualitative data from key informant interviews conducted during site visits in 2014 and 2015 and focus groups hosted in 2015

Our quantitative analyses were the most rigorous possible for all 10 awardees. Ideally, we would have identified a control or comparison group for each awardee, thereby allowing us to examine what might have happened had the HCIA-funded program not been implemented (that is, the counterfactual). In fact, working closely with the awardee staff, we were able to identify such groups for six awardees: CHCS, FPHNY, HLN, KMHS, MMC, and ValueOptions (Table ES.2). For five of these awardees, we were able to conduct difference-in-differences analyses—an approach that allows for reasonably strong conclusions about a program's impact on outcomes of interest.[2] (We also expect to conduct difference-in-differences analyses for ValueOptions. Delays in obtaining appropriate data files from the awardee meant that we could not complete these analyses in time to include our findings in this report.)

Although we were able to apply the difference-in-differences analytic approach to estimate program impacts on key outcomes, we were able to focus only on certain subgroups of participants for some awardees (Table ES.2). For example, for HLN, we could only analyze the program's impacts on Medicare beneficiaries because of the lag in availability of Medicaid data and slow enrollment progress during the first two years of the program. For FPHNY, our analyses included only 17 percent of all participants—a limitation imposed by participant characteristics (about 40 percent were uninsured) and data availability (the program was unable to provide valid Medicaid identifiers for the 34 percent of the participants they indicated were enrolled in Medicaid).

In the individual awardee chapters, we note several data limitations that affected our analyses. For example, Medicaid claims files do not include psychiatric stays for adults in institutions for mental disease (IMDs). This limitation means that quantitative analyses may undercount hospitalizations and readmissions for Medicaid-enrolled program participants.

For some awardees, the number of individuals included in our analyses was large relative to all participants, but small for analytic purposes. For example, CHCS enrolled 261 individuals into its intervention group and a similar number into its control group. Although we were able to include most of these participants in our analyses, the size of the intervention group meant the program would have had to be extremely effective for us to detect statistically significant changes relative to the control group during the study period. In this case, as for some other awardees with low enrollment, we faced the danger of concluding that the program had no effect when in fact we might have detected an effect with a larger number of participants or a longer study period.

Overall, we urge readers to interpret our conclusions carefully in light of these and other limitations noted in the report.

[2] A difference-in-differences analysis calculates a program impact by comparing the average change over time in the selected outcome variable for the treatment group with the average change over time for the control or comparison group. Thus, it examine how the difference between the two groups differs over time.

Table ES.2. Evaluation design features for behavioral health/substance abuse awardees

Awardee	Intervention group	Control or comparison group	Data sources for outcomes	Outcome measures	Key limitations
Awardees with control or comparison group					
CHCS	Homeless adults	Awardee randomly assigned homeless adults to control group	Awardee-provided baseline and follow-up assessments	Mental health (SF-6D, Brief Symptom Index, et alia) and physical health (weight loss and blood pressure) status measures	Data on expenditures, hospitalizations, and ED visits not available
FPHNY	Adults in New York City diagnosed with psychosis or severe mental illness	Similar New York State Medicaid beneficiaries	New York State Medicaid data files, 2010–2015	Expenditures hospitalizations, ED visits	Medicaid analysis only; small proportion of participants in analyses
HLN	Adults with behavioral health needs in areas with shortages of behavioral health clinicians	Similar Medicare FFS beneficiaries, residing in 10 nearby states	Medicare data files, 2010–2015	Expenditures hospitalizations, ED visits	Medicare analysis only; small proportion of participants in analyses
KMHS	Patients served by community mental health center in Washington State	Similar Medicare FFS beneficiaries served by other mental health providers in Washington State; similar Medicaid beneficiaries in Washington State	Medicare data files, 2010–2015; Medicaid data files, 2011–2014	Expenditures, hospitalizations, ED visits, office visits	Expenditures not available for Medicaid participants
MMC	Adults with SMI living in southwest Brooklyn	Similar Medicare FFS beneficiaries living in three similar cities; no comparison group for Medicaid population	Medicare 2010–2015; Medicaid data files 2010–2015	Expenditures, hospitalizations, readmissions, ED visits	Medicaid participants not included in impact analyses
ValueOptions	Plan members with two or more detoxification admissions in Massachusetts	Awardee-selected comparison group of similar members receiving care at non-intervention sites	Awardee-provided Medicaid claims and clinical assessment data	Expenditures, residential stays, days of intensive day treatment	Because of delays in receiving data from awardee, we will provide results of impact analyses in addendum

Awardee	Intervention group	Control or comparison group	Data sources for outcomes	Outcome measures	Key limitations
Awardees without a control or comparison group					
Feinstein	Patients with schizophrenia, recently discharged from hospital and receiving care at community treatment centers	None[a]	None[b]	None[b]	Not Applicable
Felton	Youth (14–29) with schizophrenia, schizoaffective disorder, or schizophreniform disorder	None[a]	Electronic Health Record (EHR) extracts including hospitalization and ED visit data entered by Felton staff based on counties' mental health department data.	Hospitalizations, ED visits	Lack of comparison group; small sample; analysis limited to 12-months; lack of data on stability and independence, two of Felton's primary goals for participants
ICSI	High-risk adult patients with Medicare or Medicaid coverage who have active depression plus uncontrolled diabetes or cardiovascular disease or both	None[a]	Medicare 2010–2015; Electronic Health Records (EHR) 2013-2015	For Medicare enrollees: ED visits, hospitalizations and hospital readmissions, and total expenditures; For all participants: Depression score (PHQ-9), blood pressure control, and HbA1c control	Lack of comparison group; Medicare analysis only; Inconsistent recording of EHR outcome metrics
Vinfen		None[a]	None[c]	None[c]	Not Applicable

Notes: For each awardee for which we have a comparison or control group, we conducted multivariate longitudinal analysis of intervention and comparison or control group outcomes or costs, controlling for factors specific to that awardee using a "difference-in-differences" analysis. We excluded Medicare Advantage participants from our Medicare analyses because expenditures and utilization data for this population are not included in the available Medicare administrative data. Similarly, we excluded Part D pharmacy services and expenditures from our analyses due to lack of available data. Only CHCS had a randomized control group.

[a] We were unable to identify and obtain sufficient data to develop an appropriate comparison group for this awardee.

[b] We were unable to use Medicaid or Medicare data to assess pre-post changes or program impacts for Feinstein due to insufficient sample sizes and challenges associated with obtaining accurate and complete data from clinical sites.

[c] We were unable to conduct quantitative analysis of the effects of Vinfen's program because we did not receive sufficient data. Vinfen conducted various descriptive and trend analyses of claims data using Medicaid claims data from MassHealth. Vinfen sent these findings in a set of PowerPoint slides to Mathematica and CMMI in early 2016.

Because of various delays, we were unable to obtain all the data we needed in time to conduct certain analyses. Specifically, for four of the awardees for which we conducted or plan to conduct impact analyses (FPHNY, KMHS, MMC, and ValueOptions), the story remains incomplete. We expect to (1) expand on our analyses using additional data covering longer post-intervention periods or more outcome measures for FPHNY, KMHS, and MMC and (2) conduct our planned impact analysis for ValueOptions, given that we have recently obtained corrected data files. For all of these analyses, we will report our results in an addendum to this report, which we will submit to CMMI in March 2017.

For the other six awardees (CHCS, Felton, Feinstein, HLN, ICSI, and Vinfen), this report provides final evaluation results. As noted, we were able to complete a difference-in-differences analysis for CHCS. We also completed a difference-in-difference analysis for HLN for the Medicare population.[3] We completed pre-post analyses for Felton, Feinstein, and ICSI, focusing on intervention groups only (that is, no comparison groups were available). Vinfen was unable to send sufficient, individual-level quantitative data, including data for a comparison group, that we could use for an impact analysis. As a result, we based our evaluation primarily on (1) qualitative data we obtained through interviews and document review and (2) demographic data that Vinfen provided on program participants who consented to have their data used for research purposes.

Several awardees plan to conduct their own analyses of data they have collected on program participants (and, in some cases, comparison groups) and publish reports about their programs' outcomes. Their analyses may focus on different evaluation questions and use different data sources and measures than those we use in this report. For example, several awardees may examine program outcomes on clinical indices of health and mental health status by using data from medical records or participant surveys. In contrast, we focus primarily on service use and cost outcomes—outcomes that can be assessed similarly across awardees and are priorities for CMMI. Because the awardees are using different sources of data and focusing on different outcomes, they may reach conclusions that differ from the ones we present in this report.

Key findings on CMMI's four core measures

As noted, CMMI is particularly interested in program effects on total Medicare and Medicaid expenditures and three measures of service use: hospitalizations, readmissions, and emergency department (ED) visits. For participants in this group of programs, readmissions were quite rare; as a result, impact estimates would have been unreliable. Hence, our service use calculations focused on hospitalizations and ED visits.[4] Table ES.3 highlights key findings (described further in the awardee-specific chapters) for three measures: Total expenditures per person for each month enrolled in the designated program (Medicare or Medicaid), aggregate number of hospitalizations during the study period, and aggregate number of ED visits during the study period. The study period varied by awardee depending on availability of data. For FPHNY, the study period was 12 months. For the other awardees, the study period was up to two years.

[3] Due to implementation of the Transformed Medicaid Information Management System (T-MSIS), sufficient data were unavailable to support analysis of HLN Medicaid participants.

[4] MMC is an exception to this pattern. We were able to calculate readmissions for Medicare-enrolled participants and found no effects. The chapter on MMC provides further details about this analysis.

Table ES.3. Impacts of five awardees' programs on CMMI core measures: Key findings[a]

			Change relative to comparison group	
		Measure	Aggregate (expenditures in thousands)	Per beneficiary per month
FPHNY	Mobile crisis teams			
	Worked with individuals after a mental health crisis to develop and implement individualized action plan	Medicaid expenditures for Medicaid participants	$1,794 higher*	$1,838 higher*
		Hospitalizations for Medicaid participants	120 more*	0.12 more*
		ED visits for Medicaid participants	1 less	Same
FPHNY	Crisis respite centers			
	Provided alternative to hospitalization for individuals who needed temporary residential or respite care	Medicaid expenditures for Medicaid participants	$3,143 lower*	$1,797 lower*
		Hospitalizations for Medicaid participants	374 less*	0.21 less*
		ED visits for Medicaid participants	126 more	0.07 more
HLN				
	Provided behavioral care services via telehealth to individuals in areas that lack access to these services	Medicare expenditures for participants in FFS Medicare	$260 higher	$128 higher
		Hospitalizations for participants in FFS Medicare	101 less	0.05 less
		ED visits for participants in FFS Medicare	20 more	0.01 more
KMHS				
	Integrated primary care and care for co-occurring physical disorders with mental health services	Medicare expenditures for participants in FFS Medicare	$5,144 lower*	$266 lower*
		Hospitalizations for participants in FFS Medicare	297 less*	0.02 less*
		ED visits for participants in FFS Medicare	546 less*	0.03 less*
		Hospitalizations for Medicaid participants	199 less*	0.00 less*
		ED visits for Medicaid participants	1,592 more*	0.03 more*
MMC				
	Coordinated mental and physical health care through advanced HIT	Medicare expenditures for participants in FFS Medicare	$26 lower	$3.44 lower
		Hospitalizations for participants in FFS Medicare	39 less	0.01 less
		ED visits for participants in FFS Medicare	71 less	0.01 less

Source: Analyses of Medicare and Medicaid data. See awardee-specific chapters for details.

* Indicates statistical significance at the p<0.10 level.

[a] Due to insufficient data, we did not conduct impact analyses for CMMI core measures for Felton and ICSI, but we did conduct pre-post analyses. In the pre-post analyses for Felton, rates of hospitalizations and ED visits decreased from baseline to the first intervention period and continued to decline slightly through the second baseline period. For ICSI, pre-post analyses showed no significant differences between average baseline and post-intervention expenditures; the hospitalization rate trended upward from the beginning of the baseline period through the first

intervention period and then fell consistently. The ED visit rate climbed steadily until it began falling in the 6-months prior to intervention start and continued to fall throughout the intervention period.

[b] For MMC, pre-post analyses for Medicaid-enrolled participants showed significant declines in inpatient stays and ED visits in the post-intervention period. Despite these declines, total expenditures per enrolled month continued to increase in the post-intervention period at a similar rate to that observed in the baseline period. We are conducting additional analyses to examine this pattern further.

Total expenditures. Many policymakers and program administrators are interested in knowing whether HCIA programs lower Medicare or Medicaid spending. Addressing this issue with a reasonable degree of certainty requires comparing expenditures for program participants with expenditures for a control or comparison group. We were able to develop comparison groups and obtain data on Medicare or Medicaid expenditures for four awardees (FPHNY, HLN, KMHS, and MMC).

Of these four awardees, we estimated program impacts on Medicare expenditures for three, and found the following results:

- Relative to the comparison group, HLN's program appears to have had no effect on expenditures for the 25 percent of its participants who were covered under fee-for-service (FFS) Medicare.
- Relative to the comparison group, KMHS' program appears to have reduced total expenditures for the 13 percent of its participants covered under FFS Medicare by an average of $266 per person per month.
- Relative to the comparison group, MMC's program appears to have had no effect on expenditures for the 7 percent of participants covered under FFS Medicare.

Given limitations on data availability, we were able to estimate program impacts on Medicaid expenditures only for FPHNY. This awardee's program had two components, each of which had different effects on Medicaid expenditures:[5]

- One component, providing care through crisis respite centers as an alternative to hospitalizations, appears to have reduced Medicaid expenditures for the 11 percent of participants included in the analysis by $1,797 per person per month of Medicaid enrollment, relative to the comparison group.
- The second component, providing care through mobile crisis teams, appears to have increased Medicaid expenditures for the 6 percent of participants included in the analysis by $1,838 per person per month of Medicaid enrollment, relative to the comparison group.

Hospitalizations. We calculated program impacts on rates of hospitalizations for participants in four HCIA programs and found no significant effects for HLN or MMC (for their FFS Medicare participants). We found the following significant results for FPHNY and KMHS:

[5] Our analysis for FPHNY only addressed the first 12 months following program participation. We were unable to assess longer term impacts because data became available too late for this report. We will provide results of additional analyses of program effects on the Medicaid population in an addendum to this report.

- During the study period, FPHNY crisis respite centers appear to have reduced hospitalizations by 374 for Medicaid participants included in the analysis relative to the comparison group.
- During the study period, FPHNY mobile crisis teams appear to have increased hospitalizations by 120 for Medicaid participants included in the analysis relative to the comparison group.
- During the study period, KMHS' program appears to have reduced hospitalizations for Medicare participants by 297 and for Medicaid participants by 199 relative to the comparison groups.

ED visits. We calculated program impacts on ED visit rates for participants in four awardees' programs and found no significant impacts for FPHNY, HLN, or MMC (for their FFS Medicare participants). We did find significant impacts for KMHS. Specifically, during the study period, KMHS appears to have reduced ED visits for Medicare participants by 546 and increased ED visits for Medicaid participants by 1,592 relative to comparison groups.[6]

Other important findings

Although we focused much of our attention on CMMI's four core measures, our evaluations of each awardee gave us many other insights into different outcomes, such as mental health status and access to care; use of HIT to support care coordination; implementation challenges and corresponding solutions; and factors contributing to workforce satisfaction. These findings, which are likely to be of great interest to the field overall, emerged from the analysis and synthesis of our quantitative, qualitative, and survey data and are described in the report's awardee-specific chapters. Here, we note three cross-awardee observations that may be of particular interest to policymakers and program administrators.

The challenges encountered in implementing a program are broadly linked to the phase of innovation the program is in. Each of the 10 programs we evaluated had its own distinctive history with respect to its place in the long-term process of developing a new model of service delivery. For example, MMC's extensive history of learning from and understanding the perspectives of its partners substantially helped initial implementation of its program; as one staff member noted, "By the time we really launched this program, we had a group of organizations and frankly, the leadership of those organizations really liked one another and worked well together, and I think that made a huge, huge difference."

To provide a framework for understanding the administrative context of these programs, we identified five phases of innovation and located each awardee's progress over the course of the HCIA within these phases (Table ES.4).

Some awardees, including CHCS and FPHNY, implemented programs that were essentially new efforts for them, even though the program's underlying theories of action were familiar to the staff responsible for implementing the new program. Generally speaking, these awardees

[6] We also were able to calculate impacts on office visit rates and found that, during the study period, KMHS's program reduced office visits by 2,560 for Medicare participants relative to the comparison group.

faced the challenges of developing new procedures to support recruitment and program operations. Other awardees, such as Feinstein and ICSI, had refined and matured their models locally through prior work and used HCIA funding to expand the model to other clinical locations, often quite distant from the awardee's home site. For these awardees, implementation challenges typically involved the development of tracking and reporting mechanisms to monitor program fidelity, given that these other sites usually had to adapt program parameters to fit local conditions. In the awardee-specific chapters, we describe each program's implementation challenges and corresponding solutions.

Table ES.4. Awardees' position in the phases of innovation

	Phases of innovation				
Awardees	Formulate idea	Pilot test to assess feasibility	Implement revised program and evaluate locally	Refine, implement broadly, evaluate across multiple sites	Adopt, institutionalize, sustain
CHCS	√	√			
Felton	√	√	√		
Feinstein	√	√	√	√	
FPHNY	√	√	√		
HLN	√	√	√	√	
ICSI	√	√	√	√	
KMHS	√	√	√	√	√
MMC	√	√	√	√	√
ValueOptions	√	√	√		
Vinfen	√	√	√		

Source: Awardee-specific analysis; details on the awardee's place in the innovation phase are in the awardee-specific chapters.

In most cases, substantial resources are required to train and support the workforce needed to implement new models of service delivery. The first round of the HCIA awards emphasized the role of the workforce, and for good reason: "Innovation" in health care typically means that some providers have to change their behaviors and, for many, these changes are difficult. Awardees in this group (Feinstein, FPHNY, ICSI, KMHS, and ValueOptions, for example) underscored the importance of making a substantial commitment to helping providers (1) integrate new procedures into clinical practice (such as alerting patients to newly available telehealth opportunities), (2) relate to new members of the clinical team, such as peer support specialists, and (3) think beyond traditional mental health boundaries by, for example, considering patients' physical health conditions.

KMHS is a good example of this general observation. For this awardee, the leaders recognized—based on their previous efforts to integrate consideration of substance use disorder into a clinical environment used to focusing solely on mental health disorders—that substantial training and support were needed to change "business as usual." As a result, this awardee used the first six months of the HCIA funding to train staff throughout the entire organization in the procedures needed to implement the program. In addition, KMHS continued widespread training

efforts after the program began to ensure key procedural changes were maintained and to address related implementation issues. KMHS's program was one of the most successful in this group of awardees with respect to having positive effects on expenditures and service use—an outcome that was likely due, in part, to the comprehensive and sustained training that was a fundamental component of its program model.

In many instances, the decision to sustain a program is made before rigorous quantitative evidence is available on program effects or is not based solely on such evidence. Most program developers and directors are deeply committed to the programs' underlying conceptual models. In many cases, the HCIA represents one of many sequential (or, in a few instances, simultaneous) funding sources that advance the development and spread of these models. From program leaders' perspectives, the award's value lies, in part, in giving program staff the time and money to work through implementation challenges and find additional funds to keep refining their programs. Results from rigorous evaluation results may come too late in the process and, if they are equivocal or negative, may pose obstacles to obtaining more funding. For example, MMC built sustainability efforts into its operational plan from the outset of the award, suggesting that program leaders were determined to continue the program in some form regardless of evaluation results. Positive findings from MMC's internal monitoring of program operations helped program leaders secure ongoing support; our preliminary evidence that this program had no effect on total Medicare expenditures and may have increased total Medicaid expenditures appears to have had negligible influence on their decision-making process. In contrast, KMHS leaders struggled to find support for some program components, despite strong evidence suggesting the program had positive effects on expenditures and service use.

Conclusions

The HCIA awardees in behavioral health and substance abuse implemented programs with the common aim of improving health outcomes and service delivery and reducing costs of care for individuals with mental illness and substance use disorders. Although the overall evidence is mixed, our evaluation indicates that some programs achieved some of these goals. These results, coupled with an exceptionally broad array of "lessons learned," could give staff at CMMI and other federal and state agencies ideas for initiatives that build on the work of these awardees. For example, further synthesis of awardees' experiences around integrating mental health and primary care services for individuals with serious mental illness could contribute to other initiatives related to this topic, such as efforts underway through the Innovation Accelerator Program.

These awardees received more than $80 million through an initiative that has yielded substantial experience in implementing different approaches to improving care for individuals with mental health and substance use disorders. For most awardees, key program elements will be sustained in some fashion, potentially bringing further returns to the government's substantial investment.

I. INTRODUCTION

A. The HCIA initiative

The Affordable Care Act authorized the Center for Medicare & Medicaid Innovation (CMMI) in the Centers for Medicare & Medicaid Services (CMS) to test innovative health care payment and service delivery models that have the potential to lower spending on Medicare, Medicaid, and the Children's Health Insurance Program while maintaining or improving beneficiaries' health. As part of CMMI's efforts, the first round of the Health Care Innovation Awards (HCIA) initiative gave 107 organizations the funding to implement a broad range of service delivery models (https://innovation.cms.gov/initiatives/Health-Care-Innovation-Awards/). The models address groups of beneficiaries with poor clinical outcomes or heavy utilization of services. CMMI will examine the evidence about the implementation and outcomes, and might promote replication of the most promising models of care. Consequently, rigorous evaluation of the HCIA initiative is vital for CMMI to achieve its mission.

B. Overview of the behavioral health awardees

In the first round of the HCIA initiative, 10 awardees implemented programs focused primarily on individuals with mental health and substance use disorders (Table I.1). The projects shared some cross-cutting themes (for example, training staff to coordinate care and using information technology to plan or monitor services) but they involved different subgroups of the priority population—such as individuals with schizophrenia or with serious mental illness and a chronic physical condition. The awardees implemented their programs in a range of community-based settings, including primary care practices and mental health clinics. The number of direct participants varied from 140 to more than 2,700 across the 10 awardees.

The three-year awards began in early July 2012. For three awardees (The Felton Institute, Institute for Clinical Systems Improvement, and Kitsap Mental Health Services), the project period ended three years later, on June 30, 2015. Four awardees (Feinstein Institute for Medical Research, HealthLinkNow, Maimonides Medical Center, and Vinfen Corporation) received four-month, no-cost extensions to close out their programs. Two awardees (Fund for Public Health in New York and ValueOptions, Inc.) received no-cost extensions of six months; one (Center for Health Care Services) received a 12-month, no-cost extension. The extensions for the latter three allowed the awardees to complete their own evaluations and transition their projects to more sustainable sources of funding.

Several awardees conducted analyses of data that they collected on program participants (and, in some cases, control or comparison groups) and plan to publish, or have published, reports about their programs' outcomes. Awardees' own analyses may focus on evaluation questions, use data sources, and calculate outcome measures that are quite different from those in this report. For example, several awardees examined the outcomes of their programs on clinical indices of health and mental health status by using data from medical records, participant surveys, or other sources. In contrast, we focused primarily on service use and cost outcomes—outcomes that are priorities for CMMI and that we can assess similarly for all the awardees. Because the awardees used different data sources and focused on different outcomes variables, they may reach conclusions that differ substantially from the ones presented in this report.

Table I.1. Behavioral health and substance abuse awardees

Awardee (name abbreviation used in report)	Overview of program (dollars awarded[a])	Program population (target number of direct participants[b])
Center for Health Care Services (CHCS)	Integrated primary care clinic into behavioral health service setting ($4,557,969)	Adults in San Antonio, Texas, who are homeless (260)
The Felton Institute (Felton)	Implemented an integrated treatment model to improve intervention for psychosis ($4,703,817)	Patients (ages 14–29) with symptoms of schizophrenia, schizoaffective disorder, or schizophreniform disorder (140)
Feinstein Institute for Medical Research (Feinstein)	Improved treatment of schizophrenia through training, care management, and new technology ($9,380,855)	Patients with schizophrenia who were recently discharged from the hospital and are receiving care at a community intervention center in one of eight states (770)
Fund for Public Health in New York (FPHNY)	Provided crisis intervention services to facilitate early engagement and continuity of care, combining community-based care with access to primary care ($17,608,085)	Individuals in Manhattan, Brooklyn, the Bronx, and Queens who have been diagnosed with psychosis or severe mental illness (2,232)
HealthLinkNow (HLN)	Provided behavioral care services via telemedicine to individuals in rural areas who lack access to these services ($7,718,636)	Patients with behavioral health needs in rural areas in Montana, Washington, and Wyoming with shortages of behavioral health clinicians (1,534)
Institute for Clinical Systems Improvement (ICSI)	Implemented collaborative care management model for patients with depression and diabetes or cardiovascular disease ($17,999,635)	High-risk adult patients with Medicare or Medicaid coverage in one of eight states who have depression and diabetes or cardiovascular disease (2,704)
Kitsap Mental Health Services (KMHS)	Integrated primary health care for individuals with severe mental illness ($1,858,437)	Patients served by community mental health center in Kitsap County, Washington (not applicable[c])
Maimonides Medical Center (MMC)	Coordinated mental and physical health care through advanced health information technology ($14,842,826)	Adults with serious mental illness living in southwest Brooklyn (500)
ValueOptions, Inc. (ValueOptions)	Provided support for recovery through reinforcement-based treatment model ($2,760,737)	Plan members in Massachusetts with two or more detoxification admissions (1,492)
Vinfen Corporation (Vinfen)	Integrated health care services into existing behavioral health outreach teams in community ($2,942,962)	Individuals in Boston with serious mental illness (470)

Note: In this report, we usually use the acronym or name abbreviations indicated in parentheses to designate the awardees. In subsequent tables, we list awardees in alphabetical order based on their full names, as we do here.

[a] Dollar amounts accessed from http://innovation.cms.gov/Files/x/HCIA-Project-Profiles.pdf.

[b] Awardees' self-reported enrollment goals as specified in their applications or quarterly reports to CMMI's technical support contractor (the Lewin Group).

[c] KMHS did not define a specific enrollment target for its Race to Health! program because KMHS staff intended the program to reach all patients who used KMHS' outpatient services during the study period.

C. Evaluation goals

In September 2013, CMMI contracted with Mathematica Policy Research to evaluate the 10 projects in Table I.1. Our evaluation, which concludes in September 2017, has three broad and interrelated goals.

First, using diverse sources of data, we have responded to a series of specific evaluation questions that CMMI asked us to address. Previous reports—the first and second annual report—provide the results of early analyses conducted to answer these questions (see https://innovation.cms.gov/Files/reports/HCIA-BHSA-FirstEvalRpt.pdf and https://innovation.cms.gov/Files/reports/hcia-bhsa-secondevalrpt.pdf). In these reports, we paid particular attention to CMMI's four core quantitative measures of program effectiveness:

- Total Medicare and Medicaid expenditures
- Hospitalization rates
- Hospital readmission rates
- Rates of emergency department (ED) use

The second goal involves identifying general lessons learned about successful projects based on a synthesis of findings across awardees. For example, in our previous reports, we discussed common challenges that awardees faced in implementing their projects and the solutions they developed to address them.

In this report, we focus on the third goal: "telling the story" of each awardee by describing their program objectives, implementation experiences, and outcomes, using CMMI's four core measures to the extent possible.

D. Evaluation methods

Our evaluation used a mixed-methods approach that combined quantitative and qualitative analyses. For our quantitative analyses, we conducted as rigorous an evaluation as possible for all 10 awardees. Ideally, we would have identified a control or comparison group for each awardee, thereby providing an opportunity to understand what might have happened had the HCIA-funded program not been implemented (that is, the counterfactual). In fact, working closely with the awardee staff, we were able to identify control or comparison groups for only five awardees: CHCS, FPHNY, HLN, KMHS, and MMC (Table I.2). As a result, for these awardees, we conducted difference-in-differences analyses—an approach that allows for reasonably strong conclusions about a program's impact on outcomes of interest.[7] (We expected to conduct difference-in-differences analyses for ValueOptions, but problems in the initial data files that we received from the awardee and delays in obtaining corrected files meant we were unable to complete these analyses in time to include our findings in this report.)

[7] Difference-in-differences is a statistical technique which calculates the effect of an intervention on an outcome by comparing the average change over time for the outcome on the treatment group to that for a comparison or control group. It is intended to mitigate the effect of differences between the treatment and comparison or control group that are unrelated to the intervention.

Table I.2. Evaluation design features for six awardees

Awardee	Intervention group	Control or comparison group	Data sources for outcomes	Outcome measures	Key limitations
CHCS	Homeless adults	Awardee randomly assigned potential participants to control group	Awardee-provided baseline and follow-up assessments	Mental health status (for example, SF-6D, Brief Symptom Index) and physical health status (weight loss and blood pressure)	Data on expenditures, hospitalizations, and ED visits not available
FPHNY	Adults in NYC diagnosed with psychosis or severe mental illness	Similar New York State Medicaid beneficiaries	New York State Medicaid data files, 2010–2015	Expenditures, hospitalizations, ED visits	Medicaid analysis only; small proportion of participants in analyses
HLN	Adults with behavioral health needs in areas with shortages of behavioral health clinicians	Similar Medicare beneficiaries residing in 10 nearby states	Medicare data files, 2010–2015	Expenditures, hospitalizations, ED visits	Medicare analysis only; small proportion of participants in analyses
KMHS	Patients served by community mental health center in Washington State	Similar Medicare beneficiaries served by other mental health providers in Washington State; similar Medicaid beneficiaries in Washington State	Medicare data files, 2010–2015; Medicaid data files, 2011–2014	Expenditures, hospitalizations, ED visits, office visits	Expenditures not available for Medicaid-enrolled participants
MMC	Adults with serious mental illness living in southwest Brooklyn	Similar Medicare beneficiaries living in three similar cities; no comparison group for Medicaid population	Medicare 2010–2015; Medicaid data files 2010–2015	Expenditures, hospitalizations, readmissions, ED visits	Medicaid participants not included in impact analyses
ValueOptions	Plan members with two or more detoxification admissions in Massachusetts	Awardee-selected comparison group of similar members receiving care at other sites	Awardee-provided Medicaid claims and clinical assessment data	ED visits, residential stays, days of intensive day treatment, and total expenditures	Because of delays in receiving data from the awardee, we will provide results of impact analyses in an addendum

Notes: For each awardee for which we have a comparison or control group, we conducted multivariate longitudinal analysis of intervention and comparison or control group outcomes or costs controlling for factors specific to that awardee using a "difference-in-differences" paradigm. Medicare Advantage participants are excluded from our Medicare analyses because expenditures and utilization data for this population are not included in the available Medicare administrative data. Similarly, Part D pharmacy services and expenditures are excluded from our analyses because we did not have sufficient data.

Although we were able to apply the difference-in-differences analytic approach to estimate program impacts on key outcomes, we could focus only on certain subgroups of participants for some of these awardees—as Table I.2 notes and as we explain in the awardee-specific chapters. For example, for HLN, we could conduct analyses only for the program's Medicare beneficiaries because of the lag in availability of Medicaid data and slow enrollment progress in the first two years of the program. For FPHNY, our analyses included only 17 percent of all participants—a limitation related to participant characteristics (about 40 percent were uninsured) and data availability (the program did not provide valid Medicaid identifiers for 34 percent of the participants it indicated were Medicaid enrolled).

For some awardees, the number of individuals included in our analyses was proportionately large relative to all program participants but small for analytic purposes. For example, CHCS enrolled 261 individuals into its intervention group and about the same number into its control group. Although we were able to include most of these individuals in our analyses, the group size meant that the program would have had to be extremely effective for us to detect significant changes in the intervention group relative to the control group. In this case, we faced the possibility of concluding that the program had no effect when, in fact, we might have detected an effect with a larger number of participants.

Overall, we urge readers to interpret our conclusions carefully in light of these and other limitations that we note in each chapter.

Because of various delays, we were unable to obtain all the data we needed in time to conduct certain analyses for this report. Specifically, for four awardees (FPHNY, KMHS, MMC, and ValueOptions), the story remains incomplete. We expect to (1) extend our analyses using additional data covering a longer post-intervention period or more outcome measures for FPHNY, KMHS, and MMC and (2) conduct our planned impact analysis for ValueOptions, given that we have recently obtained corrected data files. For all of these analyses, we will report our results in an addendum to this report, which we will submit to CMMI in March 2017.

For the other six awardees (CHCS, Felton, Feinstein, HLN, ICSI, and Vinfen), this report provides final evaluation results. As noted, we completed a difference-in-differences analysis for CHCS. We also completed a difference-in-difference analysis for HLN for the Medicare population.[8] We completed pre-post analyses for Felton, Feinstein, and ICSI, focusing on their intervention groups only (that is, no comparison groups were available). Vinfen was unable to send sufficient, individual-level quantitative data, including data for a comparison group, that we could use for an impact analysis. As a result, we based our evaluation primarily on (1) qualitative data we obtained through interviews and document review and (2) demographic data that Vinfen provided on program participants who consented to have their data used for research purposes.

Overall, we drew on five types of data for our evaluation:

1. **Enrollment data**. We obtained enrollment data from CMMI's technical assistance contractor (the Lewin Group) and used these numbers to determine each awardee's progress

[8] Due to implementation of the Transformed Medicaid Information Management System, sufficient data were unavailable to support analysis of HLN Medicaid participants.

toward its final enrollment goal. Awardees specified an enrollment goal in their early reports; many revised these goals as they implemented their projects.

2. **Medicare and Medicaid claims data.** We used these data primarily to estimate program impacts on CMMI's four core measures and other outcomes. We were unable to obtain these data for all awardees because some could not provide the information we needed to identify participants' Medicaid or Medicare data or because of lags in the availability of Medicaid data. In some cases, we were able to obtain these data for only a small number of participants and, as a result, did not conduct analyses because they would have yielded unreliable findings.[9]

3. **Awardee's survey and administrative data.** Some awardees were able to provide adequate data from their electronic health records, surveys, or clinical assessments of participants. Generally, we used awardee data to assess program implementation (for example, by identifying what services were delivered to whom). In some cases, we used them to assess outcomes, such as changes in participants' symptoms and functional status. In the case of one awardee, CHCS, all the data we used to evaluate the program came from the awardee.

4. **Workforce surveys.** We conducted a workforce survey in 2014 and 2015 that provided information about staff burnout and stress, job satisfaction, and perceptions of training and job support, and included findings from our analysis of survey data in our second annual report and the eighth quarterly report. In this report, we draw on survey results only to help identify lessons learned from the overall evaluation.

5. **Qualitative data from interviews and focus groups.** We conducted interviews with key informants during site visits to awardees in spring 2014 and 2015. During these site visits, we met with awardee leaders and staff, program participants, and other stakeholders to learn more about the implementation process and their experiences with various components of the programs. During the visits in 2015, we also conducted focus groups with staff and, where possible, with program participants and nonparticipants. We held focus groups with these two groups to understand differences in their experiences with care.

E. Road map to the report

The following chapters (II–XI) contain findings from our evaluation of the programs that the awardees implemented. In each awardee narrative, we:

- Explain the program's implementation context
- Report the program's process through specific phases of innovation
- Describe enrollment outcomes and participants' demographic characteristics
- Define the methods and data we used for the evaluation
- Present results of our quantitative and qualitative analyses
- Highlight findings from our workforce survey

[9] Medicaid claims files do not include psychiatric stays for adults in institutions for mental disorders (IMDs). Consequently our quantitative analyses may undercount hospitalizations and readmissions for Medicaid-enrolled program participants.

- Describe whether and how the awardee is sustaining its program
- Conclude with a summary of lessons learned

The technical appendix (Appendix A) provides additional details about our quantitative analyses.

This page has been left blank for double-sided copying.

II. CENTER FOR HEALTH CARE SERVICES

Findings from Mathematica's evaluation of CHCS's Project HEALTH

- Our analysis of quantitative data indicated that Project HEALTH had no impact on select measures of mental health status; however, the result may reflect major analytic limitations.
- Our analysis of qualitative data suggested several positive effects on participants' health and well-being, and revealed that participants used more medical and nonmedical services than they did before the program.
- Program staff told us they successfully implemented the peer support model because the program had (1) a strong training curriculum, (2) active supervision built on trust and flexibility, (3) the commitment of senior management, and (4) strong lines of communication.
- Program staff were able to find new funding opportunities that allowed them to sustain Project HEALTH.

A. Introduction

The Center for Health Care Services (CHCS), a mental health care provider based in San Antonio, Texas, used Health Care Innovation Awards (HCIA) funding to implement Project HEALTH (Homeless Engagement Addressing Limitations to Healthcare) on the Haven for Hope campus, a campus designed for homeless individuals that includes a traditional homeless shelter and a range of mental health and social services. The program targeted homeless adults with serious mental illness—or serious mental illness combined with a substance use disorder—who also had or were at risk of developing a chronic physical disease. Project HEALTH had the following two key components:

- **Delivery of primary care services.** Project HEALTH developed a multidisciplinary care team that included behavioral health staff from the existing wellness center, and primary care staff drawn from a newly established primary care clinic. The wellness center and primary care clinic were both located on the Haven for Hope campus.
- **Peer support.** Project HEALTH assigned participants to a peer support specialist who helped them build and sustain their readiness for change, motivation, and compliance with the treatment plan.

By implementing Project HEALTH, CHCS sought to improve access to health care and management of chronic conditions, which staff hypothesized would reduce emergency department (ED) visits and psychiatric and medical hospitalizations.[10] CHCS leaders hypothesized that this shift in service utilization patterns would reduce health care costs. In collaboration with the University of Texas at Austin, CHCS offered the program to a randomly selected group of potential participants. Our evaluation of the project was therefore based on a

[10] In June 2015, CHCS received a no-cost extension from CMMI to extend its program for 12 months. As a result, CHCS was able to maintain Project HEALTH's services through October 2015 and continue its HCIA-funded evaluation, analysis, and dissemination through June 2016.

randomized controlled trial design—that is, the evaluation included an intervention group and a control group.

We drew on the following data sources for the findings presented in this chapter:

- **Enrollment data** submitted by CHCS to the reporting website maintained by CMMI's technical assistance contractor (the Lewin Group) for the HCIA, Round 1
- **Qualitative data** collected by Mathematica during site visits and telephone interviews through June 2016. We conducted in-depth interviews with key awardee staff, members of the workforce, and other stakeholders. We also convened focus groups with members of the workforce and with participants in the program and control group.
- **Quantitative data** that CHCS collected from program and control group participants, using various survey instruments and health status assessment measures; CHCS then submitted the data to Mathematica through December 2015. As described below, we used the data to conduct an analysis of the program's impact on psychological stress, capacity for self-management, feelings of hope, and capacity for life change.
- **Workforce survey data** collected by Mathematica in spring 2015. We designed the survey primarily to gather information on (1) workforce stress and burnout, (2) job satisfaction, and (3) perceptions about the effectiveness of the training and support that staff received to help them function effectively in their jobs.

This chapter is a comprehensive and final summary of the findings of our evaluation of CHCS's HCIA program.

1. Overview of administrative context

Before the HCIA, members of the CHCS leadership team identified two significant challenges in their ability to address the range of health and social needs of Haven for Hope residents. First, staff members observed that the majority of residents seeking mental health services from CHCS on the campus had significant physical health conditions and medical needs that CHCS was not equipped to handle. Campus residents in need of primary care sought these services at the neighboring federally qualified health center (FQHC). However, many CHCS staff and residents considered the FQHC inadequate for addressing their physical health needs. Many of the participants we interviewed said they waited a long time for services and sensed that the FQHC staff did not respect or know how to work effectively with homeless clients. In addition, interview respondents perceived that the FQHC was not well-equipped to provide follow-up care for individuals with chronic mental health conditions, noting the lack of a mechanism for coordinating care with the behavioral health and social services provided by CHCS.

Second, CHCS staff members told us how difficult it was to keep clients engaged in services. Many clients did not make it a priority to obtain or follow up with health and social services because they were struggling to obtain daily necessities, such as food, clothing, and shelter. Untreated mental illness and substance abuse also prevented many clients from seeking assistance and staying engaged once they got it. In addition, CHCS lacked a strong coordination mechanism that linked various service programs, allowed for cross-service communication, and

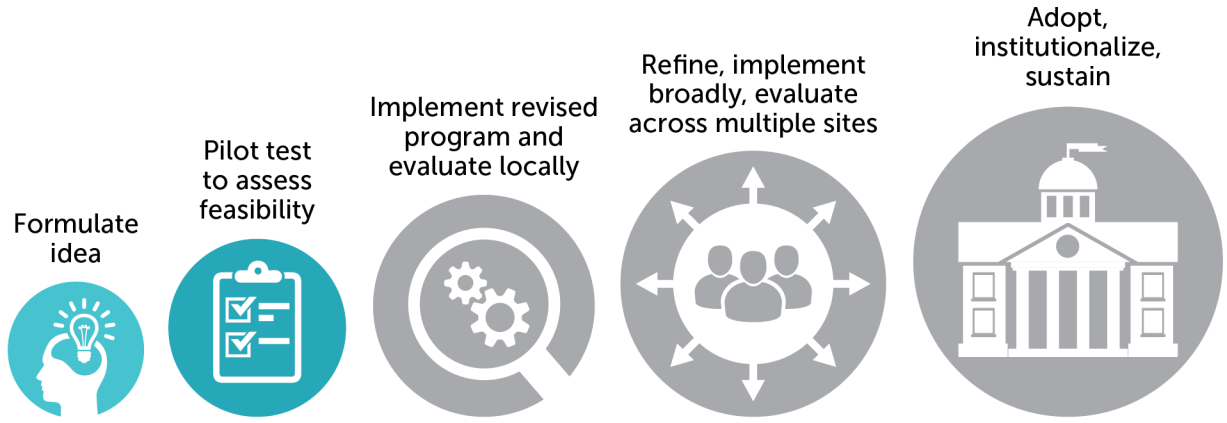
kept clients engaged. As a result, clients were often lost to follow-up and isolated from the services that could ultimately lead them to greater stability and well-being.

To address these challenges, the CHCS leadership team adopted a two-pronged strategy for converting its service delivery model into an integrated and coordinated system of care. First, CHCS would establish an on-site primary care clinic and integrate it with the existing behavioral health service structure, giving clients a comprehensive approach to care. Second, believing that improved service effectiveness required greater coordination of client care, CHCS leaders would assemble a care coordination team of peer support specialists—individuals who had lived with homelessness, had personal experience in coping with mental health and substance abuse conditions, or both—who would help clients remain engaged in services and promote their readiness for change. CHCS leaders believed that, in combination, these two innovation components would lead to more efficient service use and better client outcomes. CHCS viewed the HCIA as the mechanism for pursuing its goal.

2. Progression through phases of innovation

The HCIA-supported Project HEALTH was a pilot program; a small number of randomly selected residents were the mechanism through which CHCS tested the impact of (1) a primary care clinic integrated with behavioral health services and (2) a peer support program. With Project HEALTH, CHCS moved through the first two phases of innovation in Figure II.1.

Figure II.1. Phases of program innovation: CHCS

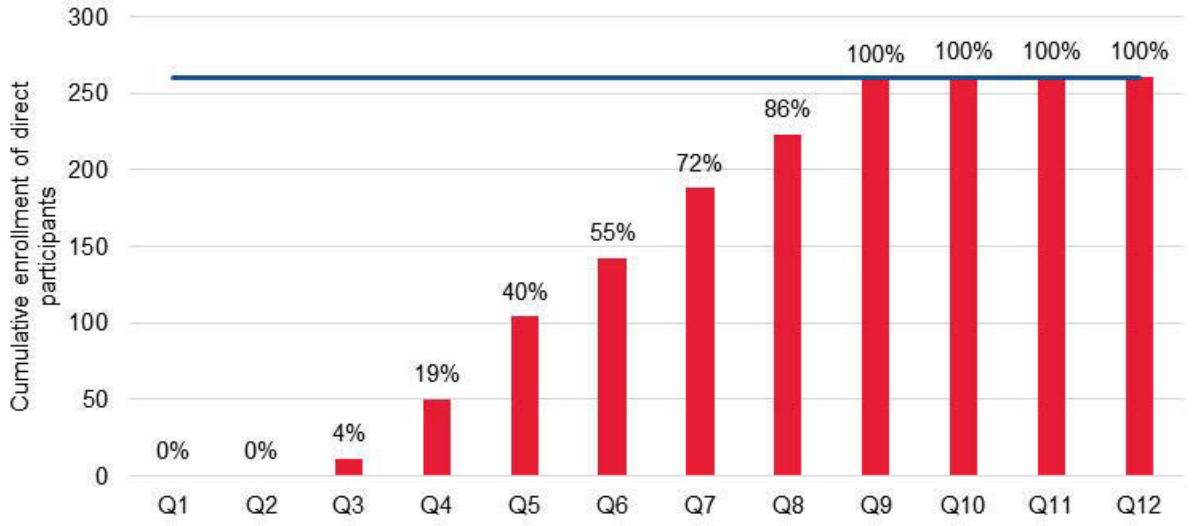


3. Enrollment

By the ninth quarter following receipt of the HCIA award, CHCS successfully reached its enrollment target of 260 direct participants (Figure II.2). CHCS defined a direct participant as an individual randomly assigned to receive access to the program, including the primary care clinic and a peer support specialist. CHCS also enrolled 259 participants in a control group. Control group members received standard care, including access to preexisting on-site behavioral health services (but no access to the primary care clinic) and staff assistance with linkages and referrals to existing social services and resources (but no access to a peer support specialist). CHCS

recruited all intervention and control group members from Prospects Courtyard, a safe outdoor sleeping area on the Haven for Hope campus.

Figure II.2. Percent of target enrollment achieved by quarter, Q1–Q12



Source: Awardee's enrollment data reported to the website maintained by CMMI's technical assistance contractor (the Lewin Group).

Note: As seen in the horizontal line, CHCS's target enrollment was 260 unique participants.

4. Participants' demographic characteristics

As shown in Table II.1, the majority of program participants were under age 45 (55 percent), male (57 percent), unemployed (93 percent), and uninsured (73 percent). Demographic information on control group members appears in Appendix A.

B. Methods

In this section, we present the quantitative methods we used in the impact analyses discussed in Section C (summative findings) as well as the qualitative methods we used to identify the findings presented in Sections D and E (findings about the workforce and sustainability).

1. Quantitative methods

As noted, CHCS included a randomly assigned control group as part of its program design, which enabled us to use an experimental design to conduct impact analyses. For these analyses, we used data from survey assessments and health status measures that CHCS staff administered to program participants and control group members at baseline and at two follow-up time points[11] —around 6 and 12 months post-enrollment. We restricted samples for each outcome regression to program participants and control group members for whom we had data at baseline

[11] Follow-up assessment dates did not necessarily occur exactly 6 and 12 months after a participant or control group member's enrollment date, in large part because of difficulties in reaching a person for follow-up. For "6 months," CHCS measured outcomes at a point generally between 5 and 7 months after enrollment. Similarly, for "12 months," CHCS measured outcomes generally between 11 and 13 (or more) months after enrollment.

and at least one of the two follow-up points. This restriction excluded approximately 20 percent of each of the two groups that did not have longitudinal data.

Project HEALTH hypothesized that making primary care services more available would reduce the number of hospitalizations and ED visits, among other outcomes. We attempted to use Medicare or Medicaid claims to examine changes in service utilization. However, only 69 individuals in the intervention group were enrolled in Medicare, Medicaid, or both, representing just over one-fourth of the 261 individuals enrolled in the program (Table II.1). Therefore, we did not have sufficient sample size to examine these outcomes using Medicare or Medicaid claims data. We were unable to identify an alternative data source that fully encompassed the population and provided information on the outcomes of interest. Therefore, we were not able to analyze hospitalizations, ED use, or health care expenditures for CHCS.

Table II.1. Demographic characteristics of Project HEALTH participants

	Number	Percent
Total	261	100
Age		
18–34	78	29.9
35–44	64	24.5
45–54	87	33.3
55 or older	32	12.3
Gender		
Male	148	56.7
Female	113	43.3
Insurance coverage		
Medicaid, non-dual	40	15.4
Medicare, non-dual	12	4.6
Dual	17	6.5
Other[a]	191	73.5
Employment status		
Employed	18	6.9
Not employed	242	93.1
Living situation		
Homeless	230	89.8
Not homeless or other[b]	26	10.2
Education level		
Less than high school	27	10.4
Some high school	62	23.9
High school diploma or GED	107	41.3
More than high school/GED	63	24.3

Source: EHR data provided by CHCS, November 2014.

[a] "Other" refers to intervention group members with neither Medicaid nor Medicare. Due to the format of the data provided, we cannot assess whether these members are uninsured or have some other form of insurance; however, based on conversations with staff, we assume that many of these individuals are uninsured.

[b] "Other" refers to correctional facilities, group quarters, and other living situations not explicitly classified as homeless or not homeless.

In lieu of examining these outcomes, we examined four self-reported survey scores and two health status measurements from the assessment data collected by CHCS staff. These survey

scores measure psychological stress, capacity for self-management, feelings of hope, and capacity for life change. Staff at CHCS expected the program to affect these scores, which come from the following instruments:

- **Brief Symptom Index 18 (BSI-18).** The BSI-18 is a self-report measurement tool that screens for psychological distress and psychiatric disorders, with higher composite scores (ranging from 0 to 72) indicating higher distress levels.
- **University of Rhode Island Change Assessment (URICA).** The URICA assesses motivation for change through scores on four stages of change: pre-contemplation, contemplation, action, and maintenance. Higher committed action composite scores, calculated by subtracting the contemplation score from the action score (ranging from -4 to 4), indicate a higher motivation to change.
- **Short Form 36 Health Survey, Version 1 (SF-36).** The SF-36 assesses quality-of-life measures, including physical function, physical and emotional role limitations, general health, pain, emotional well-being, social functioning, and energy or fatigue. Higher SF-6D composite scores, derived from selected SF-36 instrument items as a single measure of preference-based health status (ranging from 0 to 1), indicate better quality-of-life ratings.
- **Adult Hope Scale (AHS).** The AHS measures a participant's feelings of hope on two subscales: (1) pathways—measuring the ability to plan to achieve goals—and (2) agency—measuring the ability to initiate and sustain use of those pathways. When added together, the two subscales create a global hope score (ranging from 8 to 64), with higher scores indicating increased feelings of hope.

We also assessed two health status outcome measures: weight loss (using body mass index [BMI]) and blood pressure (both systolic and diastolic). We selected measures that were applicable to the CHCS program, were feasible to construct with the assessment data provided by CHCS, and demonstrated strong evidence of clinical importance. To examine the impact of Project HEALTH on these health status measures, we limited the analysis sample to only those program participants and control group members with a suboptimal measure score at baseline. Specifically, we examined weight loss among individuals with a BMI meeting the definition of overweight or obese[12], and examined reductions in blood pressure among those who were hypertensive.[13]

We then examined the changes in these measures during the first 12 months of the intervention based on the HEDIS convention of using the last measurement within 12 months. That is, we examined impacts from baseline to either their 6 or 12 month follow-up (whichever

[12] The National Institutes of Health defines overweight as a BMI of 25 to 29.9 kg/m^2 and obese as a BMI of 30 kg/m^2 or greater (National Institutes of Health. "Classification of Overweight and Obesity by BMI, Waist Circumference, and Associated Disease Risks." Accessed August 29, 2016 at [http://www.nhlbi.nih.gov/health/educational/lose_wt/BMI/bmi_dis.htm].

[13] The HEDIS measure definition of high blood pressure specifies a cutoff of 140/90 for adults ages 18 to 59 and for adults ages 60 to 85 with diabetes; and 150/90 for adults ages 60 to 85 without diabetes (National Committee for Quality Assurance (NCQA). The State of Health Care Quality 2015. Washington, DC: NCQA, 2015). CHCS data did not specify whether a participant or a control group member had diabetes; thus, we used a more conservative cutoff of 140/90 for all participants. The extra distinction is unlikely to have made a substantial difference in our analysis because only 5 percent of our sample with available blood pressure measures were age 60 or older.

was later to allow for more time for improvement). We examined the proportion of intervention and control group members who lost 5 percent of BMI and who were no longer hypertensive, and we looked at the continuous changes in BMI and in systolic and diastolic blood pressure.

More information on these measures, the regressions used in the impact analyses, and information on the comparability of intervention and control group characteristics appears in Appendix A.

2. Qualitative methods

We collected qualitative data during site visits to the Haven for Hope campus in March 2014 and March 2015. During both visits, we conducted in-depth interviews with CHCS leaders, members of the workforce, and other stakeholders to discuss respondents' perceptions of program effects, implementation challenges, the level of workforce satisfaction, and relevant internal and external contextual factors. During the second visit, we convened a focus group with members of the workforce and with members of the intervention and control group. We conducted a telephone interview with CHCS leaders in May 2016 to discuss the program's sustainability.

C. Summative findings

1. Survey assessment outcomes

None of the estimated impacts of Project HEALTH was statistically significant for the four survey assessment scores (Table II.2). Both the intervention and control groups showed statistically significant improvements in all four scores from baseline to each follow-up measurement point; however, the improvements for the intervention group were not statistically different from those observed for the comparison group. In other words, from baseline through follow-up, the program participants reported a statistically significant reduction in psychological stress and increases in their capacity for self-management, feelings of hope, and capacity for life change. However, we cannot attribute these positive outcomes to the impact of the intervention, because the control group members had comparably significant improvements in their outcomes.

2. Health status outcomes

We found that Project HEALTH had no significant impact on weight loss or hypertension control (Table II.3). Slightly over one-fourth of intervention and control group members who were overweight or obese at baseline lost at least 5 percent of their body weight; however, the difference between the two groups was not significant. Similarly, among those who were hypertensive at baseline, 41 percent of intervention group members and 33 percent of control group members were no longer hypertensive at follow-up. The difference between the groups was not significant.

Similarly, among those who were overweight or obese or hypertensive at baseline, the intervention had no significant impact on average BMI and blood pressure in the follow-up period (Table II.3). There was a statistically significant gain in BMI among intervention group members, along with statistically significant reductions in systolic blood pressure for both groups.

Table II.2. Estimated impacts of Project HEALTH based on survey assessments

	Intervention group members	Control group members	Estimated impact[a]		
			Value	Percent	*p*-value
AHS global hope					
Average score at enrollment	41.76	42.62			
Change in score 6 months after enrollment	2.60*	1.80**	0.80	1.84	0.465
Change in score 12 months after enrollment	3.20**	2.39**	0.81	1.86	0.505
BSI-18 global distress					
Average score at enrollment	30.69	31.23			
Change in score 6 months after enrollment	-9.85**	-9.68**	-0.18	-0.84	0.913
Change in score 12 months after enrollment	-11.75**	-11.50**	-0.26	-1.22	0.891
SF-6D					
Average score at enrollment	0.57	0.57			
Change in score 6 months after enrollment	0.04**	0.05**	-0.01	-1.45	0.487
Change in score 12 months after enrollment	0.05**	0.06**	-0.01	-1.45	0.551
URICA committed action					
Average score at enrollment	-0.28	-0.26			
Change in score 6 months after enrollment	0.14**	0.15*	-0.02	14.92	0.820
Change in score 12 months after enrollment	0.16**	0.18**	-0.02	17.64	0.790

Source: Mathematica analysis of survey assessment data provided by CHCS, December 2015.

Notes: We report regression-adjusted average scores at baseline and changes from baseline to 6- or 12-month follow-up assessment. Asterisks indicate statistically significant changes from baseline to follow-up, as described below.

To be included in the analysis, individuals in both intervention and control groups must have had a baseline score as well as the corresponding follow-up. All regression models control for age, gender, employment status, insurance status, education status, living situation, month of enrollment, and duration of services used. We define 6- and 12-month follow-up as the dates when an individual received his or her 6- and 12-month follow-up survey measurements, respectively. Due to data availability, the number of intervention and control group members included in regressions varied, from between 168 and 217 intervention members and between 160 and 209 control group members.

[a] We derived the impact estimates in Stata by using the margins command to compare the difference between the baseline and follow-up period means for the intervention and control groups.

†Significantly different from zero at the 0.10 level, two-tailed test.

*Significantly different from zero at the 0.05 level, two-tailed test.

**Significantly different from zero at the 0.01 level, two-tailed test.

Table II.3. Impact estimates for changes in health status

	Intervention group members	Control group members	Estimated impact[a]		
			Value	Percent	*p*-value
If overweight or obese at baseline, change in BMI					
Percent overweight or obese at baseline	55.82	60.78			
Percent of whom lost at least 5 percent of their body weight[b]	27.43	25.81	1.62	5.84	0.780
Average BMI at baseline	33.13	32.60			
Average change in BMI from baseline[b]	0.53†	-0.01	0.55	1.66	0.168
If hypertensive at baseline, changes in systolic and diastolic blood pressure (BP) measures					
Percent hypertensive at baseline	22.22	23.56			
Percent of whom were no longer hyptertensive[c]	41.46	33.18	8.28	20.0	0.436
Systolic BP					
Average at baseline	158.02	153.29			
Average change in systolic BP from baseline[d]	-7.21†	-7.61†	0.40	0.27	0.945
Diastolic BP					
Average at baseline	94.41	94.10			
Average change in diastolic BP from baseline[d]	-0.28	-0.83	0.55	0.59	0.879

Source: Mathematica analysis of survey assessment data provided by CHCS, November 2015.

Note: We present group-level means. We do not adjust baseline means and percentages, but average changes from baseline are regression-adjusted. Asterisks indicate statistically significant changes from baseline to follow-up, as described below.

To be included in the analysis, individuals in both intervention and control groups must have had the underlying condition at baseline. All regression models control for age, gender, employment status, insurance status, education status, living situation, and quarter of enrollment. We define 6-month and 12-month follow-up as the dates when an individual received his or her 6-month and 12-month follow-up survey measurements, respectively.

[a] We derived the impact estimates in Stata by using the margins command to compare the difference between the baseline and follow-up period means for the intervention and control groups.

[b] Denominator is individuals who were overweight or obese at baseline and were included in the regression; n = 113 intervention group members, n = 117 control group members

[c] Denominator is individuals who were hypertensive at baseline and were included in the regression; n = 41 intervention group members, n = 45 control group members

[d] Denominator is individuals who were hypertensive at baseline and were included in the regression; n = 43 intervention group members, n = 46 control group members

†Significantly different from zero at the 0.10 level, two-tailed test.

*Significantly different from zero at the 0.05 level, two-tailed test.

**Significantly different from zero at the 0 .01 level, two-tailed test.

3. Analytic limitations

The experimental design, including intervention and control groups, enabled us to conduct an impact analysis. However, we note several limitations in our analytic approach.

Small sample size. Our analysis was limited to participants assessed both at baseline and at least one of the follow-up points. CHCS completed 6- and 12-month follow-up surveys with about 80 percent of the individuals in each group at baseline (fewer than 120 individuals for most

analyses). Overall, the sample size limited our ability to detect small differences between the intervention and control groups.

Limited time frame. The period for which we were able to obtain data for our sample was limited. We examined outcomes for up to one year following enrollment, which may not be a long enough follow-up period to detect effects.

Lack of data on utilization and expenditures. CHCS staff hypothesized that an increase in the availability of primary care services to the intervention group would lead to a reduction in hospitalizations and ED visits for that group. In the absence of an adequate data source, we could not assess program impact on these or other health service and expenditure outcomes.

4. Qualitative findings on perceived effects

All respondents, including CHCS leaders, members of the workforce, and participants, believed that Project HEALTH had a positive effect on participants' physical health, mental health, and overall well-being. Anecdotally, staff felt that regular access to the preventive and primary care team was resulting in fewer hospitalizations and ED visits. Peer support specialists noted that some participants were accustomed to relying on the ED to address their basic health care needs and that changing this behavior required peers to follow up persistently. Primary care staff struggled to find specialty care providers in the community willing to take their clients; however, staff provided several examples of participants who successfully received much-needed specialty services, such as surgery. Several participants described how the program helped them adhere to their prescribed medications and attend scheduled appointments. According to one participant, "They would always go out of their way to help me get to my doctors' appointments with all three different places . . . they would hunt me down to make sure I go. They are very persistent."

CHCS leaders and members of the workforce also reported improved participant engagement and follow-up with mental health and substance abuse services. Many respondents attributed greater use of behavioral health services to the role of the peer support specialist. For example, according to one peer, the relationship he established with a participant led that individual to confess that he often felt suicidal. The peer established a level of trust with this participant and helped him commit to seeking out mental health services. One participant relied on his peer during tough times because he knew the peer had faced similar challenges. "If I was having a bad day I could just call him and talk to him . . . the situation I was going through, he has already been through, so he knew how to talk to me in a certain voice and a certain manner that you're not going to get from talking to your buddy at [the shelter] . . . he just knew how to calm me down."

According to peers, the majority of participants used alcohol or drugs, which they viewed as a major barrier to accessing needed physical and mental health services. Peers strongly believed that (1) participants viewed them as role models because of their own lived experience with homelessness or behavioral health challenges and (2) this dynamic prompted many participants to seek sobriety and change.

Peer support specialists emphasized their role in helping participants address essential needs other than health care, such as housing, clothing, and employment. For many participants, these

needs were a higher priority than their physical and mental health conditions. Peers used their network of community resources to help participants obtain a range of services outside the health sphere. For example, peers helped many participants transition into campus wellness dormitories or off-campus housing. One peer described the importance of helping a participant obtain a pair of shoes: "That's what was preventing him from getting a job, from moving on and getting out of here and getting back into normal life. We help with simple things like that, which make a huge difference to these folks."

One participant described how her peer had been instrumental in helping her obtain custody of her children, commenting, "[My peer] makes me feel important as an individual. She makes me feel like I'm loved, she makes me feel like I'm cared about. Somebody actually worries about me and my kids, it's not just me . . . she worries about my family, and it means a lot." According to peers, establishing these types of relationships took time and persistence but was necessary to be effective in helping participants address the challenges they faced.

D. Findings about the workforce

Nearly all of the Project HEALTH workforce roles were new to CHCS, including:

- **Peer support specialists.** The seven peers, who had direct experience with mental illness or homelessness (or both), played a central and crucial role in connecting Project HEALTH participants to all other services, coordinating services, and building the participant's motivation and readiness to change.
- **Community guest specialists.**[14] The specialists, many of whom also had direct experience with mental illness or homelessness, were responsible for recruiting and enrolling participants from the campus shelter. The eight guest specialists also administered assessments to intervention and control group members at regular intervals.
- **Health navigators.** Two navigators, both of whom had an advanced degree in a related field, supervised and provided guidance to the peer support specialists and the community guest specialists. The navigators helped the peers and guest specialists think through challenges, identify resources, and brainstorm ideas.
- **Primary care physician and licensed vocational nurse.** The physician and nurse provided primary care and worked closely with the psychiatrist, other pre-existing behavioral health staff, and the peer support specialists to coordinate client care. This "care team" communicated regularly by email and telephone and held both regular in-person meetings and impromptu conversations.

Most members of the Project HEALTH workforce, which consisted almost entirely of peer support specialists and community guest specialists, said they were satisfied with their jobs. In response to the 2015 workforce survey, 95 percent of staff reported they were either moderately or extremely satisfied with their current job. Consistent with their general satisfaction with their work and dedication to their clients, the Project HEALTH survey respondents reported high

[14] "Community guest specialist" was a pre-existing job title at CHCS. Prior to Project HEALTH, these staff provided direct client services. Under the HCIA project, however, the community guest specialists were responsible purely for the study-related activities described in this chapter.

levels of personal accomplishment and low levels of depersonalization and emotional exhaustion. Peers and community guest specialists substantiated these survey results in interviews and focus groups, noting that, although their jobs may be draining, they usually do not feel overwhelmed by stress. However, members of the Project HEALTH leadership team recognized that the work can take an emotional toll on people who provide direct services to participants. "It's emotionally very challenging work to be on those front lines . . . a lot of our frontline staff, man, they are just out there. They are just there with people's emotions and their breakdowns and their—everything. All the time. They're amazing."

In general, all respondents perceived the peer workforce program as successful in its ability to assist participants and improve overall service delivery. As noted below, four critical workforce lessons emerged from CHCS's experience with the peer workforce.

CHCS benefited from collaborating with an organization that specialized in providing peer support training and certification. All peer support specialists received intensive training from the Via Hope Texas Mental Health Resource, which provides recovery-oriented peer certification training. The program trains staff in a range of techniques, including motivational interviewing, wellness self-management, chronic care models, shared decision making, crisis intervention, and person-centered recovery-based care. Administrative staff relied heavily on Via Hope to help address questions about peer training or supervision and to provide general support and consultation on implementing a peer support program.

In addition, CHCS leaders provided frontline staff with training in how to manage stress and maintain a work-life balance. The peers and community guest specialists we interviewed for the evaluation believed that the training sessions were helpful. "Most of the trainings we have are geared towards self-care. We have to balance the job with our personal life. From 8:30 to 5:30 it is intense and we deal with intense situations. But at 5:31, that's my time." Supervisors also instructed staff not to answer calls and texts that came in over the weekend from participants. Peers and community guest specialists explained that most participants learned not to contact staff outside business hours. "We do get calls and texts over the weekend. We're instructed not to answer. And most of the time I don't. The participants need to learn about boundaries too. It's called modeling."

The peers and community guest specialists we interviewed for the evaluation explained that the training helped them develop useful skills such as motivational interviewing and cultural competency. Although peers received training in crisis intervention, several noted that they would have benefited from more advanced training. One peer explained why this training was relevant to the work at hand, noting "There have been so many situations here where people just lose it, or they're not on their medication, or something like that. You have to know how to talk to these men and women. You have to be calm. There's a lot of things involved, the people that are around you. You need to know what to do. That would be very useful."

CHCS leaders believed that providing a peer workforce with active supervision, flexibility, and support reduced burnout. Management and frontline staff emphasized that the health navigators, who provided direct and active supervision, were a critical outlet for a workforce at risk of burnout and that they played an important role in helping staff think through and navigate challenging situations, such as setting and maintaining boundaries with clients.

Peers also noted how the supervising health navigators supported them by providing time off to deal with stress. As one peer explained, "There's always going to be stress, but in my opinion, if it gets to a point where it's becoming too much, we could always go to our supervisor and say, 'Look, I need a day.'" One peer said it was critical for peers to take time off to reflect on their experiences. "I'll go out in the backyard with my dog, and I think about things that I said to people. And then, could I have said it differently for better results, and that helps me not burn out." Peers also pointed to the freedom that navigators accorded them in providing services. As one peer explained, "I think part of it is the autonomy that we've been given by the navigators. They've pretty much let us do our job. They know we're qualified. They've been with us from day one. They trust us. We trust each other."

Demonstrated commitment among organizational leaders helped ensure that peers felt respected and valued by staff. Peers reported that, early on, they struggled to be viewed as credible work colleagues by other campus workforce staff; the situation adversely affected their morale and effectiveness. CHCS's top leadership team conducted campus-wide outreach to ensure that peers were treated equally. Peers recognized and appreciated the CHCS leadership and management teams' commitment to ensuring that they felt supported and respected. Consequently, peers noted that as the program and their role in it became more well-known throughout the greater CHCS campus, they were seen as a unique and valuable resource. One peer told us how she felt when the psychiatrist allowed her to participate in an initial evaluation with a client. "This is optional and it feels like she trusts me, especially with confidentiality. It's very rewarding that she accepts me in that role. And it helps me understand where the person is."

Communication and collaboration among peers and other program staff was critical to overall effectiveness. Staff repeatedly cited communication and teamwork as two factors essential to the program's success, with peers and clinicians relying on each other to think through challenges and brainstorm solutions. Clinicians and peers were required to meet regularly to ensure communication and strong working dynamics. In the long term, all members of the workforce, including the clinicians, viewed the required meetings as valuable; in fact, the workforce began to rely heavily on the peers to follow up with clients about treatment or to provide important contextual information on family situations, housing, or other important background information. In addition, CHCS adopted a fully integrated EHR system used by both mental and physical health care staff, although not without significant technical challenges. The system allowed mental and physical health staff to collaborate more effectively.

E. Program sustainability and spread

Our May 2016 telephone discussion with awardee leaders revealed that, upon the expiration of HCIA funding, CHCS successfully sustained and expanded use of the primary care clinic and peer support program. According to CHCS staff, two primary factors led to sustainability: funding availability and community outreach and buy-in. Both factors were supported by the commitment and dedication of the CHCS leadership team.

1. Availability of funding mechanisms

Funding to support CHCS's integrated care model is scarce. The behavioral health services provided by CHCS on the Haven for Hope campus primarily target the uninsured—Texas did not expand Medicaid to cover childless homeless adults—and thus the mechanisms for cost

reimbursement are few. Consequently, CHCS staff are continuously pursuing federal, state, local, and private funding opportunities to support their service programs.

In January 2015, funds from the state's Delivery System Reform Incentive Payment (DSRIP) program, part of the broader CMS Section 1115 waiver demonstration, became available to support the primary care clinic, including the primary care physician and registered nurse. Additional funding from a mix of other sources, such as state and local grants, also supports the program. The clinic now provides services to selected clients beyond those who participated in Project HEALTH.

CHCS has integrated most of the peers into other existing service programs supported by a variety of sources, including some administrative funds and local grants. Senior staff strongly believe that the HCIA-supported peer program made health and social services more accessible to clients, and the organization continues to pursue funding opportunities to sustain and eventually expand peer support. According to one senior member of the implementation team:

> I just really love that the Center has taken [the peer program] and run with it. I've been at organizations before where [the attitude is], "Funding ran out, close this up, do the next thing." Whereas here [the attitude is], "We figured this out through a research project, and it really worked. It's really something we want to sustain." There's a big push for the program, and a lot of program administrators have been very adamant that this can't disappear. This is a really important component. It wasn't here before, but since it's been built in, we've become dependent on it. We rely heavily on the peers.

2. Community outreach and buy-in

CHCS's success in obtaining grants and other sources of funding relies, in part, on a clear demonstration of how CHCS's integrated and coordinated service system positively affects the community as a whole, particularly other local service providers. For example, with the provision of on-campus integrated health services, high-cost clients are less likely to end up in other organizations' systems. To foster collaboration, members of the CHCS leadership team frequently conduct outreach to community partners, such as the FQHC, local and county administrators, hospital systems, and correctional facilities and other criminal justice agencies, as well as to other nonprofit organizations.

CHCS leaders view these efforts as essential for long-term sustainability. According to a senior CHCS staff member, "Before, we didn't know how much it would cost to [provide integrated care]. Now we can say, 'Hey, if you want us to serve 200 people, this is about how much it's costing us just to get people stable.' Then we can have an intelligent conversation with the community and say, 'This is what we need from you.' Then we can engage the hospitals and say, 'If you just give us $100,000, you're going to save $1 million because [the consumers] are not going to be in your emergency room.' So that's the sustainability piece. It's really dependent on what the data shows in the end. It all comes down to money."

3. Challenges to long-term sustainability

CHCS is implementing a sustainability plan for both components of its program, but senior leaders expect to face barriers to long-term sustainability and expansion. The most significant barrier will be funding. CHCS does not have a permanent and dedicated source of funds to support either its workforce of peers or its primary care clinic. Senior staff will need to work continuously to identify and pursue public and private funding opportunities.

Although CHCS is using DSRIP funds to support the primary care clinic, CHCS staff reported that Texas's decision not to expand Medicaid hinders the ability of their clients to receive physical health services from specialists. The primary care clinic and its staff are not equipped to provide specialty care, such as cancer treatment, or technical services, such as imaging and laboratory work. Clinic staff spend a significant amount of time in outreach to the local community of specialty health care providers; however, few specialists are willing to take clients without any mechanism for reimbursement.

One member of the implementation team asked, "How do you serve a group of individuals with no money, no insurance, no Medicaid, and linked to no state reimbursement?" Without a reimbursement mechanism for primary care services, CHCS provides services within a set budget. As a result, staff at the primary care clinic carefully consider the type and cost of the services they recommend to clients. As one staff member reported,

> A lot of times, our [primary care physician] PCP identifies a problem that requires specialty care, like cancer, but there's no provider willing to take [that patient]. The PCP often faces tough questions, like, "Do I treat this person for this condition that I'm aware of, even though I don't really feel capable of managing it? Do I give this person this medication that costs $1,000 a month because I know that's going to help them, even though it will end up taking away from other people that need medication?" There's only so much money in the pot.

F. Lessons learned

Quantitative findings indicate that Project HEALTH had no impact on outcomes measuring mental health status. However, it is important to consider several important analytic limitations. Specifically, the impact analysis relied on a small sample and was confined to a 12-month time frame that may not be long enough to observe noticeable impact. In addition, we did not have access to data for which there was reported anecdotal evidence of positive effect, such as hospitalizations and ED visits, and on the use of other nonmedical services, such as housing.

The quantitative analysis demonstrated that both intervention and control group members showed significant improvements in assessment measures of psychological stress, capacity for self-management, feelings of hope, and capacity for life change. Control group members were neither assigned to a peer support specialist nor supposed to have access to primary care clinic services. However, control group members did have access to the community guest specialists, who were responsible for the initial engagement and recruitment of participants and for administering the baseline and follow-up assessments. Control group members also visited the primary care clinic for baseline and follow-up health assessments. To some extent, control group members might have benefited from these various interactions.

Qualitative data suggested that Project HEALTH had several positive effects on participants' health and well-being and helped participants meet some of their social service needs, such as housing. Interview and focus group participants primarily noted anecdotal changes in service use and access as well as participants' willingness to engage in addressing health issues and making life changes. It is unclear whether these participants would have taken steps or accessed services without the availability of the peer support and clinic; it is possible that these participants' determination would have allowed them to succeed without the program components. However, one theme from our focus group with control group members was the perception that the status quo makes it hard to access desired services, even when an individual is motivated to seek help.

The perceived success of Project HEALTH is linked, in part, to successful reliance on and management of the peer support specialists. In addition to improved client outcomes, CHCS leaders believed that the peer workforce enhanced the organization's service model by operating as a bridge linking various services and programs. Respondents reported that the success of the peer program was attributable to CHCS's collaboration with Via Hope, which provided intensive peer support training and certification, active supervision built on trust and flexibility, commitment among the organization's most senior leaders, and strong lines of communication among all staff.

CHCS has sustained the primary care clinic by using CMS's DSRIP and other state and local funds. The peer support program continues to operate with administrative funds and local grants. Long-term sustainability requires ongoing funding, the commitment and dedication of the organization's leaders, and community outreach and buy-in. The lack of a dedicated funding stream or reimbursement mechanism for primary care and peer-led care coordination services requires CHCS's unceasing dedication to pursuing temporary funds, such as grants. CHCS's success in obtaining grants and funding relies, in part, on proving that the integrated and coordinated service system model has a positive impact on other local and state service providers. A team of organizational leaders, committed and dedicated to improving the lives of the homeless, is driving efforts to engage community stakeholders and pursue funding opportunities.

III. THE FELTON INSTITUTE

Findings from Mathematica's Evaluation of the Felton Institute's HCIA Program

- We were unable to conduct a rigorous impact evaluation of Felton's PREP program because (1) key data were unavailable, (2) we could not identify a comparison group, and (3) the number of enrolled individuals at any one site was too small to support the analysis.
- Our analysis of qualitative data indicated that PREP relied on an efficient referral system established through partnerships with school districts and mental health service providers, who are often the first to observe the signs associated with schizophrenia.
- As with most programs, successful implementation of the PREP program required the right workforce. Identifying highly qualified staff in rural counties and communities was challenging, however, and so was retaining highly trained staff in a competitive market.
- Mental health departments in four counties sustained Felton's PREP program with funding from the state's Mental Health Services Act. Leaders in these counties told us they believed the evidence justified this continuing investment.

A. Introduction

The Felton Institute (formerly the Family Service Agency of San Francisco) is a nonprofit provider of social and mental health services in San Francisco. In collaboration with the University of California, San Francisco (UCSF), Felton developed the Prevention and Recovery in Early Psychosis (PREP) program. An outpatient program, PREP uses a team approach to offer an integrated suite of evidence-based medication and psychosocial interventions to stabilize and promote the remission of early onset schizophrenia in individuals ages 14 to 29.

The PREP program has the following components:

- Community education to promote early identification of psychosis and prompt referral to the program
- Early diagnosis
- Algorithm-guided medication management that promotes long-term compliance with medication regimens, as well as stable remission
- Integrated, evidence-based psychosocial interventions proven to be effective with the target population, including cognitive behavioral therapy for early psychosis (CBTp), multifamily psycho-education groups, and the individual placement and support model of supported employment

PREP services are administered by a workforce of about seven health care professionals, including two or three therapists, a psychiatrist or psychiatric nurse, a case advocate, a family partner, and an employment and education specialist. These staff members work as a team to administer the program's evidence-based practices. Staff receive intensive training and ongoing remote clinical support and supervision from a team of experts that Felton assembled through its collaboration with UCSF. The ultimate clinical objective of the program is to help participants become stable, goal-oriented, and independent within two years of enrollment.

Felton first launched the PREP program in San Francisco in 2007, replicated it in Alameda County in 2010 and in San Mateo County in 2012, and used Health Care Innovation Awards (HCIA) funding to launch the program in San Joaquin County and Monterey County beginning in mid-2012. HCIA funding for PREP in these last two counties ended on June 30, 2015. Although Felton leaders and the clinical support team primarily work in Felton's San Francisco office, the PREP teams are in offices maintained by Felton in each county with an established PREP program. As of the date of this report, PREP programs continue to operate in four of the five counties—San Francisco, Alameda, San Mateo, and Monterey.[15]

We draw on the following data sources for this chapter:

- **Enrollment data** submitted by Felton to the reporting website maintained by CMMI's technical assistance contractor (the Lewin Group) for HCIA, Round 1.
- **Qualitative data** collected by Mathematica during site visits and telephone interviews conducted through June 2016, including in-depth interviews with key awardee staff, members of the workforce, and other stakeholders. We also convened a focus group with members of the workforce and a separate group with program participants.
- **Quantitative data** that Felton extracted from its internal client tracking database and submitted to Mathematica in July 2015. As described below, we used these data to conduct a pre-post analysis of hospitalizations and emergency department (ED) visits. We could not conduct an impact analysis because data limitations prevented us from identifying an appropriate comparison group.

This chapter is a comprehensive and final summary of the findings of our evaluation of Felton's HCIA program.

1. Overview of administrative context

The PREP model is based on the belief that treating an individual before the psychosis seriously progresses will have a likely outcome of long-term stability and quality of life. The effectiveness of the program is thus conditional on early diagnosis. Most individuals targeted by PREP are school-aged youth, as young as 14. County mental health departments, which are often the first point of contact for youth with symptoms, lack the resources and capacity to address the unique and complex needs of this population. Felton's PREP model seeks to bridge this gap by giving participating counties a prepackaged program that incorporates a combination of evidence-based psychosocial practices, a trained and structured team to administer the program, and ongoing clinical support and supervision.

One member of the PREP team explained the county's appreciation of PREP: "The county has a big interest in us [PREP] because they don't have an early intervention program. We get referrals from county staff who have really troubled clients; there is a sense of relief that our program is an option." A member of Felton's leadership team echoed the sentiment and emphasized that PREP is a model of specialized services that its staff are uniquely and

[15] As we discuss in Section E, San Joaquin County did not renew its contract with Felton to provide PREP services. According to staff at Felton, the county contracted with a different organization that will provide early psychosis services as well as non-specialized mental health services, such as crisis residential and inpatient services.

intensively trained to provide. “We’re offering an evidence-based program that nobody else is offering. There isn’t a pool of people on the market trained in the combination of skills that we’ve trained our staff in.”

PREP targets youth in the early stages of disease, so school personnel, family members, and local mental health providers are often the first to observe signs associated with the disease. Stigma, fear, and cultural norms however, prevent many individuals from seeking and receiving treatment. The PREP program’s success has relied on its significant efforts to engage and educate the local community. In the smaller rural communities, the presence of PREP has come as a relief, particularly to the county mental health providers who view PREP as a resource for referring any clients with complex behavioral health needs. At times, this has led the providers to refer clients with serious mental illnesses that PREP does not treat, such as bipolar disorder, highlighting the importance of community outreach and education.

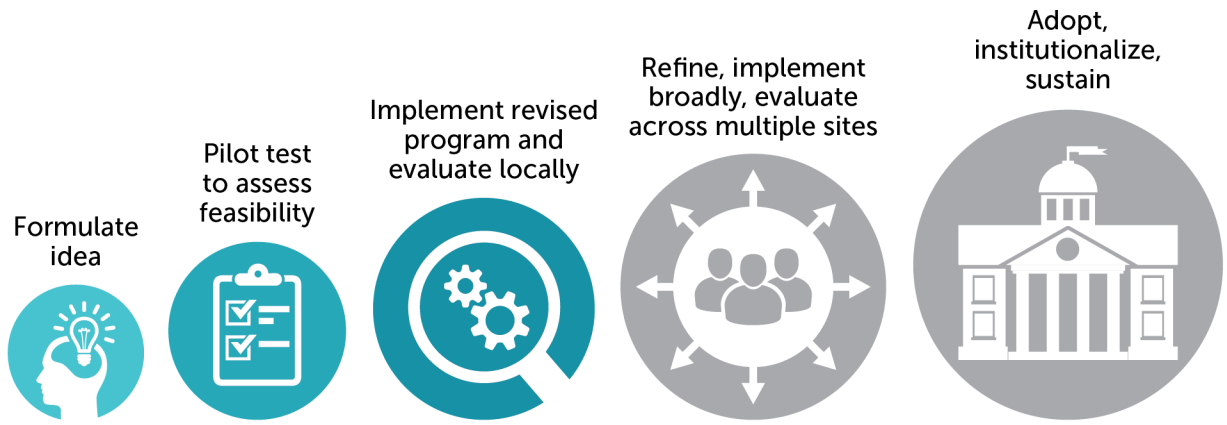
Since launching the first program in San Francisco in 2007, Felton has established a reputation of success for innovation and implementation, which led to the program’s expansion into four more counties. A county representative highlighted Felton’s reputation and established program as factors contributing to county buy-in: “I think what was attractive about the Felton Institute was that they really had an established and effective model in place. They've been working in other counties and there was a history of this working well, so they could say to us, ‘We have found that if you do the multifamily group and the CBT for psychosis, and you do a good assessment at enrollment, and you provide only low level medication, that people respond well, require fewer services, and often recover.’ What county wouldn’t want that?”

Although county leaders may recognize the potential of a program like PREP, funding limitations are a major barrier to broader adoption of the program. County departments of mental health have limited resources to fund a range of services; PREP and other programs to treat early psychosis, which target a relatively small population, must compete for limited funds. Much of Felton’s success is rooted in California’s 2004 passage of Proposition 63, known as the Mental Health Services Act (MHSA), which dedicates a portion of state income taxes to support a broad continuum of mental health prevention and early intervention service needs. Each county mental health department receives MHSA funding and has the flexibility to determine how those funds are spent. Felton’s PREP program has been established and sustained in San Francisco, Alameda, and San Mateo Counties using MHSA funds (as well as funds provided through other mechanisms, such as Medicaid reimbursement).

2. Progression through phases of innovation

Felton’s PREP model accomplished the first three phases of innovation in Figure III.1. Before the HCIA, Felton developed and tested PREP through established programs in San Francisco, Alameda, and San Mateo Counties. Since it was launched in San Francisco, the first site, in 2007, Felton has made minor changes to the model based on internal assessments of implementation and outcomes. Felton successfully expanded the model into Alameda and San Mateo Counties in 2010 and 2012, respectively. That launching was a milestone for PREP, because the attributes of these two counties challenged the model in new ways. Compared to the original three counties, they were larger geographically, made up of mostly rural communities, and had large Hispanic populations, among other demographic distinctions.

Figure III.1. Phases of program innovation: Felton



3. Enrollment

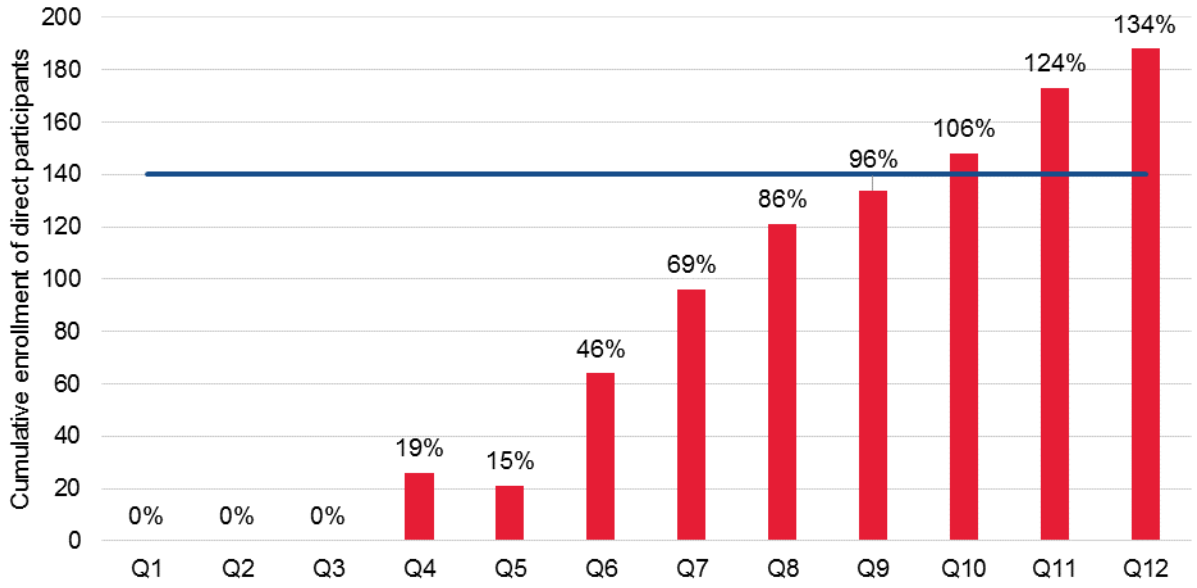
Felton defined an enrolled participant as someone who completed the Structured Clinical Interview for DSM-5 (SCID) following a referral to the program. Felton's enrollment began in the fourth quarter (following a three-quarter planning period), rose steadily through Quarter 9, and surpassed the enrollment goal of 140 in Quarter 10. By the end of the program's 12th quarter (June 30, 2015), Felton had administered the SCID—which staff used for diagnosis—to 188 unique participants in the two HCIA-funded sites (Figure III.2).

Some individuals who completed the SCID, however, received a diagnosis that was not one of the three diagnoses necessary to be eligible for PREP services. These individuals were referred to the appropriate county mental health provider. Among the 188 participants who completed a SCID in the two HCIA-supported county programs, 72 received PREP program services and had data sufficient for quantitative analysis. That is, data were available for the 12-month period following enrollment, which was the time period targeted for analysis.

The small number of participants in the two HCIA-funded counties meant we could not perform rigorous quantitative analyses of program effects. In consultation with CMS, we decided to include data from participants served by all five county programs in order to improve the strength of the analysis. As a result, we use data from 280 participants who received PREP services in any of the five counties between July 1, 2012, and December 31, 2014.

The PREP program in San Francisco, however, is different from the one in the other four counties because it serves individuals in the prodromal phase of the disorder—that is, they are pre-symptomatic. The other four county programs do not serve this population. Individuals in the prodromal phase are less likely than other participants to have a pre-enrollment hospitalization or ED visit. Because trends in hospitalization and ED visits are the focus of our analysis, we present findings for San Francisco program participants separately from the pooled sample of participants served by the other four county programs.

Figure III.2. Percent of target enrollment achieved by quarter, Q1–Q12



Source: Awardee's enrollment data reported to the website maintained by CMMI's technical assistance contractor (the Lewin Group).

Note: The blue horizontal line represents Felton's target enrollment of 140 unique participants.

4. Participants' demographic characteristics

As shown in Table III.1, the majority of participants in the four county sample were between the ages of 19 and 34 (79 percent), male (75 percent), and diagnosed with either schizophrenia, schizoaffective disorder, or schizophreniform (88 percent). The sample was racially diverse, with the largest percent of participants identifying as African American or Asian/Pacific Islander (41 percent). In contrast, the San Francisco program had a smaller proportion of participants who were between the ages of 19 and 34 (68 percent) or had one of the three diagnoses (62 percent). These differences reflect San Francisco's emphasis on the prodromal phase, which means participants are likely to be younger and less likely to have received one of the three diagnoses yet. Although reliable insurance data were unavailable, Felton staff believe the large majority of participants were Medicaid eligible.

Table III.1. Demographic characteristics of PREP program participants

	4-county total[a]		San Francisco[b]	
	Number	Percent	Number	Percent
Total	187	100	93	100
Age				
14-18	40	21.4	30	32.3
19–34	147	78.6	63	67.7
Gender				
Female	47	25.1	28	30.4
Male or transgender[c]	140	74.9	65	69.6
Race				
African American or Asian/PI[c]	77	41.2	34	36.6
Latino	53	28.3	24	25.8
White, non-Hispanic	44	23.5	35	37.6
Other or unknown	13	7	0	0
Diagnosis				
Schizophrenia, schizoaffective disorder, or schizophreniform disorder	164	87.7	58	62.4
Other psychosis	23	12.3	35	37.6

Source: Data collected by Felton staff and maintained in Felton's client tracking database.

[a] Includes Alameda, Monterey, San Joaquin, and San Mateo Counties.

[b] San Francisco is reported separately because of different program enrollment criteria.

[c] In compliance with CMS policy prohibiting report of cells with fewer than 11 individuals. Transgender individuals are reported with male and African Americans and Asian/Pacific Islanders together.

B. Methods

1. Quantitative methods

We used data extracts from Felton's internal Cloud-based Integrated Reporting and Charting Environment (CIRCE) to conduct a pre-post analysis of participants' psychiatric-related hospitalizations and ED visits. Felton developed CIRCE to track demographic data and the results of surveys and assessments. CIRCE also has data on psychiatric-related hospitalizations and ED visits, which Felton staff obtain from the counties' mental health departments. The CIRCE extract provided to Mathematica for this analysis included psychiatric-related hospitalization and ED visit data for the 12 months pre- and post- PREP program enrollment for all participants enrolled in PREP between July 1, 2012, and December 31, 2014. Because these data were manually entered into CIRCE and we were unable to conduct any quality reviews, readers should interpret the results cautiously. More information on the data sources and a description of our process for constructing the outcomes presented here, along with any important caveats and notations, can be found in Appendix A.

2. Qualitative methods

We collected qualitative data during site visits to each of the HCIA-supported county program offices. We visited the San Joaquin County program office in Stockton in May 2014 and the Monterey County program office in Salinas in May 2015. Both visits included in-depth interviews with key awardee staff, members of the workforce, and other stakeholders to discuss

respondents' perceptions of program effects, implementation barriers and challenges, workforce satisfaction, and relevant internal and external contextual factors. During the second visit, we convened a focus group with members of the workforce and with program participants. We conducted a phone interview with key awardee staff in May 2016 to discuss program sustainability.

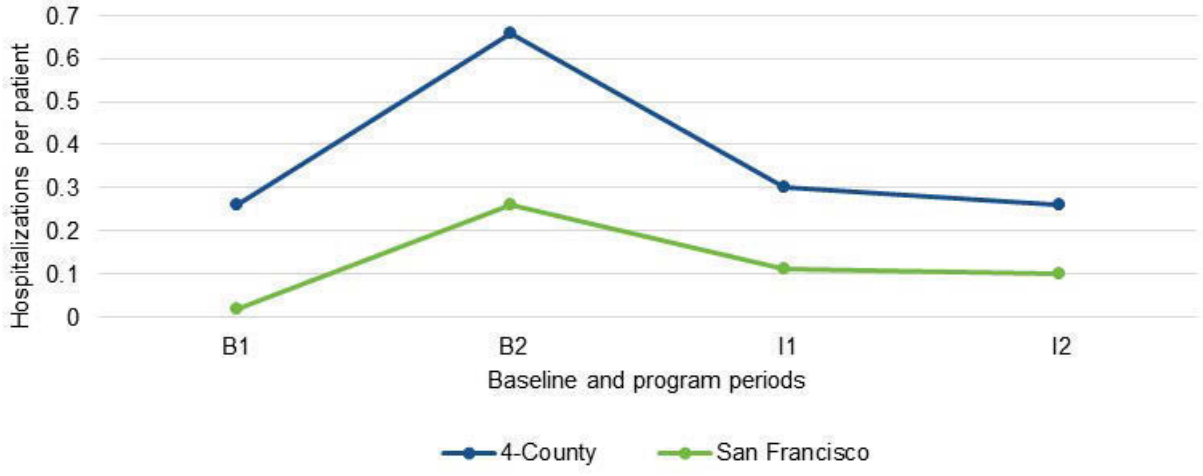
C. Summative findings

1. Effects estimates for selected core measures

For both the four-county and San Francisco-only samples, rates of psychiatric-related hospitalization (Figure III.3) increased during the baseline period. For the four-county group, the rate increased from .26 admissions per participant to .66 admissions per participant. For the San Francisco group, the rates increased from .02 admissions per participant to .26 admissions per participant. The spike in hospitalizations likely reflects a worsening of symptoms prior to enrollment. For many participants, the increase in hospitalizations may have led them to be referred to the PREP program. For both groups, the rates decreased from baseline to the first intervention period and continued to decline slightly through the second baseline period, approaching the rates found at the beginning of our observations. Without a comparison group, we do not know whether these changes would have taken place in the absence of the PREP program.

We observed roughly similar patterns for psychiatric-related ED visits (Figure III.4). In the San Francisco sample, the lowest hospitalization and ED rates were in the first baseline period. This is probably because the program in San Francisco had so many participants in the prodromal phase of the disease, and they were less likely to have had a pre-enrollment hospitalization or ED visit.

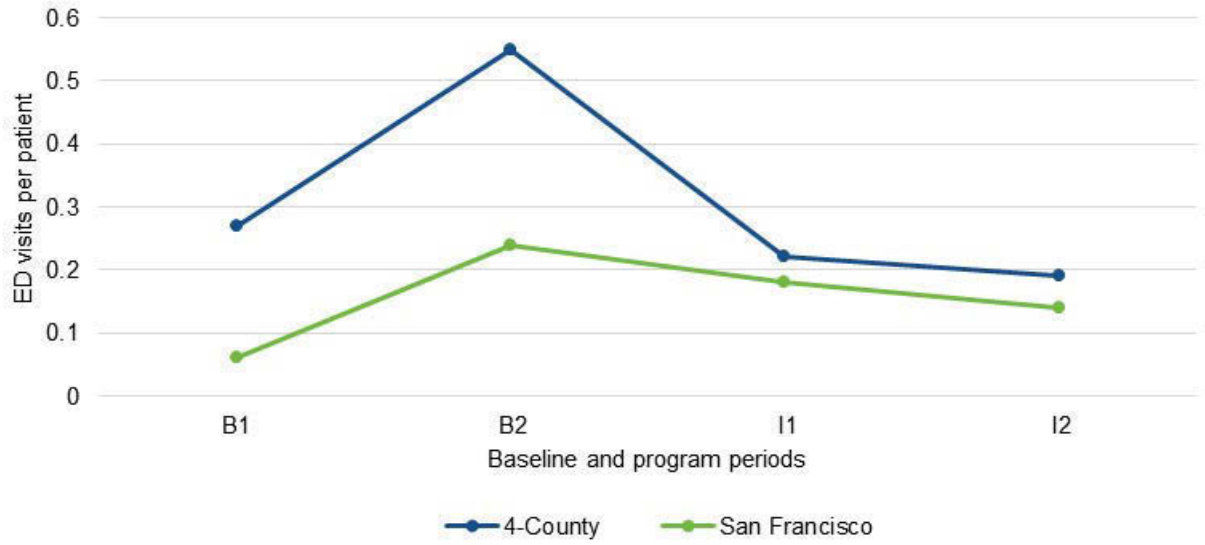
Figure III.3. Psychiatric-related hospitalizations per patient per 6-month period



Source: Data extracted by Felton staff from county mental health department systems and maintained in Felton's client tracking database.

Note: B1 = first baseline period; B2 = second baseline period; I1 = first intervention period; I2 = second intervention period.

Figure III.4. Psychiatric-related ED visits per patient per 6-month period



Source: Data extracted by Felton staff from county mental health department systems and maintained in Felton's client tracking database.

Note: B1 = first baseline period; B2 = second baseline period; I1 = first intervention period; I2 = second intervention period.

2. Analytic limitations

These trends do not give us conclusive information on the program's impact, and they should be viewed with caution for several reasons. First, the analysis is based on a relatively small sample: 187 participants in the four counties and 84 participants in San Francisco. Like other programs designed to treat psychosis in its early stages, PREP targets a narrowly defined population; a small sample is inevitably a consequence of a small target population. In addition, we were unable to confirm whether the sample is representative of the target population.

Second, we were unable to verify the quality of the data collection process or the completeness of the data. The data on psychiatric-related hospitalizations and ED visits that are maintained in Felton's CIRCE system are extracted from another data source—the county mental health department's data system. Because Felton staff in each of the five county programs extract these data manually, we were unable to assess data quality and consistency. In addition, Felton staff were unable to confirm the completeness of the hospitalization and ED data. For example, if those visits or hospitalizations were in private facilities or the facilities of other counties, they may not be reflected in the data—although Felton staff expressed confidence that most of these admissions were represented.

Third, we did not have data on the measures of stability and independence, two of Felton's primary goals for participants. In addition, our analysis was limited to a 12-month intervention period; we do not know how participants fare in the long term.

Finally, we were unable to develop a comparison group, so we do not know what would have taken place in the absence of the program. Participants were likely referred to PREP during

a particularly difficult time in the course of their disorder, as indicated by the high rates of hospitalizations and ED visits in the baseline period before they enrolled in PREP. These rates could have been lower without PREP, or with a less intensive program.

3. Qualitative findings on perceived effects

Although findings from our quantitative analysis are inconclusive, our analysis of qualitative data indicate that staff believe the PREP program had positive effects on hospitalizations, ED visits, and other quality of life indicators. All the stakeholders we interviewed—including members of the workforce, county representatives, and participants—strongly believed the participants' health and quality of life improved as a direct result of the program. Within a relatively short period of time, staff observed significant changes in participants' behavior. In the words of one staff member:

> I was attending some [program] graduations and… the gratitude they have for the program—especially, I remember one client that said, "I was having trouble at school. I was having problems finishing high school; and thanks to PREP, I was able to complete high school, and now I'm in college." So it's something that made a big impact on this guy.

D. Findings about the workforce

Three critical lessons emerged from Felton's HCIA experience. First, successful implementation requires assembling a workforce with the right attributes. Five attributes are particularly important in an effective team:

- Experience working with the target population
- An ability to think outside the box
- A willingness to address issues they were not necessarily trained to deal with
- Roots in the local community
- Cultural perspectives that correspond with those of the participants

Second, identifying highly qualified candidates in rural communities is a significant challenge, as is retaining highly trained staff in a competitive market. Third, providing remote training and clinical support is challenging, but essential to program effectiveness and work quality. We discuss each of these lessons below.

1. Assembling an effective team

Felton staff believe the success of the PREP model relies, in part, on assembling the right team. Team members we interviewed suggested that all members of the PREP team, not just the psychiatrist and clinical therapists, should have some experience working with people who have psychosis. They also believe that effectiveness is a result of every member of the team believing that each individual client can overcome his or her current condition. As one team member noted, "You must be able to see the individual as going through an illness, not as someone who is incapacitated."

Although staff are provided with intensive training and coaching from Felton's clinical training and support team, they are expected to be willing to deal with issues other than the clinical issues they were trained to address. PREP participants often have unstable housing, school-related issues, social service needs, transportation constraints, and family problems. As a result, PREP team members, regardless of their individual role, often work with participants to resolve a broad range of challenges. As one team member puts it: "There are a lot of things that we've had to do that don't pertain to our role. You end up having to deal with things you didn't sign up for and so you have to be willing to go beyond your trained role. For example, I had a participant whose family had a bad housing situation. So the participants and I went for a walk and looked for new housing options."

Members of Felton's workforce agreed that it helps if staff have roots in the local community and an understanding of the local population they serve. Those team members who were not from the county emphasized that it was critical for them to learn about the county and its resources. Doing so allowed them to serve clients more effectively and avoid the impression of being an outsider. Felton leaders emphasized that it was important for staff to be able to provide services through each client's cultural lens, which can only be learned through an understanding of the community. A member of Felton's leadership explains,

> We learned to incorporate more of the cultural responsiveness skills … whether it is having families that are not open to accessing care due to mistrust, or having a large undocumented client base and working to reassure family members that we are not the government. We are helping the young person make meaning and sense of what is happening to them, and helping them succeed in life. These are things you won't find in the books.

2. Recruiting and retaining rural team members

Identifying and retaining qualified staff remains a challenge for Felton, in part due to the intense competition for the highly skilled workforce that Felton assembles. The rural landscape of San Joaquin and Monterey Counties introduced workforce challenges that Felton did not encounter when launching programs in the original counties, which were more urban. In particular, there was a smaller pool of qualified candidates for the workforce in these communities, and few individuals willing to relocate for the positions.

In addition, San Joaquin and Monterey encompass larger geographic territory than the original three counties do, and staff must travel long distances to provide services. This burden contributed to workforce burnout and challenges retaining staff. Furthermore, the intensive and specialized training that staff receive make them attractive candidates to other organizations in the field. As a nonprofit organization, Felton is limited in its ability to be competitive, and consequently, the organization has struggled to retain staff who are enticed by more lucrative employment opportunities. Felton leaders have identified this issue as a priority.

3. Providing remote training

Although there are challenges in providing effective training and clinical supervision remotely, Felton established an effective model. Through its partnership with UCSF and an

investment in a high-tech videoconferencing system, Felton established a clinical training and support team that can provide remote support and oversight to all PREP field staff. Felton has sustained its intense training and clinical support program, in part, by using the train-the-trainer model, through which members of the local PREP office are trained to become onsite trainers. Members of Felton's workforce expressed their general satisfaction with the training they received, although at times they felt challenged by its level of intensity. They said the trainings taught them valuable skills that enhance clients' experience and improve outcomes. The intensity and frequency of the trainings were sometimes a stress to staff, as was the inability to have direct access to members of the core training team, who were not based in their office. However, these challenges did not seem to greatly reduce the perceived effectiveness of the training.

E. Program sustainability and spread

Our May 2016 phone discussion with awardee leaders revealed that after HCIA funding ended, Felton partnered with Monterey County to develop a sustainability plan for PREP by using funds available through MHSA and Medicaid reimbursement. The leadership team reported that, on the other hand, San Joaquin County did not sustain PREP; that county sought to use its limited MHSA funds for services beyond those designed to treat early psychosis. PREP programs continue to operate in San Francisco, Alameda, and San Mateo Counties, and Felton plans to explore ways to expand PREP into new counties in the future.

Staff at Felton identified several factors that contributed to PREP sustainment in Monterey, including a strong relationship with county leaders, a unique organizational structure in the county mental health department, expanded program eligibility criteria, and the availability of a long-term funding mechanism.

First, robust communication with county staff was critical to successful implementation and sustainment. In particular, the county's director of behavioral health services was a strong advocate for early psychosis treatment, and thus Felton had a true and committed partner. A second, related factor was the unique organizational structure of the county mental health department. Most counties in the state operate a three-tiered division structure, with separate divisions overseeing programs targeted to children, adults, and older adults. Monterey County, however, includes a division overseeing transition-aged youth, which is the target population of Felton's PREP program. This arrangement provided Felton with a single liaison who was invested in PREP's success, further strengthening the partnership.

A third factor that led to sustainment in Monterey County was Felton's willingness to expand program eligibility criteria to include individuals who had their first psychotic episode within five years of the referral instead of two. The cutoff time frame is somewhat arbitrary; what matters most is treating the individual before the illness progresses. Felton's clinical team prefers the two-year cutoff and maintains this cutoff in the other three county programs. However, the rural landscape of Monterey County prevents many individuals from accessing services after a psychotic episode, and thus more individuals go untreated following an episode. Expanding eligibility criteria to five years allows PREP to identify individuals who might benefit from its services and might otherwise go untreated.

The fourth, and perhaps most critical factor, is the availability of MHSA funds. MHSA is the primary funding mechanisms for the PREP program, and is particularly critical to resource-

challenged counties like Monterey. In addition, California's decision to expand Medicaid in accordance with the Affordable Care Act ensures a higher level of reimbursement for county-provided services.

Felton encountered numerous challenges in sustaining the PREP program in San Joaquin County. These included a change in county leadership, communication barriers with county staff, and insufficient county resources. Felton identified an advocate within the San Joaquin County mental health system to partner on PREP launching and implementation; however, a change in leadership required Felton to develop new partners mid-implementation. In addition, Felton did not have a single dedicated liaison within San Joaquin County. Unlike the Monterey system, the San Joaquin mental health system has no division for transition-aged youth, requiring Felton to work and coordinate with staff in two different divisions: youth and adult. This structure hindered communication during implementation and prevented Felton from establishing a stronger partnership with critical mid-level county staff.

Limited resources for competing priorities were another barrier to sustainment in San Joaquin. The county opted to use MHSA funds to contract with an organization that will provide both early psychosis services and non-specialized mental health services, such as crisis residential and inpatient services. Members of Felton's leadership team believe the county's decision will make early psychosis services available that are less intensive than PREP's program. However, they also believe the county's experience with the HCIA-supported PREP program helped advance the county's interest in providing early treatment for psychosis.

F. Lessons learned

An innovative program like PREP, which treats psychosis in its early stages, has the potential to give Medicaid beneficiaries with serious mental illness better health outcomes and enhanced quality of life at lower cost. There is extensive evidence that these programs are effective in improving clinical outcomes and lowering costs.[16] There is not enough evidence, however, that these programs have an impact on costs to Medicaid and Medicare programs specifically. To obtain such evidence, programs must be implemented in a way that enables rigorous evaluation. Limited data availability, the lack of a comparison group, and small sample size prevented us from conducting a rigorous impact evaluation of Felton's PREP program.

Analysis of qualitative data suggests that Felton implemented the program successfully. Five key factors contributed to this success: establishing an effective and efficient referral system, community outreach and education, a strong and dedicated workforce, committed county leadership, and long-term funding availability. A referral system established through partnerships with staff at various community entities, particularly school districts and mental health service providers, is critical because these are the people who are often the first to observe signs associated with the disease. Community outreach and education also have been essential to

[16] See list of published articles related to the National Institute of Mental Health's Recovery After an Initial Schizophrenia Episode (RA1SE). Available at http://www.nimh.nih.gov/health/topics/schizophrenia/raise/published-articles.shtml. Accessed August 9, 2016.

Felton's long-term success, particularly in smaller rural communities that have hitherto lacked the capacity to provide early psychosis services.

Our findings suggest that Felton has been able to identify and assemble a workforce that can effectively implement the PREP program. PREP is a team-based intervention, and successful implementation hinges on assembling a team of professionals who have experience with the target population, are flexible, can think outside the box, and if possible, have roots in the local community. However, hiring and retaining qualified staff has been a significant challenge for Felton. To help establish community partnerships and buy-in, Felton preferred to hire staff rooted in the local communities, but the pool of qualified candidates in these rural communities was relatively shallow. For several positions, Felton hired individuals who were either willing to relocate or commute long distances several days a week from outside the county.

A more significant and long-term challenge is staff retention. PREP team members receive extensive specialized training and certification on the various evidence-based practices, particularly CBTp. The knowledge and skill sets obtained in Felton's training program are not commonly available in the United States, and as a result, staff are often lured by organizations that can provide more competitive employment opportunities. Felton, like many nonprofit organizations, has not found a way to combat this challenge effectively.

Long-term sustainability of the PREP program relies on both dedicated county leadership and the availability of adequate funding. The PREP program is appealing to counties with limited capacity for addressing the significant and diverse needs of this high-cost population. However, resources for funding such programs are limited. In California, MHSA—which has provided funding to county mental health departments across the state—has proven to be an effective mechanism for supporting the PREP program. At the end of the HCIA funding period, both San Joaquin and Monterey Counties wanted to keep providing early psychosis treatment, but only Monterey opted to use its MHSA funding to renew its contract with Felton. These different outcomes can be attributed, in part, to the stronger partnership Felton maintained with leaders and mid-level staff in the Monterey County mental health department. San Joaquin County's decision not to sustain the PREP program appears to reflect a desire to use its limited MHSA funding to cover a broader range of services, not a lack of interest in treating early psychosis.

The lack of rigorous quantitative evidence of PREP's success has not prevented its continuation; the program is sustained in four counties. Analysis of qualitative data suggests that, for some participants, key outcome measures showed positive trends—information that Felton's leaders hope will support long-term sustainment. For now, leaders in four counties have decided the program is worth continuing as long as MHSA stays in place.

This page has been left blank for double-sided copying.

IV. THE FEINSTEIN INSTITUTE FOR MEDICAL RESEARCH

Findings from Mathematica's Evaluation of Feinstein's HCIA Program

- The Feinstein Institute for Medical Research implemented a program of complementary technologies and interventions that was designed to lower the costs of care and improving the mental health of 770 individuals across 10 sites.
- We were unable to assess program impacts due to a lack of complete quantitative data. However, our evaluation identified several implementation lessons that similar programs should consider.
- A program comprising an array of complementary and innovative technologies requires a well-structured road map for initiating participant use of technologies and a binding element (such as the innovation's relapse prevention planning process) to bring the components together.
- Financing arrangements that support innovative technologies and additional staff time, such as capitation, and organizational characteristics, such as affiliation with an inpatient unit, would help sustain a program similar to Feinstein's.

A. Introduction

1. Program goals

The Feinstein Institute for Medical Research, the research division of the North Shore–Long Island Jewish Health System in New York,[17] used funding from the Health Care Innovation Awards (HCIA) to implement the Improving Care–Reducing Costs (ICRC) program. This program was designed to lower the cost of care and improve the mental health status of 770 individuals with schizophrenia, schizoaffective disorder, or a psychotic disorder not otherwise specified, who were at risk for hospital readmissions. In addition to relapse prevention counseling to reduce the risk of readmissions due to mental health symptoms, the program used innovative technologies, such as an interactive smartphone application and web-based psychotherapy, to improve disease management and care.

To implement the ICRC program, Feinstein partnered with 10 community mental health centers to serve program participants in eight states: Florida, Indiana, Michigan, Missouri, New Hampshire, New Mexico, New York, and Oregon. The ICRC program trained and deployed a new cadre of health care workers at each clinic, known as mental health/health technology (MH/HT) case managers. These workers provided relapse prevention counseling and ongoing support to participants during the six-month program period. They also facilitated participant use of complementary technologies and services. In addition, the treatment team at each site included a prescriber, who was supported by an electronic decision assistant. Participants were discharged from the program at the conclusion of the six-month period.

To help assemble the package of complementary technologies and services for participants, Feinstein partnered with technology developers and researchers at six other institutions: Boston

[17] North Shore–Long Island Jewish Health System is now called Northwell Health.

University, Dartmouth College, the Nathan S. Klein Institute for Psychiatric Research, Proteus Digital Health, the University of Minnesota, and the University of Pittsburgh.

Each program site targeted adults with schizophrenia and associated disorders. Participants met the following criteria:

- Age 18 to 60
- Clinical diagnosis of schizophrenia, schizoaffective disorder, or a psychotic disorder not otherwise specified
- Currently hospitalized or within 30 days of a psychiatric hospitalization at enrollment
- Enrolled in or eligible for a Medicaid program, privately insured, or uninsured

The findings in this chapter are based on data that we collected or received by June 1, 2015, as well as enrollment data reported throughout the award period. We drew on the following data sources for this chapter:

- Enrollment data submitted by Feinstein to the reporting website that was maintained CMMI's technical assistance contractor (the Lewin Group)
- Quantitative data collected from 6 of the 10 implementing sites
- Qualitative data, including information gathered during telephone interviews and in-person site visits: In 2014 and 2015, we conducted in-depth interviews with key awardee staff, members of the workforce, and other stakeholders and, in 2015, convened focus groups with members of the workforce and with ICRC participants.
- Workforce survey data collected in 2014 and 2015

We do not report findings on CMMI's four core measures, because we were unable to gather enough quantitative data to do so. As described in Section B, we experienced challenges associated with small sample sizes and obtaining accurate and complete data from Feinstein. This report provides final evaluation results for Feinstein.

2. Overview of administrative context

Feinstein conducts clinical research in partnership with providers and patients across the health system's facilities. Before the ICRC program, leaders had significant experience with conducting research on schizophrenia and relapse prevention, in partnership with mental health sites across the country.

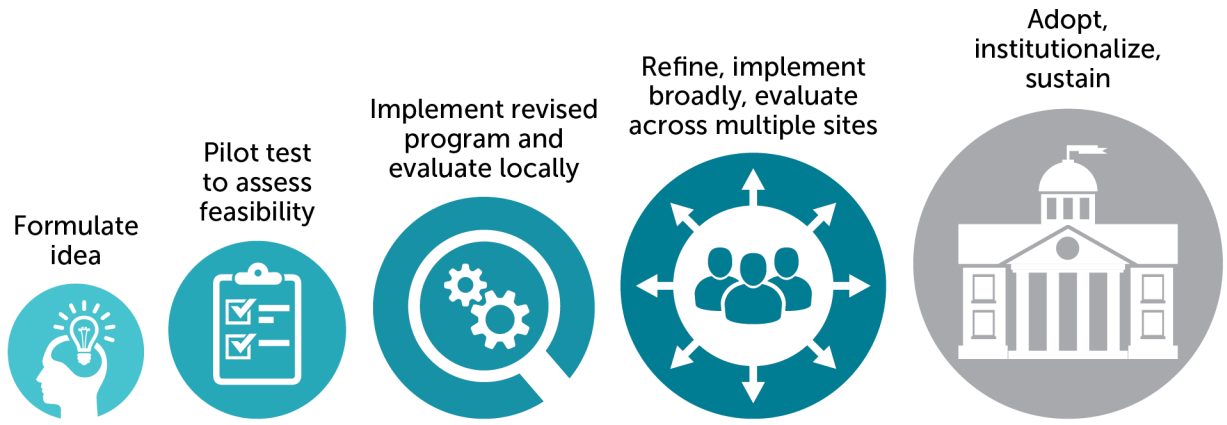
Feinstein created a central administrative leadership and support structure to facilitate implementation of the ICRC program at the 10 participating mental health centers. Feinstein selected implementing sites based on shared characteristics, such as past participation in similar research projects, as well as leadership buy-in for participation in the program. However, the policy context in which the sites operated varied by state. One of the central components of the ICRC program, the MH/HT case manager role, functioned slightly differently depending upon each state's reimbursement policies. In states where Medicaid did not reimburse case management, MH/HT case managers often provided both technology and relapse prevention

support in addition to traditional case management services. However, in states where Medicaid reimbursed case management, some MH/HT case managers provided only the program's technology and relapse prevention services, and collaborated with the participant's established case manager, who provided the traditional case management.

3. Progression through phases of innovation

The ICRC program progressed through four of the five phases of innovation over the course of the award period (Figure IV.1). The program itself was newly formulated and pilot tested at one site. The program was then refined and implemented at the other nine sites. Program leaders assessed outcomes at these nine sites by collecting and analyzing quantitative data throughout the duration of the program and by completing qualitative client interviews as participants were discharged. During our interviews, program leaders reported that the results of their internal evaluation indicated a reduction in total cost of care and hospital readmissions, among other improvements.

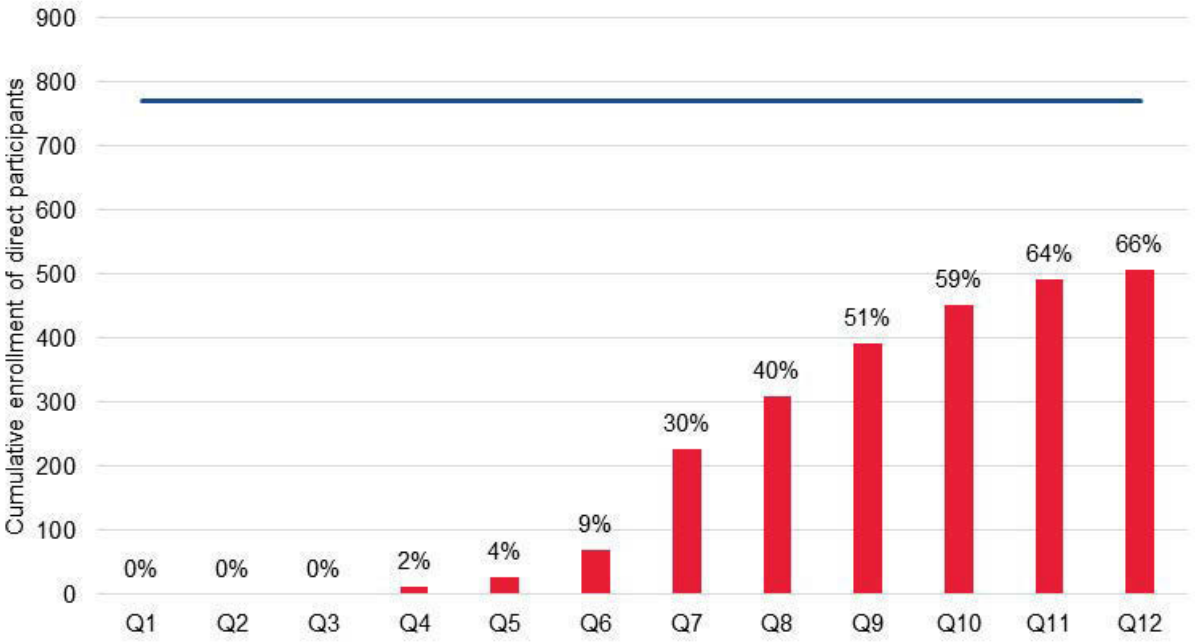
Figure IV.1. Phases of program innovation: Feinstein



4. Enrollment

By the end of the 12th quarter of the HCIA award period (June 30, 2015), Feinstein and its partners had enrolled 506 participants. Recruitment for the program was initially delayed because Feinstein staff were notified of the award of HCIA funding later than expected. Despite efforts to enhance enrollment, Feinstein ultimately did not reach its goal of 770 participants across the 10 sites (Figure IV.2).

Figure IV.2. Percent of target enrollment achieved by quarter, Q1–Q12



Source: Quarterly reports submitted to the website maintained by CMMI's technical assistance contractor (the Lewin Group).

Note: The blue horizontal line represents Feinstein's enrollment target of 770 unique participants.

Our quantitative data collection for Feinstein was limited to 6 of 10 sites. The remaining 4 sites had limited potential for quantitative analysis because they were located in states with high managed care penetration and poor reporting of managed care encounter data. At the selected sites, most participants were between the ages of 18 and 34 (59.9 percent) and were male (60.6 percent) (Table IV.1).

Table IV.1. Demographic characteristics of Feinstein participants at selected sites

	Number of participants	Percentage of participants
Total	312	100
Age		
18–34	188	59.9
35–44	63	20.1
45–54	45	14.3
55–64	18	5.7
Gender		
Female	148	39.4
Male	228	60.6

Source: Mathematica analysis of Feinstein enrollment data.

Notes: The total reflects participants through October 16, 2015, for whom a program enrollment date and demographic characteristics were available. Data were missing for 66 of the 378 participants for whom the sites provided data. These data are from six sites: Zucker Hillside Hospital, the Center for Rural and Community Behavioral Health, the Community Mental Health Center Inc., Henderson Behavioral Health, PeaceHealth, and the Mental Health Center of Greater Manchester.

B. Methods

1. Quantitative analytic methods

We were unable to use Medicaid or Medicare data to assess pre-post changes or program impacts for Feinstein due to insufficient sample sizes and challenges associated with obtaining accurate and complete data. We requested Medicare and Medicaid identification numbers and program enrollment dates for participants from 6 of the 10 Feinstein clinical sites. After discussing this issue with CMMI, we decided not to request such information from the other 4 sites because they were located in states that have high Medicaid managed care penetration and poor reporting of Medicaid managed care encounter data to CMS. Therefore, their data would likely be unusable. Across the 6 sites, we received insurance status information for 378 participants (Table IV.2).

Table IV.2. Feinstein participants with data on insurance status

Site (location)	Number of participants	Percent of participants
CMHC (IN)	31	8.2
CRCBH (NM)	18	4.8
Henderson (FL)	65	17.2
MHCGM (NH)	48	12.7
PeaceHealth (OR)	53	14.0
ZHH (NY)	163	43.1
Total	378	100

Source: Mathematica analysis of Feinstein enrollment data.

Notes: The total reflects participants through October 16, 2015, for whom data on insurance status was available. These data are from six sites: Zucker Hillside Hospital (ZHH), the Center for Rural and Community Behavioral Health, Community Mental Health Center Inc. (CMHC), Henderson Behavioral Health, PeaceHealth, and the Mental Health Center of Greater Manchester (MHCGM).

The data we received from Feinstein contained Medicare identification numbers for 121 program participants. However, only 106 of these identification numbers were valid when compared with the Medicare claims. Of the 106 individuals with valid identification numbers, 83 had a date of enrollment into the ICRC program. Of these 83 with an enrollment date, only 74 matched to Medicare administrative records. As such, we could conduct pre-post analyses on only 74 Medicare participants.

To identify Medicaid enrollees among the ICRC participants, we worked with data from Feinstein's largest single site (ZHH) and the only one in New York State. This site had a comparatively large sample size and reasonably high-quality Medicaid data. ZHH reported a total of 163 participants across all insurance categories: Medicaid, Medicare, private insurance, or no insurance. The data we received from Feinstein included Medicaid identification numbers for 89 of the 163 participants. Of these 89 participants, 73 had a valid program enrollment date. Of the 73 participants with valid enrollment dates, we were able to match 54 participants to New York State Medicaid enrollment data at the time they joined Feinstein's program.

In collaboration with CMMI, we decided not to conduct pre-post or impact analyses using Medicaid or Medicare claims for two reasons. First, these sample sizes would likely yield unreliable results and the chances of finding strong positive results were minimal. Consequently,

we could have concluded that the program had no effect when, in fact, the evaluation was not a fair test of its effects. Second, the findings from this small population probably would not be generalizable to all ICRC participants. For example, participants with accurate identification numbers may have had the most contact with the program because they had many more symptoms than other participants. Overall, the available data did not justify conducting quantitative analyses.

2. Qualitative methods

In spring 2014, we conducted in-person interviews with Feinstein's central program team and all program staff at one site (ZHH). We conducted telephone interviews with small groups of staff in each program role (MH/HT case managers, prescribers, and site project directors) at a subset of program sites. We also conducted telephone interviews with researchers who developed the program's technologies and its relapse prevention counseling component, as well as the psychologists who facilitated online forums for participants and their families on the program's daily support website. In total, we conducted 15 interviews with 25 individuals, including staff from 7 of the 10 sites, in 2014.

In spring 2015, we made a second visit to ZHH to interview the central program team and the site's program staff. We also conducted telephone interviews with program staff at a second site (PeaceHealth). In total, we conducted 14 interviews with 21 people in 2015. In addition, we held an in-person focus group with 9 program participants and a virtual focus group with 13 current and former MH/HT case managers from 9 of the 10 sites.

During both site visits, we assessed respondents' perceptions of program impacts, implementation barriers and challenges, workforce satisfaction, and relevant internal and external contexts. Although we gathered in-depth qualitative data from individuals in all of the program's workforce roles during each site visit, we were unable to collect qualitative data from all participating sites because of the large number of sites in multiple states. Therefore, our conclusions may not be representative of all sites. However, they do provide insight into what program staff and participants think about the program's implementation and effects.

C. Summative findings

As noted, we did not complete impact analyses for the Feinstein program. As such, we are unable to present effect estimates for any of CMMI's core measures. Below, we present a brief overview of our qualitative findings on the program's effects.

- Program leaders and frontline staff believed the program improved desired health outcomes for participants. As evidence of impact, leaders cited internal data that demonstrated reductions in hospital readmission rates both while participants were enrolled in the program and after the program ended. Staff mentioned symptom improvement among participants as a primary outcome of the program.
- Program staff mentioned the benefits of the program's health technologies for participants, noting that (1) the on-demand availability of the interventions, (2) the increased connectivity with providers and other family and social supports, and (3) the technologies that enhance coping were key factors in maintaining participant stability and improving outcomes. However, the program did encounter challenges in implementing at least one of its

components. Feinstein had intended to deliver the ingestible sensor component of the ICRC's health technology program to 100 participants at the ZHH site; however, it was unable to acquire the technology in time to provide it during the HCIA award period. Program leaders noted that Proteus, the company that developed the ingestible sensor and personal monitoring patch technology, was sold to another technology company (Otsuka Pharmaceuticals), which planned to refine the technology before releasing it for use.

- Nearly all interview respondents cited the MH/HT care manager role as integral to the program's success. However, the role of the MH/HT varied slightly across sites. As noted, differences in states' reimbursement practices for case management led to some differences among the sites in terms of the MH/HT case manager's responsibilities.

D. Findings about the workforce

Overall, ICRC staff appeared to value the program and its impact on participants. According to interviews and the workforce survey, program staff found the trainings they received helpful for their work. They appreciated the hands-on nature of the trainings, specifically underscoring the importance of learning techniques to manage difficult symptoms experienced by the target population. Most staff reported feeling supported in their roles, noting that the program's trainers and central leadership team were responsive to their questions and requests and helpful when challenges arose.

Interviews and workforce surveys also revealed that staff generally felt very satisfied by their work, citing (1) the overall benefit they thought the program had on client outcomes and (2) the innovative nature of the program as the primary contributors to their satisfaction. Staff identified several major barriers to their job performance, including characteristics of the target population, challenges related to the use of technology and patient engagement, and difficulty in transitioning clients after the six-month program ended.

E. Program sustainability and spread

Feinstein program leaders indicated during our second site visit (spring 2015) that they did not intend to sustain any aspects of the ICRC program after the HCIA funding concluded, noting that changes in payment policies and reimbursement mechanisms would be necessary to support continuation and expansion of the program. However, some sites and providers intended to incorporate the program's underlying concepts and the lessons learned from particular components into future practice. For example, several providers mentioned that they intended to use the core steps of the program's relapse prevention counseling and plan (such as setting goals and recognizing warning signs) with clients in the future.

Although Feinstein was unable to sustain the ICRC program, leaders suggested that sustaining a program similar to this one would require an organization that is structured so that program service costs are eventually offset by cost reductions elsewhere. In addition, Feinstein staff suggested that organizations that receive capitated payments or similar types of financing may be better positioned to sustain the program because these types of financing mechanisms allow for some flexibility to cover enhanced provider time for prescribers' use of the decision assistant as well as case managers to implement the program.

F. Lessons learned

In the absence of sufficient quantitative data, we were unable to assess program impact. Based on the available qualitative evidence, however, we can draw the following conclusions about implementation that may be useful for future initiatives of this nature:

- Program leaders suggested that a program such as the ICRC program is likely to be most successful in clinics structured and financed in specific ways. Program leaders and staff at participating sites suggested that the program was likely to be most successful at sites that offered a full continuum of acute and outpatient behavioral health services. Clinics with both inpatient and outpatient units that are paid a capitated rate (and are therefore responsible for a client's inpatient and outpatient care) are most likely to benefit from reductions in readmissions. In addition, leaders noted that sites that offer a full continuum of services or that are affiliated with larger hospitals or health systems tended to have greater access to resources as well as an easier time identifying and enrolling participants. Sites without such affiliations had to conduct more extensive outreach to get referrals and recruit participants from local hospitals and other external providers.
- Program leaders suggested that in some states, current financing arrangements for the program's target population did not accommodate the extended prescriber and case manager time the program requires, nor the administrative or "nonproductive" time needed to implement the program. Given this, leaders found themselves unable to sustain the program under current policy and reimbursement structures.
- A program comprising an array of complementary and innovative technologies requires a well-structured road map for initiation of the technologies as well as a binding element. In this case, the MH/HT case manager served as that element by using structured relapse prevention counseling and pulling the diverse technologies together to ensure they were tailored to each client. However, it is difficult to determine whether perceived improvements in patient outcomes were due to any specific program components or were the result of an additive effect when multiple components were combined.

V. FUND FOR PUBLIC HEALTH IN NEW YORK[18]

Findings from Mathematica's Evaluation of the Fund for Public Health in New York's HCIA Program

- For Medicaid enrollees, Parachute NYC's crisis respite services were associated with fewer hospitalizations and lower costs, but its need-adapted mobile crisis teams (NA-MCTs) were associated with more hospitalizations and higher costs.
- Peer support was a key feature of the program, but integrating peer support specialists into the NA-MCTs and establishing their role in Parachute NYC were challenging.
- New York's Medicaid reform offered a unique opportunity to build a sustainable payment model into state legislation.
- As of April 2016, all components of the Parachute NYC program are being sustained, with some modifications.

A. Introduction

The Fund for Public Health in New York (FPHNY), a nonprofit organization dedicated to improving the health and well-being of city residents, partnered with the Division of Mental Hygiene in New York City's Department of Health and Mental Hygiene to implement Parachute NYC. This project focused on adults in New York City who experienced a mental health crisis and an episode of psychosis or severe mental illness. Parachute NYC was designed to give them better care at a lower cost by moving beyond the crisis model of care and focusing on patient-centered care; long-term, community-integrated treatment; and better access to primary care services.

Parachute NYC had three main components:

- **Need-adapted mobile crisis teams (NA-MCTs).** NA-MCTs consisted of clinicians and peers who provided in-home mental health services to participants in each of four boroughs (Brooklyn, the Bronx, Manhattan, and Queens). The NA-MCTs provided psychosocial education, psychotherapy, peer support, and referral to community services.
- **Crisis respite centers (CRCs).** CRCs provided a supportive, safe environment for individuals experiencing or anticipating a psychiatric crisis. Throughout the four boroughs, the CRCs offered 24-hour peer support, education in self-advocacy, and training in self-help. The CRCs were designed to be a short-term alternative to hospitalization where participants could stay for up to fourteen days.
- **Support line.** The citywide "warm support line" was a confidential phone service operated by peer staff; it offered counseling and referral services to callers in emotional distress.

[18] We thank New York State Department of Health (NYSDOH) for providing Medicaid data to support these analyses. The findings and conclusions presented are those of Mathematica Policy Research alone and not those of NYSDOH.

Parachute NYC was the first large-scale implementation of the need-adapted treatment model (NATM) in the United States. This model integrated a multidisciplinary team with the client's personal support network and incorporated ongoing support and follow-up. FPHNY designed the Parachute NYC model to test the hypothesis that adding intentional peer support (IPS) to mobile crisis teams would help avoid hospitalizations and use of the emergency department (ED).

Although the Parachute NYC program introduced the CRCs, the mobile crisis teams were in place before the program began. These teams, managed by different health service agencies in New York City, provided rapid assessment and short-term, in-home counseling and referrals to people who were experiencing a psychiatric crisis. FPHNY contracted with these agencies, and Parachute NYC trained the teams on the new IPS and NA-MCT treatment modalities. To help fulfill the program's mission, the mobile crisis teams incorporated peers into their teams and treatment practices. The teams offered enhanced Parachute NYC services to program participants, while continuing to provide their traditional short-term services to clients who were not participating in the program.

This chapter is a comprehensive summary of the findings of our evaluation of Parachute NYC to date. We will present results of additional analyses in an addendum to this report. To conduct our evaluation, we drew on the following data sources:

- Enrollment data submitted by FPHNY to the reporting website maintained by CMMI's technical assistance contractor (the Lewin Group) for the HCIA, Round 1.
- Data from workforce surveys conducted in the spring of 2014 and 2015 to gather information about staff burnout and stress, job satisfaction, and perceptions of training and support.
- Quantitative data on participant enrollment, utilization, and expenditures, extracted from Medicaid claims and enrollment information using participant identifiers provided by FPHNY.
- Qualitative data, including phone interviews and in-person site visits conducted in spring 2014 and 2015. We conducted in-depth interviews with awardee leaders, members of the workforce, and other stakeholders. In 2015, we also convened focus groups with peers, clinicians, and program participants.

1. Overview of administrative context

FPHNY is a nonprofit organization formed by the New York City Department of Health and Mental Hygiene. City staff directed and managed FPHNY projects, and FPHNY staff provided administrative and fiscal oversight. This relationship with the City department enabled FPHNY to advocate for key policy changes in New York, while also allowing the organization to be more flexible and agile than a government agency can typically be.

FPHNY contracted with various agencies to provide the Parachute NYC CRC and NA-MCT services. Each agency was independently managed and largely responsible for its own recruitment and hiring. The CRCs and NA-MCTs had separate staff and served different populations in different boroughs. The four NA-MCTs consisted of both peer support specialists

and an interdisciplinary team of clinicians (psychologists, social workers, psychiatrists, and family therapists) who provided therapeutic expertise. The NA-MCTs worked with participants to develop and implement an individualized action plan, consulted with family members, and provided advocacy and guidance to participants as they navigated the health care system.

In the CRCs, peer support specialists were the main workforce. The CRCs provided an alternative to hospitalization for individuals who needed temporary residential or respite care. CRC peers provided 24-hour support, education in self-advocacy, and training on self-help. Peers in the Manhattan CRC ran the support line in addition to performing their CRC duties.

Each of the CRCs and NA-MCTs had supervisors on staff. FPHNY also hired a consultant, trainers, and a program evaluation team to provide more support and guidance. In the final year of the award, FPHNY contracted with a registered nurse to provide basic physical health screening, health promotion, and education to all the CRCs.

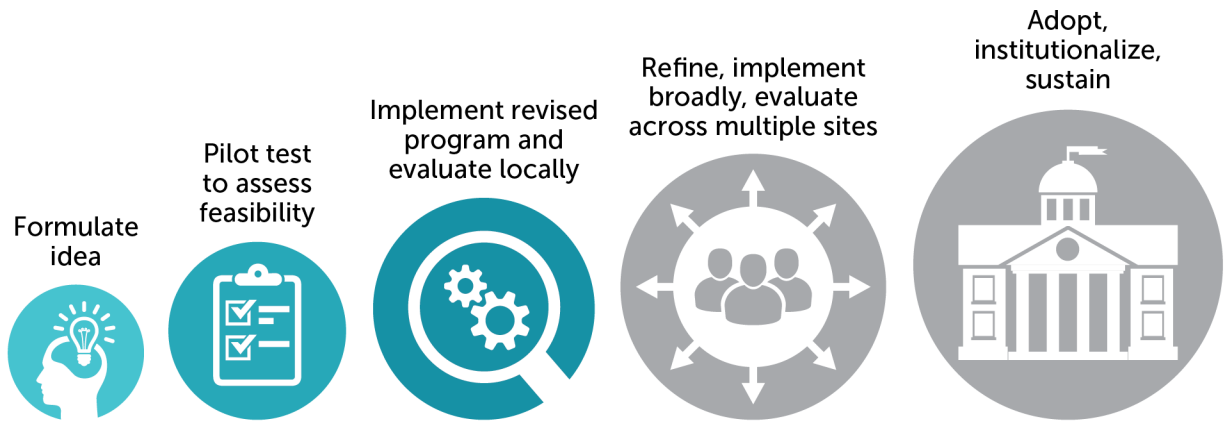
Each CRC and NA-MCT also hired peer support specialists at the start of the program. Individuals in the peer role had personal experience with mental illness and recovery. Before the program began, the existing mobile crisis teams were composed exclusively of clinicians. The Parachute NYC program incorporated peer staff into these teams.

2. Progression through phases of innovation

Over the course of the award period, Parachute NYC moved through the first three phases of innovation delineated in Figure V.1. Drawing from existing models overseas and in the United States, Parachute NYC leaders developed the core concepts of the program, adapting it to New York's unique environment. FPHNY used HCIA funding to establish the new CRC, NA-MCT, and support line services. The awardee adapted the existing mobile crisis teams to the NA-MCT model and hired peer specialists for the teams. Parachute NYC also established new CRCs staffed by peers. This was the first time peer specialists had been integrated into mobile crisis team and crisis respite center services in New York City, as well as the first use of the NATM.

Although the project did not include a formal pilot phase, FPHNY rolled out Parachute NYC services in Manhattan slightly ahead of the other three boroughs. This allowed program leaders to learn from the experiences of staff and administrators in Manhattan. Over the course of the program, Parachute NYC leaders made changes to improve the program training and enrollment criteria. For example, the CRCs were initially intended for people experiencing psychosis-related crisis, but when these criteria resulted in low referral rates, services were quickly expanded to include individuals experiencing emotional crisis, depression, and anxiety. By the end of the award period, FPHNY had established its program and demonstrated the viability of its service model within the target populations, achieving higher referral rates and consistent service use. The awardee plans to continue all aspects of the innovation, but will make some changes to the implementation approach.

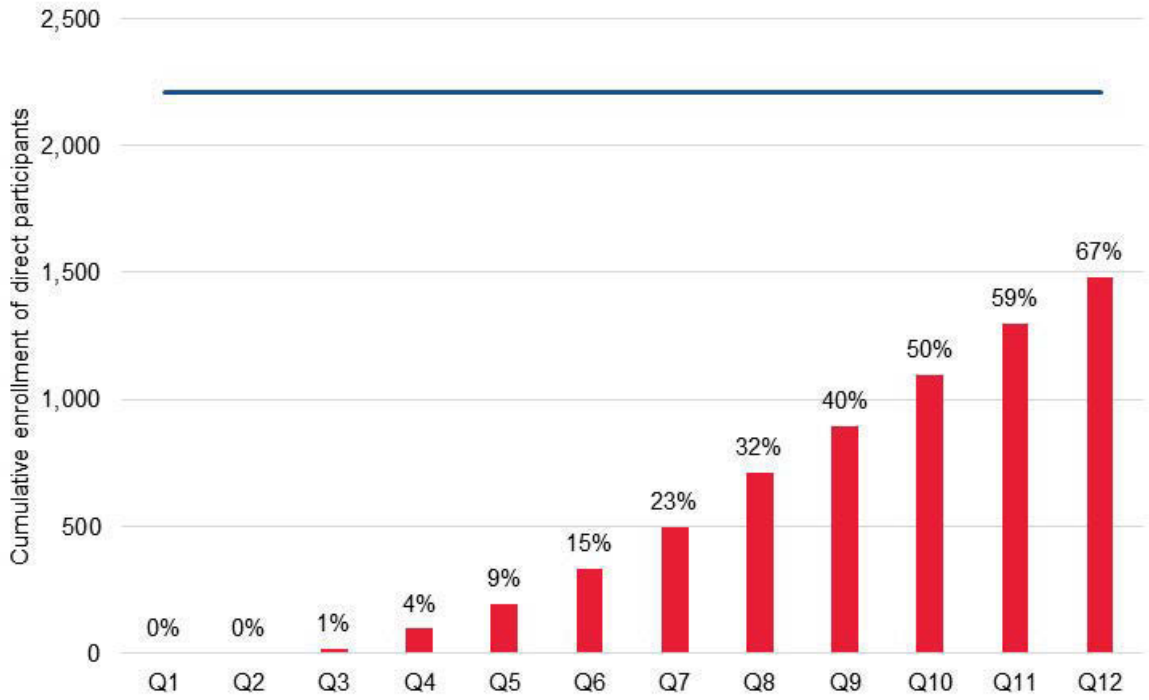
Figure V.1. Phases of program innovation: FPHNY



3. Enrollment

By the end of the 12th quarter (June 30, 2015), Parachute NYC had provided services to 1,480 participants, or 67 percent of its enrollment target of 2,208 (Figure V.2). Project leaders said that, in retrospect, their initial enrollment goal was overly optimistic.

Figure V.2. Percent of target enrollment achieved by quarter, Q1-Q12



Source: Mathematica analysis of program enrollment data provided by FPHNY.

Note: The blue horizontal line represents FPHNY's enrollment target of 2,208 unique participants.

4. Participants' demographic characteristics

When they first used Parachute NYC's services, almost all participants were under age 65; about two-thirds were under age 45 (Table V.1). About half the participants were enrolled in Medicaid; about 40 percent either had no information available on their insurance status or had no insurance coverage; almost 6 percent had private insurance; and fewer than 4 percent were Medicare beneficiaries.

Table V.1. Demographic characteristics at initial service use, FPHNY participants, January 2013–June 2015

	Number	Percent
Total	1,480	100.0
Age[a]		
Less than 18	72	4.9
18–34	671	45.4
35–44	234	15.8
45–54	292	19.7
55–64	185	12.5
65 or older	25	1.7
Gender[b]		
Female	631	43.0
Male	838	57.0
Insurance coverage[c]		
Medicaid, non-dual	767	51.8
Medicare, non-dual	31	2.1
Dual	19	1.3
Private insurance	83	5.6
Other (includes other insurance, unknown and uninsured)[d]	580	39.2

Source: Program administrative data provided by FPHNY, October 2015.

[a] Information on age was missing for one member of the intervention group.

[b] Information on gender was missing for 11 members of the intervention group.

[c] FPHNY provided binary indicators for whether an intervention group member had Medicaid, Medicare, private insurance, and/or "other" insurance. If a member of the intervention group was identified as having multiple types of insurance, we categorized the person as follows: dual (19 members with both Medicaid and Medicare), Medicare (one member with Medicare and "other" insurance), private (34 members with private insurance and either Medicaid or "other"), or "other" (3 members with other insurance and both Medicaid and private insurance).

[d] Of the intervention group members classified as "other," 206 (13.9 percent of the total sample) had missing insurance status. Nearly everyone else in this group had no insurance coverage indicated; however, the reliability of this reported information was not verified.

B. Methods

1. Quantitative methods

We conducted an impact analysis with a difference-in-differences time series model and a matched comparison group. For this analysis, we used Medicaid claims and administrative data provided by the New York State Department of Health. Therefore, the analysis was limited to intervention group members enrolled in Medicaid. The analysis focused on three of the four core

outcomes that CMMI prioritized for all HCIA awardees: total Medicaid expenditures,[19] hospitalizations, and ED visits.

FPHNY staff reported that 767 participants were on Medicaid. However, when we searched the New York State Medicaid files, we discovered that the FPHNY and state files had consistent identifying information for only 507 of these individuals.[20] Of these 507, we had to exclude 260 (slightly more than half), leaving an analytic sample of 247 participants. We excluded the 260 individuals for at least one of the following reasons:

- They were dually enrolled in Medicare, did not have full coverage for the state's Medicaid benefit package, or had third-party coverage. We took this step to ensure that all individuals in our analytic sample had consistent coverage during the analysis period for services in the NY state Medicaid benefit package including hospitalizations and ED visits.
- They were not enrolled in Medicaid for at least six months before and at least five months after they received Parachute NYC services. This ensured that we had enough data in the pre- and post-intervention periods for each individual included in the analysis.[21]
- They did not have a behavioral health diagnosis in the claims data.

We then took the following steps to identify a well-matched comparison group of Medicaid beneficiaries:

- We selected a pool of potential comparison group members who resembled the intervention population in the following respects: (1) resided in New York City, (2) were not enrolled in Medicare, and (3) had at least one claim with a behavioral health diagnosis between January 2012 and December 2014.
- For each potential comparison group member in each month between January 2013 and January 2015, we identified use of (1) inpatient, (2) ED, (3) psychiatric, or (4) non-psychiatric office services with a behavioral health diagnosis. Then, from among the months in which the comparison group member received a behavioral health service, we randomly selected a pseudo-enrollment month (that is, a month they could have been enrolled in Parachute NYC if it had been available to them) for each comparison group member. This random selection was weighted such that the distribution of program enrollment and pseudo-enrollment months for intervention and potential comparison group members, respectively, were proportionally similar across calendar months. For the intervention and potential comparison pool, the program enrollment month and the pseudo-enrollment month were deemed the first month in the intervention period in our analysis.

[19] These expenditures include both fee-for-service and managed care payments. When service level payment information was not available for managed care covered services, we estimated these payment amounts based on fee-for-service payment guidelines.

[20] Identifying information was deemed consistent if three out of four of the following fields matched: gender, day of birth, month of birth, and year of birth.

[21] We will be able to substantially expand the size of the analytic sample in future analyses because additional data covering a longer time period will become available. We will report our findings in the addendum to this report, which we will submit to CMMI in March 2017.

- We then retained only those potential comparison pool members who had full-benefit Medicaid enrollment and Medicaid as their primary payer for at least six months before and at least five months after the pseudo-enrollment month.[22] Based on these criteria, over 258,000 individuals were included in the potential comparison pool.
- After defining the intervention group and the potential comparison pool, we used matching methods to select a narrower comparison group that was comparable to the intervention population in the baseline period. We matched up to 20 members of the comparison pool to each intervention group member, using a two-stage process:
  - In the first stage, the matching algorithm matched the intervention group members who first used CRCs to members of the comparison pool with an inpatient stay in their pseudo-enrollment month.[23]
  - In the second stage, we used the remaining potential comparison pool members (excluding those beneficiaries that were matched to the CRC intervention subgroup) to search for matches for the intervention group members who first used the NA-MCT.
- We matched on the following characteristics:
  - Program enrollment or pseudo-enrollment month and type of BH service used in that month (inpatient, ED, psychiatric, or non-psychiatric office service)
  - Volume of BH service use (again, inpatient, ED visits, psychiatric, and non-psychiatric office service) in the year prior to program enrollment or pseudo-enrollment
  - Mental and physical health diagnoses
  - Demographics (age, gender, race/ethnicity)
  - Disability status
  - Chronic Illness & Disability Payment System (CDPS) condition indicators
  - Full continuous year of Medicaid enrollment prior to program enrollment or pseudo-enrollment
- We ran matching diagnostic statistics, which indicated a strong match.

Further information on the methodology used to construct groups for the impact analysis is in Appendix A.

2. Qualitative methods

We conducted site visits to Parachute NYC in spring 2014 and 2015. During our spring 2014 site visits, we conducted in-depth interviews with Parachute NYC administrators at the Department of Health and Mental Hygiene; CRC and NA-MCT supervisors; and peer support specialists, clinicians, and other stakeholders such as trainers and members of the evaluation

[22] This ensured some data for analysis for each comparison group member in the pre- and post-intervention periods.

[23] Intervention group members using CRC services at enrollment were matched to comparison pool individuals with an inpatient stay in their pseudo-enrollment month because CRC services were provided to individuals who required out-of-home care substituting for hospitalization.

team. In total, we conducted 11 interviews with 24 individuals, including seven peer specialists and three clinicians.

In spring 2015, we conducted a second round of in-depth interviews with the same types of team members we talked to the previous year (but not necessarily the same people, because of staff turnover or changes in staff roles). This included 12 interviews with 22 individuals, including nine peer specialists and four clinicians. During the second site visit, we also convened three focus groups, one with eight adults who received Parachute NYC services, one with seven clinicians, and one with eight peer specialists. Overlap between interviews and focus groups was limited to two clinicians and two peers who each participated in both an interview and a focus group.

During the interviews, we discussed implementation progress and challenges, staffing, target population, program resources, workforce development, and program leadership. In 2015, we added questions about perceived program effects and sustainability plans.

C. Summative findings

1. Descriptive analyses

Before developing impact estimates for the intervention and comparison groups, we analyzed the trends in regression-adjusted means for the following core outcome measures: total Medicaid expenditures, hospitalizations, and ED visits. We examined trends in the three years before the program enrollment or pseudo-enrollment month and in the year following that month.

We conducted separate analyses for two subgroups of participants: those receiving CRC services and those receiving NA-MCT services. We decided to examine these groups separately because of the substantial differences in these two program components, and because fewer than 20 participants received both types of services. Those receiving both services were assigned to the subgroup for the service they received first. About half received CRC services first and were assigned to the CRC subgroup; the others were assigned to the NA-MCT subgroup.

Medicaid expenditures. For both the CRC and NA-MCT subgroups, average total Medicaid expenditures were similar and generally rose over time for the intervention and comparison groups in the baseline period.

- Figure V.3 shows total per-person Medicaid expenditures during the baseline and post-intervention periods for the CRC subgroup. The average difference between the intervention and comparison groups during the baseline period is close to zero. At the I1 measurement point, however, expenditures trend sharply upward for the comparison group, but not for the intervention group. It is possible that, had they not received CRC services, the intervention group's expenditures would have continued to be like the comparison group's expenditures (as they were in the baseline period). This did not happen, suggesting that the CRC intervention may have prevented the upward trend. Because of the way we selected comparison group members, they (like the intervention group) were having a crisis at the start of the I1 period. Unlike the intervention group, however, members of the comparison group would have received costly hospital or ER services to deal with the crisis; in contrast, the intervention group would have received less expensive care at the CRC. Then when the

crisis passed, the comparison group goes back to previous service use patterns—as indicated at the I2 measurement point. Overall, this pattern suggests that, in the absence of the CRC, intervention participants would also have used more costly inpatient and emergency services.

- Figure V.4 shows total Medicaid expenditures during the baseline and post-intervention period for the NA-MCT subgroup. As for the CRC, the trend lines for both the NA-MCT and comparison groups are similar during the baseline period. At the I1 measurement point, however, the intervention group's line trends upward, leading to a significant difference from the comparison group. Although expenditures decrease for both groups at I2, the difference between them remains significantly different than the average difference between them during the baseline period. This suggests provision of NA-MCT services may have led to an increase in per-person Medicaid expenditures.

Figure V.3. Total Medicaid expenditures per FPHNY CRC participant per 6-month period

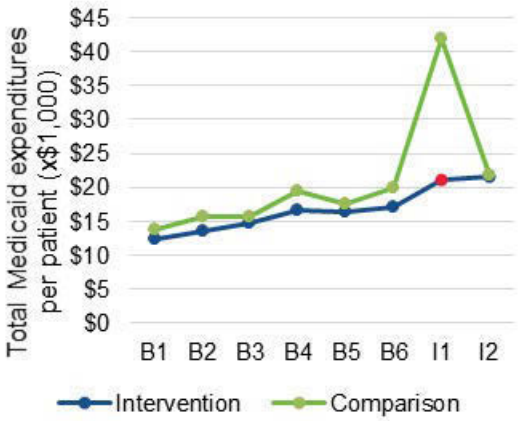
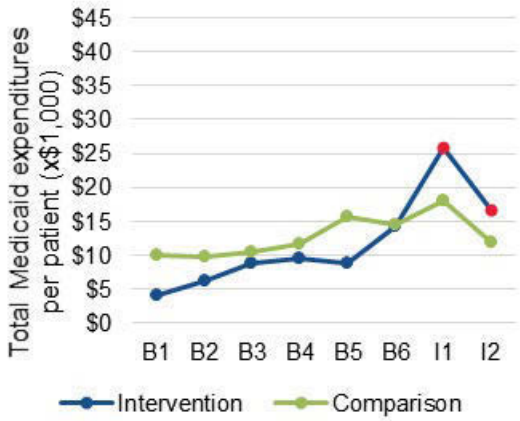


Figure V.4. Total Medicaid expenditures per FPHNY NA-MCT participant per 6-month period



Source: Mathematica analysis of Medicaid administrative data, January 2010–June 2015.

Note: Means are regression-adjusted and are given in thousands of dollars per person. The regression model controlled for age (linear and squared), gender, race/ethnicity, whether 12 months of baseline data were available, disability status, Chronic Illness & Disability Payment System condition indicators, calendar month and year of program enrollment, and diagnoses at enrollment. Red dots indicate significant difference-in-differences estimates for the given intervention period relative to the average over all baseline periods.

Hospitalizations. The pattern of hospitalization rates (see Figures V.5 and V.6) was similar to that of expenditures, suggesting that hospitalization rates may be driving expenditure patterns. Once again, during the baseline periods, the difference between the intervention and comparison groups was small. For the CRC subgroup, the difference increased at I1 because of the growing hospitalization rate for the comparison group. For the NA-MCT subgroup analysis, the difference widened because the hospitalization rate for the intervention group rose higher than the comparison group's.

Figure V.5. Hospitalizations per FPHNY CRC participant per 6-month period

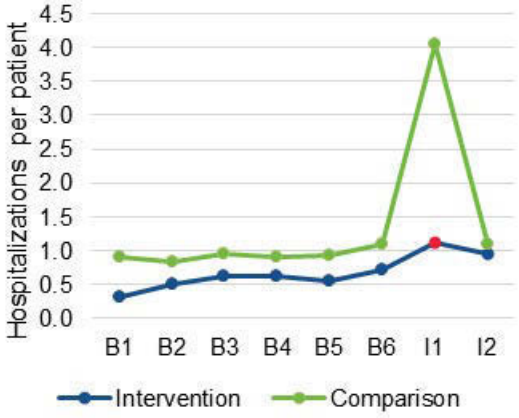
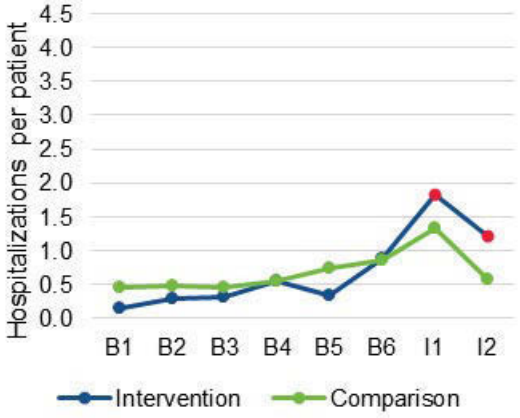


Figure V.6. Hospitalizations per FPHNY NA-MCT participant per 6-month period



Source: Mathematica analysis of Medicaid administrative data, January 2010–June 2015.

Note: Means are regression-adjusted. The regression model controlled for age (linear and squared), gender, race/ethnicity, whether 12 months of baseline data were available, disability status, CDPS condition indicators, calendar month and year of program enrollment, and diagnoses at enrollment. Red dots indicate significant difference-in-differences estimates for the given intervention period relative to the average over all baseline periods.

ED visits. The trend lines for the intervention and comparison groups closely parallel each other for both the CRC and NA-MCT participants (Figures V.7 and V.8). The ED visit rate for both groups increased during the baseline period, then declined after program enrollment or pseudo-enrollment. However, the differences were not significant, thus offering no evidence that FPHNY's program had any effect on ED visits.

Figure V.7. ED visits per FPHNY CRC participant per 6-month period

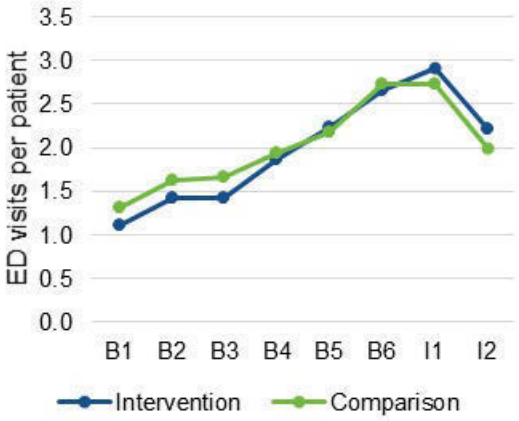
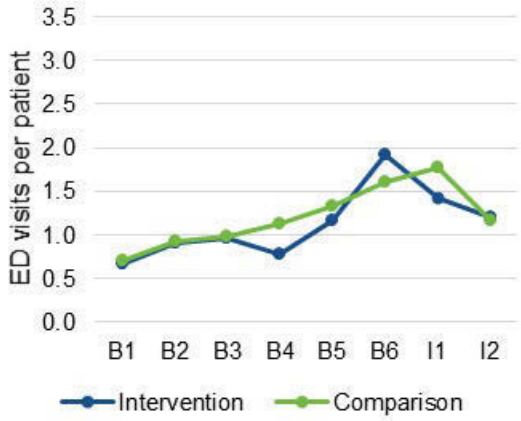


Figure V.8. ED visits per FPHNY NA-MCT participant per 6-month period



Source: Mathematica analysis of Medicaid administrative data, January 2010–June 2015.

Note: Means are regression-adjusted. The regression model controlled for age (linear and squared), gender, race/ethnicity, whether 12 months of baseline data were available, disability status, CDPS condition indicators, calendar month and year of program enrollment, and diagnoses at enrollment. Red dots indicate significant difference-in-differences estimates for the given intervention period relative to the average over all baseline periods.

2. Description of impact estimates

Findings from our impact analysis (Table V.2), which used an interrupted time-series model, corroborate the results of the descriptive analysis.[24] Specifically, we note the following:

- For all 247 participants in the analytic sample (combining data from the CRC and NA-MCT subgroups), the impact analysis revealed that the program significantly reduced the number of hospitalizations by 256 ($p < .001$).
- The substantial reduction in hospitalizations for the CRC group was partially offset by a small (but still significant) increase for the NA-MCT group.
- Possibly as a result of the changes in hospitalization rates, total Medicaid expenditures diminished significantly for the CRC group, but rose significantly for the NA-MCT group.
- Our analysis revealed no significant impacts of the program on ED use for either the CRC or the NA-MCT group.

[24] In contrast with the descriptive analysis, in which we report means for 6-month periods, the impact regressions were based on 12-month periods, including three baseline 12-month periods and one 12-month intervention period.

Table V.2. Impacts and total savings attributable to the FPHNY program

	All intervention group members		CRC only		NA-MCT only	
	Change	90% confidence interval	Change	90% confidence interval	Change	90% confidence interval
Aggregate results[a]						
Total Medicaid expenditures (in thousands)	-$1,369	[-$2,741, $1]	-$3,143	[-$4,115, -$2,172]	$1,794	[$842, $2,747]
Hospitalizations	-256	[-421, -91]	-374	[-523, -225]	120	[45, 195]
ED visits	-125	[-30, 280]	126	[-14, 266]	-1	[-63, 60]
Per beneficiary month						
Total Medicaid expenditures	-$502	[-$1,006, $1]	-$1,797	[-$2,353, -$1,242]	$1,838	[$862, $2,814]
Hospitalizations	-0.09	[-0.15, -0.03]	-0.21	[-0.30, -0.13]	0.12	[0.05, 0.20]
ED visits	0.05	[-0.01, 0.10]	0.07	[-0.01, 0.15]	0.00	[-0.06, 0.06]
Number of participants	247		163		84	
Mean number of intervention months per participant	11.0		10.7		11.6	
Approximate proportion of intervention population represented in analysis[b]	16.7%		n.a.		n.a.	
Intervention period	January 2013 through June 2015					

Source: Mathematica analysis of New York State Medicaid administrative data, January 2010–June 2015.

Note: Regression-adjusted means for intervention population based on population characteristics in I1 (N = 247). Regression model controlled for age (linear and squared), gender, race/ethnicity, whether 12 months of baseline data were available, disability status, CDPS condition indicators, calendar month and year of program enrollment, and diagnoses at enrollment. Analysis is limited to the subset of program enrollees who were Medicaid enrolled and observable in fee-for-service Medicare claims data for six months before and following intervention enrollment and who were not dually eligible for Medicaid and Medicare. The confidence intervals for all outcome measures were adjusted for multiple testing.

[a] Aggregate results are limited to the subset of intervention group members included in the analysis (247 individuals) and do not represent all program participants.

[b] We calculated the approximate proportion by dividing the number of intervention group members (247) in the analysis by the number of individuals who participated in FPHNY's program between January 2013 and June 2015 (1,480).

n.a. = not applicable.

3. Limitations of the analysis

Several limitations of the analysis should be considered when interpreting the findings:

- **Small sample size.** A small sample size and high variability—particularly in total expenditures—limits our ability to detect small differences between the intervention and comparison groups.
- **Representativeness of sample.** The 247 individuals included in our analyses are about 17 percent of the total population enrolled in Parachute NYC. The program may have had different effects on participants not included in the analyses. (For the addendum, we will provide a table comparing key characteristics of the analytic sample with all program participants.)
- **Limited time frame.** We were able to obtain data for our sample and examine outcomes for only up to one year after enrollment. Longer term impacts may differ.
- **Lack of information on site enrollment.** Data provided by FPHNY did not assign program enrollees to a given site; therefore, we could not control for unobservable differences within the intervention group that may have arisen because of differences in the way different sites implemented the program. Impacts may have been different from one site to the next. However, we likely would not have conducted site-specific analyses even if we had this information because further divisions of the analytic sample would have exacerbated the limitations imposed by small sample sizes.
- **Unobservable differences between intervention and comparison groups.** The matching methods we used to select the comparison group for this analysis may not have fully accounted for unobservable differences between the intervention and comparison groups, such as different resources available at different residential locations. These unobservable differences may bias impact estimates in unknown ways.

4. Qualitative findings on the perceived effects of the program

Results of our quantitative analyses indicated that Parachute NYC resulted in significant reductions in hospitalizations and expenditures for CRC participants, which corresponds with Parachute NYC staff members' and leaders' expectations for the program. During site visit interviews, Parachute NYC staff and leaders emphasized that one of the main goals of the CRCs was to provide an alternative to hospitalization for individuals experiencing psychiatric crisis. Many staff members believed the program would lead to lower rates of hospitalization and diminishing health care costs among participants.

Findings from the participant focus group offer insight into why and how the CRCs had positive effects. Most members of the participant focus group had received CRC services and had good things to say about them. Participants valued the CRC services and told us the CRC's environment gave them a better quality of life when they were experiencing mental health crises during which they could still take care of themselves. For example:

- Participants described the CRC as a safe escape from the stressors of their home environment. One respondent said, "It took me out of the environment that was causing the

triggers … that was the most important part." Another described his stay at the CRC as "a time-out so I could reenergize my goals and focus on me and get my life together."

- Participants particularly appreciated the homelike and relaxed atmosphere of the CRC, which they described as comfortable and welcoming. One respondent explained, "They don't bother you, if you need to talk, they talk to you ... you go in and out like you [do] regularly at home." Another respondent also described feeling at home and respected in the CRC. "They aren't telling me what to do. If anything, they [ask], 'What do you want to do?'"
- Participants appreciated the freedom offered by the CRC, particularly in comparison with the restrictions on freedom in an inpatient psychiatric facility. Several believed the CRC helped reduce their use of inpatient psychiatric or ED services. One participant commented, "It's another alternative for me because I spent … the majority of my life in and out of the hospital." Another respondent likened the inpatient hospital to jail and shared, "I welcome the Parachute program because no one [wants] to be locked in."

Participants agreed that they valued their experience in the program and it had changed their lives for the better. Several described how Parachute NYC helped bring about a positive shift in their thinking, made them feel more confident about handling their day-to-day challenges, and helped them become more aware of the stressors and triggers in their home environments. For example:

- Participants described learning new skills to recover and prevent relapse. One individual said his experience in the program allowed him to "focus on more positive ways of dealing with my life, so I can move on to a better life."
- Another respondent discussed how the CRC gave him the tools to better manage his anger and encouraged him to develop alternative coping mechanisms, such as going for a bike ride or walking his dog.

FPHNY participants also talked about the discussion groups and recreational activities that several CRCs offered. These opportunities to socialize with other CRC residents and learn from each other's experiences helped their moods and outlook on life. The topics of these discussions were generally decided by the group and included a range of themes, such as relapse prevention, stress reduction, and anger management. Participants found these discussion groups helpful and especially appreciated their informal, voluntary nature. Participants also liked the different recreational activities offered at the CRCs, such as cooking and art classes, game nights, and movie nights, and believed these activities enriched their experience at the respite center.

Several staff members thought the CRC participants' physical health habits got better as a result of their exposure to the program. One staff member also noted some small changes in participants' knowledge about their own physical health, including a heightened awareness of the importance of nutrition and regular primary care checkups. Finally, several Parachute NYC staff in the CRCs also reported that participants were more capable of advocating for themselves about their medication. Specifically, participants learned to talk to their doctors and care teams about lowering dosages to decrease the number and degree of side effects.

Although our quantitative results indicated that the NA-MCTs may have contributed to increases in hospitalizations and expenditures, qualitative results suggested that this aspect of the program may have also had positive effects on program participants' well-being and quality of life. Some staff members discussed NA-MCT participants who went back to work or school, were less isolated socially, and improved their relationships with family and friends. Several peers and clinicians said that participants often seemed more independent and "vibrant" after Parachute NYC. Staff members believed NA-MCT services gave participants hope and empowered them to change their lives. For example, a NA-MCT clinician said:

> In my experiences, the Parachute program impacts the patients that we see or the clients that we see very positively. I've seen people gain employment through this type of program. I've seen people who were not very talkative become very talkative, very expressive, and very emotional, so it's really made a promising impact on their lives where it gives them opportunity to kind of think outside the box and be more engaging.

D. Findings about the workforce

During our site visit interviews and focus groups with staff, we discussed workforce development, deployment, and training. We also conducted a workforce survey in the spring of 2014 and 2015 to gather information about staff burnout and stress, job satisfaction, and perceptions of training and support. The surveys and discussions gave us insight into the experience of FPHNY's workforce and their effectiveness in fulfilling the program's goals. Specific themes of peer integration, clinician experience, workforce training, and job satisfaction emerged from the discussions and surveys.

1. Peer integration

Parachute NYC sought to create a new workforce model to treat mental health by incorporating peer support specialists into existing mobile teams. During interviews, many staff members told us that at first, integrating peer support staff was a challenge. The program leaders were able to find strategies to address those challenges. In our 2015 interviews, respondents agreed that the integration of peers in Parachute NYC got better over time and that the role was an integral part of Parachute NYC.

Integrating peers into the CRCs and the existing mobile teams was a challenge. Several members of the NA-MCTs, both peers and clinicians, said they initially had trouble working together. Especially in the early stages of implementation, peers sensed a lack of respect for their opinions and did not think they were truly accepted in the teams or that their contributions were considered essential. Clinicians sensed they were being blamed for the peers' experiences with the mental health system. Clinicians also said they believed the peers were still in the process of recovery, which made the clinicians worry that the peers would relapse. All these issues occasionally caused tension between clinicians and peers. CRC staff also reported difficulties incorporating the peers into teams at the beginning of the project. In particular, CRC leaders had to balance holding peers accountable for job duties, such as coming to work consistently, and allowing for the fact that peers were in active recovery.

Understanding the characteristics of effective peers was important to helping peers succeed. CRC and NA-MCT leaders reported that, as the implementation moved forward, they developed a deeper knowledge of the characteristics that contributed to the success of a peer specialist. They were more successful at hiring peers who were a better fit and had the characteristics to succeed in the job. These characteristics included resourcefulness, sustained success with recovery, and an ability to balance professionalism with empathy and approachability.

NA-MCT staff clarified the role of the peers in the teams, and this helped address some of the problems. Working with peer specialists was a new experience for the clinicians, who were not always sure what the peers' roles were or how to relate to them. Several respondents identified a need to define and clarify the roles and responsibilities of the peers. This was for the benefit of both the clinicians, who would get a better sense of how to work with the peers, and the peers, who would understand better what their roles in the teams were and what supervisors expected of them. One clinician reported that the peers initially did not participate in the meetings with clients, and how they fit into the team's work was unclear. This clinician explained that the situation improved as the clinicians better understood the peers' role: "Once we had a better understanding of their role, in getting—giving—providing the peer [with] background knowledge about the case, that helps…inform what they say sometimes at the meetings."

2. Clinician experience

In addition to the initial challenges they had working with peers, clinicians found it difficult to adjust to the NATM approach to care, which was adopted without their input. They also struggled to balance using the new model with the Parachute NYC participants while practicing their traditional method of care with clients outside the program.

Conversion of the existing mobile teams to NA-MCT was an early challenge for clinicians. The dissonance between past practice, which focused on a traditional short-term crisis model, and the NA-MCT approach was a challenge for clinicians, who had to adopt the new model without being asked for their input. As one interviewee explained, "It is difficult because [the clinicians] were practicing in an existing mode of treatment and they had to change their whole mindset." Acceptance of the model improved with time, as those who were not comfortable with it left the organization and others came to appreciate it as they saw the benefits to clients.

Clinicians struggled to balance Parachute NYC with work on a traditional mobile team. For clinicians, the NATM model added new responsibilities—ongoing treatment and client support—to the mobile crisis teams. Although the NATM became the primary model used by the teams, the clinicians continued to provide the traditional services of short-term assessment and referral to clients not participating in Parachute NYC. Many clinicians on these teams struggled to balance the Parachute NYC training and implementation with their traditional work. As one of them said, "Sometimes it's kind of overwhelming and draining, and especially when you kind of have to be in two totally different mindsets." Another said, "I think it's asking too much of people to flip modes and it's difficult to build a new culture or practice around the Parachute ideas, which are different from traditional ideas … I would say definitely it was difficult and stressful for everyone involved in trying to do both."

3. Workforce training

Parachute NYC staff received different trainings to help them effectively deliver care in the new model. The clinicians and peers we interviewed offered different assessments of how useful the trainings were.

Peer staff found the trainings they received to be useful. The peer specialists were newly hired for the Parachute NYC program and trained in the models of care that were relevant to their roles within the program (IPS and Peer Health Navigation for all peers; NATM for the NA-MCT peers). Peer staff generally found the IPS trainings to be valuable in both years. In particular, they said the role-playing activities were valuable because they helped the peers apply the principles to their work and gave them the skills to effectively interact with participants. As one peer specialist said, "The trainings were pretty consistent. I never came back with something I couldn't use. They were very helpful." Staff also appreciated the regular refresher trainings in IPS and their monthly check-in meetings with the IPS trainer. As part of the IPS model, peer specialists were also trained in and practiced co-reflection, a technique that allowed peers to reflect on their practice and give each other informal feedback about how they were doing. In discussing the process, one peer said, "I don't feel that there's anything that goes on that I can't share with someone and process and work through while I'm here during my day at work."

Clinicians experienced challenges with the early training. Clinicians on the NA-MCTs were trained on the NATM model and IPS. The NATM model was a completely new approach for the clinicians, one that differed substantially from the manualized and structured approach they had traditionally used, and interviewees said they would have benefited from more opportunities to practice the approach. At the same time, some clinicians labeled the IPS training as anti-psychiatry because it was presented as an alternative to the traditional approach to mental health treatment, and the training included discussions of the drawbacks and limitations of the traditional approach. This increased their discomfort with the new NA-MCT model as a whole. Over time, the NA-MCT training materials and approaches were adapted to fit Parachute NYC and the program began providing train-the-trainer sessions, giving advanced training to staff members who had been practicing the models so they could train new staff members. In addition, the Parachute NYC staff were able to better bridge the respective IPS and NATM models by developing a crosswalk of the models to show where the principles converged, and this helped them integrate and coordinate positive use of the models within the project. Staff members believed these strategies, combined with the ongoing support the trainers provided through regular check-in meetings with staff, made later training rounds much more useful.

4. Job satisfaction

In the surveys and interviews, staff generally said they felt well supported and moderately satisfied with their jobs. They also said their work could be rewarding, believing they helped keep people out of the hospital. However, staff members also said their job was demanding emotionally, causing high levels of stress. Peer specialists reported higher levels of emotional and psychological strain than the clinicians did, and described this strain as one of the more difficult aspects of their work. As one CRC peer put it, "It gets hard to maintain an authentic emotional presence for 12 hours at a time, three times a week." In addition, their work occasionally left peer specialists open to difficult or uncomfortable reactions and emotions as they listened, responded to, and supported clients who were experiencing trauma. Staff noted that

their supervisors tried to keep them out of high-stress situations, and the peers were given time to process difficult encounters. Peers suggested that having more resources to support self-care and having more time to decompress after stressful incidents would help alleviate their job-related stress. As one peer specialist said, "We go from crisis to crisis. We don't have breathing room or have time to decompress because we have to go on to our next duty."

Each NA-MCT and CRC had weekly meetings; at the meetings, staff members could discuss current cases and solicit feedback and suggestions. These meetings were also an opportunity to revisit and review the techniques they were taught in training. Parachute NYC staff also said that regular debriefings were useful because they ensured open communication and dialogue within the CRC and NA-MCT teams.

Retention of the staff was at a relatively high level throughout the course of the program. However, several peer staff expressed concern about the limited opportunities for professional development and advancement in their roles. This affected their desire to remain in the program. Some peers expressed a desire to have supervisors who identified as peers, because their shared background would enhance communication. Program leaders acknowledged the need to identify clear paths for advancement for peers within the CRC and NA-MCT teams, and reported taking steps to address this issue. For example, several CRCs were considering a new position for a senior-level peer.

E. Program sustainability and spread

FPHNY continues to sustain all three program components with various funding sources.

- **CRCs.** The CRCs continue to provide services in all four boroughs. The program delivery model and nature of the services provided remain unchanged.
- **NA-MCTs.** NA-MCT services continue in three of the four boroughs. Brooklyn currently does not have NA-MCT coverage. The NA-MCT teams have been separated from the mobile crisis teams, and registered nursing staff were added to the teams.
- **Support line.** The support line services are now available 24/7 and are provided by text and Internet chat in addition to the phone service. The line is being operated by a different agency and is no longer staffed by peers from the Manhattan CRC.

The key sustainability approach for the mobile treatment team and CRC services was to make them reimbursable through Medicaid. Before statewide Medicaid reforms, individuals with a disability in New York State were not part of Medicaid managed care, instead receiving services through a fee-for-service arrangement. Beginning January 1, 2016, the state restructured its Medicaid system with the expectation that behavioral health services for the population with serious mental illness will be provided through managed care organizations. The mobile treatment and CRC services became reimbursable as a type of home and community-based service. These reform efforts were implemented through the state's 1915i plan. Program leaders reported that the eventual goal was to achieve coverage of approximately 75 percent of the CRC and mobile team costs from Medicaid revenue, and for the city or state to cover the rest of the operating costs. Program leaders anticipate that the funds provided by the city or state in the CRCs will be largely for uncompensated care, that is, to cover the cost of services for

participants who are not enrolled in Medicaid or are not eligible for the home and community-based services portion of Medicaid.

Originally, the state's Medicaid transformation to managed care was scheduled to take effect in January 2015. This start date would have allowed FPHNY six months overlap with the HCIA award period. Given the delay in Medicaid redesign reforms, FPHNY had to seek other sources of funding. The awardee was able to successfully secure short-term funds from the state and city to bridge the funding gap. Program leaders also noted that, although the new Medicaid billing option became available in January 2016, the ramp-up to billing Medicaid has been slow as beneficiaries were gradually enrolled in the managed care plans. Some of the CRCs have been able to successfully bill Medicaid for their services since this option became available, but this is not yet happening at the level that was hoped for. Program leaders noted that some time will likely be required before the CRC and mobile team services are fully sustainable as planned.

Starting in July 2016, the Parachute NYC support line services were incorporated into a larger behavioral health hotline and informational referral service procured by New York City.

F. Lessons learned

Overall, results from our evaluation suggest that Parachute NYC helped lower rates of Medicaid-funded hospitalizations and Medicaid expenditures for the CRC participant subgroup. In contrast, we found that Medicaid-funded hospitalizations and expenditures rose for the NA-MCT subgroup. Because these findings pertain to fewer than 20 percent of all program participants, they may not be representative of the participant group as a whole.

Qualitative data supported these results. Staff expected the CRCs to reduce hospitalizations by providing participants with a residential, community-based alternative to psychiatric hospitalization. Program leaders believed the CRCs would also lower health care expenditures by reducing the use of expensive services like hospitalization.

The majority of FPHNY participants were Medicaid beneficiaries and fewer than four percent of participants were enrolled in Medicare. Unlike Medicaid, Medicare covers the cost of inpatient mental health services for enrollees. It is therefore likely that the CRC program would have resulted in even greater cost savings if applied as an alternative to hospitalization for more Medicare beneficiaries. We also note that, because the costs of inpatient hospitalization stays are not included in Medicaid, reductions in Medicaid healthcare expenditures costs due to the CRCs are likely to be even greater than we found in our analyses.

The transformation of the workforce was an important aspect of Parachute NYC. Clinicians were trained on an entirely new model of care, and peers were incorporated into the existing mobile crisis teams. However, properly integrating the peers and establishing their role in the program was an ongoing challenge for the program. Conflicts between peer specialists and non-peer staff in the NA-MCTs were the biggest challenge to program implementation and may have impeded the effectiveness of the innovation. The experiences of Parachute NYC in establishing an integrated workforce are instructive. Programs based on an integrated peer and non-peer workforce will need to pay special attention to peer hiring and integration in order to be successful.

Our evaluation indicated that programs like Parachute NYC are more likely to be sustainable when they have the strong backing of state and local leaders and take place in a supportive policy environment. Parachute NYC benefited from the status of the New York City Department of Health and Mental Hygiene as a strong player in the legislative landscape that could effectively advocate for policy changes. In addition, New York's Medicaid reform offered a unique opportunity to build a sustainable payment model into state legislation.

VI. HEALTHLINKNOW

Findings from Mathematica's Evaluation of HealthLinkNow's (HLN's) HCIA Program

- Our analysis of quantitative data for HLN's Medicare population indicated the program had no statistically significant impact on total expenditures, hospitalizations, or emergency department visits; however, the result may reflect major analytic limitations.
- HLN faced persistent challenges meeting its enrollment targets, particularly for participants in rural and underserved areas. HLN responded by expanding the program to an additional state and dedicating significant resources to recruiting and engaging providers.
- Implementation was repeatedly delayed by the complex and lengthy provider credentialing process at many partner sites. Creation of a national centralized credentialing agency could streamline this process.
- HLN's implementation was made more difficult early on by many providers' and payers' unfamiliarity with the practice of telemedicine.
- Care navigators, who had both case management and administrative responsibilities, were an integral part of the HLN workforce.

A. Introduction

HealthLinkNow, Inc. (HLN), an organization of behavioral health providers delivering remote health care services, used HCIA funding to integrate telemedicine with a patient-centered medical home model. The goal of the program was to use integrated telemedicine and health information technology (IT) to virtually link patients with behavioral health specialists— psychiatrists, therapists, and counselors— within primary care provider (PCP) offices. HLN provided three types of services:

- Telepsychiatry, including included online psychiatric assessments, treatment planning, medication management, counseling and supportive therapy, and cognitive behavioral therapy
- Case management and care coordination by care navigators
- Provider- and patient-centered technology support from a health IT system featuring (1) an integrated electronic medical record used by behavioral health specialists and (2) a patient portal with secure, web-based patient-provider communication tools and online health education features

HLN's target population included residents of Washington, Montana, and Wyoming with any or no insurance coverage. The program served children and adults with mental health disorders ranging from depressive and anxiety disorders to dementia and substance use disorders. Care navigators were the initial point of contact for participants; they had both case management and administrative responsibilities. Care navigators coordinated with HLN's behavioral health specialists and PCP staff. They also helped participants overcome barriers to care; for example, by helping them obtain prescribed medications and sending them appointment reminders. In

addition, the care navigators were responsible for performing initial participant intake and scheduling appointments.

We draw on the following data sources for this chapter:

- Enrollment data submitted by HLN to the reporting website maintained by CMMI's technical assistance contractor (the Lewin Group) for the HCIA Round 1 initiative.
- Workforce survey data collected in spring 2014 and 2015 to gather information about staff burnout and stress, job satisfaction, and perceptions of training and support.
- Quantitative data on participant enrollment, service utilization, and expenditures, extracted from Medicare claims and enrollment information using participant identifiers provided by HLN.
- Qualitative data, including phone interviews and in-person site visits conducted in spring 2014 and 2015. Mathematica conducted in-depth interviews with awardee leaders, members of the workforce, and other stakeholders. We also convened focus groups with site coordinators, primary care providers, and program participants in 2015.

Using Medicare claims and enrollment data, we applied a difference-in-differences regression analysis to estimate the impact of HLN's telepsychiatry program on service utilization and expenditures among Medicare beneficiaries. Although we were able to construct a strong comparison group with data from neighboring states, we did not observe significant changes in the measured outcomes for Medicare beneficiaries served by HLN's programs during the study period. Our trend analysis also did not yield significant increases or decreases in cost or utilization. Constraints related to program implementation and data availability limited our analyses. This report provides final evaluation results for HLN.

1. Overview of administrative context

HLN's staffing structure consisted of a core team of care navigators and behavioral health specialists (psychiatrists, therapists, and counselors), supported by administrative staff including IT professionals, state-based project managers, program administrators, and the chief medical officer. HLN's behavioral health specialists worked remotely from their homes, and care navigators and administrative staff (except for the state-based project managers) worked in HLN's Sacramento office.

HLN's model focused on providing both telepsychiatry and case management at participating PCP offices. The awardee partnered with local providers in Washington, Montana, and Wyoming who referred patients with mental health treatment needs to HLN. HLN behavioral health specialists conducted video telepsychiatry sessions with participants in PCPs' offices. Participating provider sites were predominantly primary care offices, but also included skilled nursing facilities, hospitals, and county public health offices. Each participating site assigned a site coordinator to work with the HLN care navigators to schedule appointments and liaise with participants and providers.

HLN initially targeted provider sites in Montana and Wyoming—states with large rural areas that had shortages of psychiatrists and other behavioral health providers and faced above-

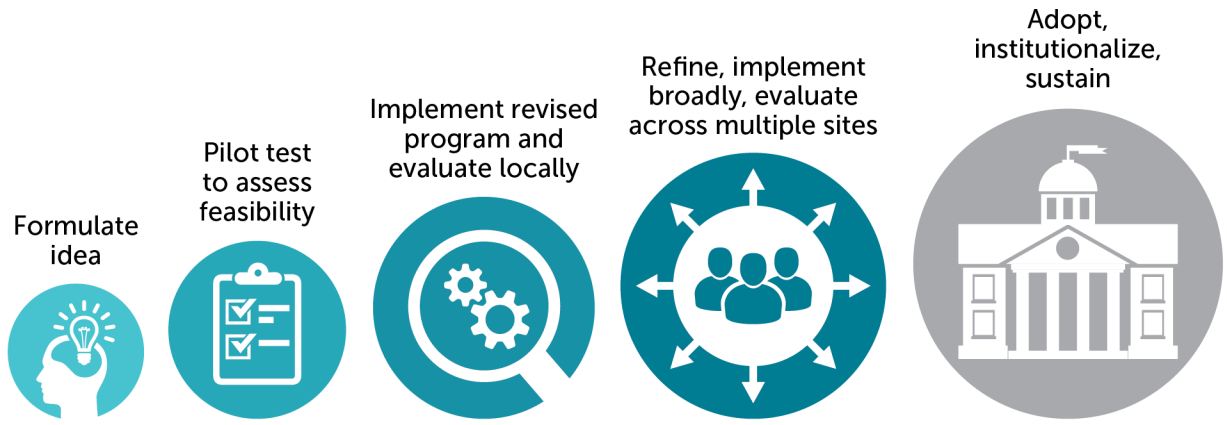
average rates of mental disorder, suicide, and substance abuse.[25,26,27] Inadequate access to mental health services may result in increased use of the emergency department (ED), unnecessary hospitalizations, and higher health care costs.

Over the course of the project, HLN's recruited sites did not refer as many participants as expected to the program. To meet enrollment targets, HLN received approval from CMS to expand the program into Washington State. HLN also increased the reach of the program by recruiting large health networks that could provide a greater volume of referrals.

2. Progression through phases of innovation

HLN's HCIA funding enabled the organization to progress through the first four phases of innovation in Figure VI.1. Using the HCIA, HLN created its program and tested the viability of the telepsychiatry service model in PCPs' offices for the first time. Over the course of the award period, HLN expanded its program to different populations and settings. HLN continues to work on establishing a sustainable business model for its services.

Figure VI.1. Phases of program innovation: HLN



3. Enrollment

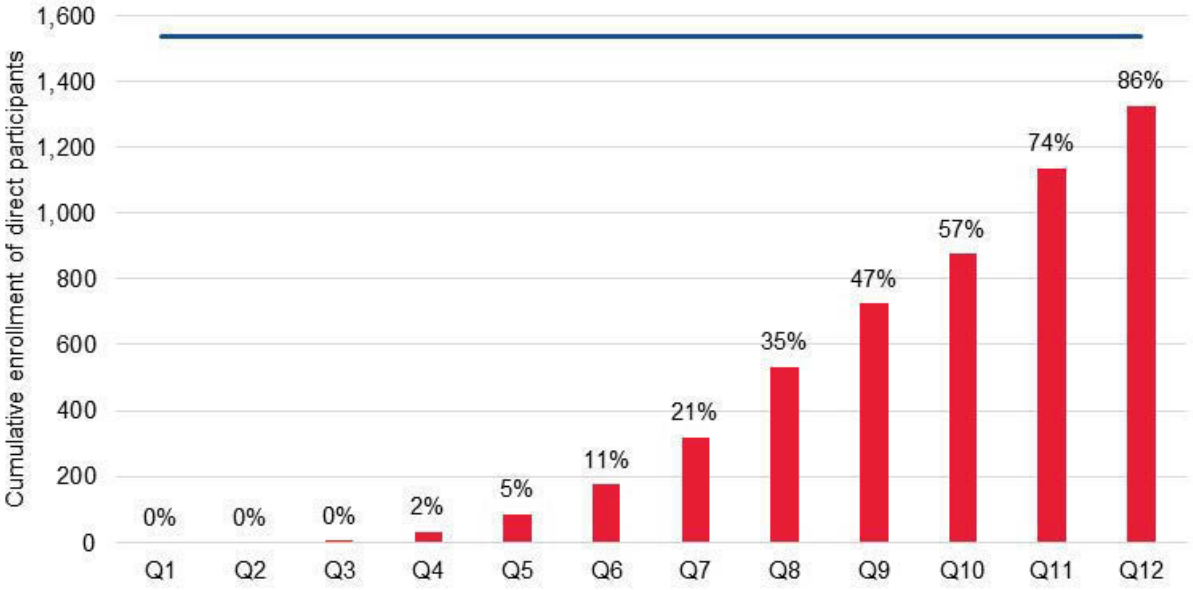
By the end of the 12th quarter (June 30, 2015), the HLN program had provided services to 1,326 participants—86 percent of its enrollment target (Figure VI.2).

[25] Wang, P.S., M. Lane, M. Olfson, H.A. Pincus, K.B. Wells, and R.C. Kessler. "Twelve-Month Use of Mental Health Services in the United States: Results from the National Comorbidity Study Replication." *Archives of General Psychiatry*, vol. 62, 2005, pp. 629–640.

[26] Substance Abuse and Mental Health Services Administration, Center for Behavioral Health Statistics and Quality. "The NSDUH Report: State Estimates of Adult Mental Illness from the 2011 and 2012 National Surveys on Drug Use and Health." Rockville, MD: Substance Abuse and Mental Health Services Administration, February 28, 2014.

[27] Centers for Disease Control and Prevention. QuickStats: Age-Adjusted Suicide Rates, by State—United States, 2012. Atlanta, GA: Centers for Disease Control and Prevention, November 14, 2014.

Figure VI.2. Percent of target enrollment achieved by quarter, Q1–Q12



Source: Mathematica analysis of program enrollment data provided by HLN.

Note: The blue horizontal line represents HLN's enrollment target of 1,534 unique participants.

4. Participants' demographic characteristics

Most HLN participants were between the ages of 18 and 64 (69 percent), and the majority (64 percent) were female (Table VI.1). Slightly more HLN participants were enrolled in Medicaid (29 percent) than in other types of insurance. Information on the rural or urban residential status of participants was not available.

Table VI.1. Demographic characteristics of HLN participants, by state

			State					
			Montana		Washington		Wyoming	
	Number of participants	Percent of participants	Number of participants	Percent of participants	Number of participants	Percent of participants	Number of participants	Percent of participants
Total	1,326	100	512	100	563	100	251	100
Age								
Under18	243	18.3	123	24.0	77	13.7	43	17.1
18–34	316	23.8	104	20.3	160	28.4	52	20.7
35–44	207	15.6	65	12.7	93	16.5	49	19.5
45–54	221	16.7	65	12.7	107	19.0	49	19.5
55–64	165	12.4	54	10.5	76	13.5	35	13.9
65 or older	174	13.1	101	19.7	50	8.9	23	9.2
Gender								
Female	843	63.6	341	66.6	350	62.2	152	60.6
Male	483	36.4	171	33.4	213	37.8	99	39.4
Insurance coverage								
Medicaid-only (non-dual)	382	28.8	188	36.7	132	23.4	62	24.7
Medicare, dual and Medicare-only (non-dual)[a]	328	24.7	172	33.6	92	16.3	64	25.5
Private insurance	310	23.4	66	12.9	205	36.4	39	15.5
Other[b]	306	23.1	86	16.8	134	23.8	86	34.3

Source: Mathematica analysis of program enrollment data provided to Mathematica by HLN staff.

[a] Dual Medicare-Medicaid beneficiaries and non-dual beneficiaries are reported here.

[b] Includes other insurance, unknown, and uninsured.

B. Methods

The goal of this evaluation is to determine whether HLN's telepsychiatry program had a measurable impact on participants' health care service use and to describe how HLN's key components—care navigation, telepsychiatry, and technology support—were implemented. Here, we describe the methods we undertook to accomplish these goals using qualitative data collected directly from HLN and quantitative data from administrative and enrollment records. We begin by summarizing the steps taken to select the HLN participants included in our quantitative analysis and to identify a matched comparison group. (Appendix A provides further details.) We then discuss our approach to examining the potential impact of the program on three core outcome measures. Finally, we describe the qualitative data collection methods we used to gather information on the implementation of HLN's program.

1. Quantitative methods

Due to limitations in the data available to support our analysis, we focused only on HLN's dual Medicare-Medicaid and non-dual (Medicare only) fee-for-service (FFS)-enrolled patient population.[28] Although we received Medicaid identifiers, we were unable to use claims information for these beneficiaries because of substantial lags in data availability. In addition, Washington State has a high penetration of managed care for behavioral health services, and this may have further limited data available for this analysis.

We obtained Medicare administrative data from the CMS Virtual Research Data Center. Of the 328 Medicare beneficiaries HLN reported serving through the HCIA, we found only 266 with available identifiers in the Medicare data. We then excluded beneficiaries who:

- Did not reside in the intervention states of Washington, Montana, or Wyoming, or in three border counties, at the time of program enrollment (18 beneficiaries)
- Were enrolled in Medicare Advantage (9 beneficiaries)
- Did not have Part A and B coverage or Medicare as their primary payer (7 beneficiaries)

The resulting analysis sample included 232 participants.

For each participant, we identified a primary care visit, acute care visit, or nursing home stay in the month before program enrollment that was likely associated with referral to the intervention program. We call these services "trigger" services. For each participant, we also identified chronic physical and mental health conditions that existed in the 12 months before enrollment—Medicare claims included diagnoses of mental health or specified somatic conditions[29] for all participants.

We then took the following steps to identify a well-matched comparison group:

[28] Due to lack of Medicaid identifiers for enrollees in Washington and significant time lags in the availability of Medicaid data for Montana and Wyoming, we did not have a large enough sample of Medicaid-covered participants to analyze. In order to conduct analyses for this report, we required data by February 2016. At that time, Medicaid data for Montana and Wyoming were available only through 2013, and only 80 Medicaid-enrolled individuals in those states participated in HLN's program through 2013. This sample size meant we could not perform meaningful impact analyses.

[29] Somatic conditions included chronic pain, fatigue, insomnia, hallucinations, or memory loss.

- We identified 10 neighboring states with characteristics similar to those of the HLN service areas. These states were North Dakota, South Dakota, Idaho, Oregon, Nevada, Utah, Colorado, Nebraska, Arizona, and New Mexico.
- We identified Medicare enrollees in these neighboring states with: (1) a primary care visit, acute care visit, or nursing home stay between March 2013 and March 2015, and (2) a behavioral health or somatic condition within the 12 months before the primary care visit, acute care visit, or nursing home stay. These beneficiaries formed the pool from which we constructed a comparison group.
- Using this pool, we matched up to 20 individual comparison group members to each intervention group member. Using a matching algorithm, we identified comparison group members who resembled the participants on a combination of characteristics predictive of future Medicare service use and expenditures: demographics, disability status, selected Hierarchical Condition Categories, dual Medicare/Medicaid enrollee status, geographic characteristics, pre-period history availability, program enrollment date, service use and expenditures in the year before enrollment, and common mental health and somatic diagnoses.

We calculated descriptive statistics and conducted interrupted time series analyses within a difference-in-differences framework to estimate program impacts on key outcomes. The evaluation period, including pre- and post-enrollment periods for each participant, was March 1, 2010, through June 30, 2015. Detailed information on the matching and impact analysis methodology are available in the Appendix A.

We present descriptive statistics on the analytic sample in Table VI.2. The sample includes 17.5 percent of the total HLN population of 1,326 participants, and the sample and the total population differ in several important ways. Compared to the total population, participants included in the analytic sample were markedly older (54.7 percent aged 65 or older in the analytic sample vs. 13.1 percent in the total sample), and more likely to be female (75.0 percent vs. 63.6 percent). These differences have important implications for interpreting our results. In particular, our results are not generalizable to all of HLN's participants and should not be viewed as a comprehensive test of the intervention's effects.

Table VI.2. Demographic characteristics of Medicare analytic sample

	Number of participants	Percent of participants
Total	232	100.0
Age[a]		
Under 35	31	13.4
35–44	35	15.1
45–54	39	16.8
55 or older	136	58.6
Gender		
Female	174	75.0
Male	58	25.0
Insurance coverage		
Medicare-only	120	49.1
Dual Medicare-Medicaid	112	50.0
Disability status[b]		
Disabled	118	49.1
Not disabled	114	50.9
Urbanicity		
Rural	82	35.3
Suburban	54	23.3
Urban	96	41.4

Source: Mathematica analysis of FFS Medicare claims data for the period from March 2013–June 2015.

[a] Fewer than 11 enrollees were in the under 18 or 65 or older categories. We therefore combined age categories to protect individual privacy.

[b] Disability status indicates that a disabling impairment was the original reason for Medicare entitlement.

2. Qualitative methods

We conducted site visits to HLN's administrative offices in spring 2014 and spring 2015. During the 2014 site visit, we conducted in-depth interviews with care navigators, behavioral health specialists, administrative staff, and other stakeholders, such as website contractors. In total, we conducted 10 interviews with 18 individuals. In 2015, we conducted a second round of in-depth interviews with the same types of team members but not necessarily the same people, because of staff turnover. This included 9 interviews with 15 individuals. In addition, we convened three focus groups during the 2015 visit. One group consisted of eight adults who either received services from HLN or were caregivers of children who received services from HLN. Another included three PCPs at participating HLN sites. The third included 10 site coordinators from the participating sites.

In the 2014 interviews, we discussed implementation progress and challenges, staffing, target population, program expenditures, workforce development, and program leadership. In 2015, we added questions about the perceived effects of the program and HLN's sustainability plans.

C. Summative findings

1. Overview of results

Our analyses found no statistically significant impacts of HLN's HCIA on total expenditures, hospitalizations, or ED visits among participants with Medicare insurance. Impacts

among Medicare enrollees living in rural areas were of particular interest for this awardee, but we were not able to conduct such analyses due to small sample sizes.[30]

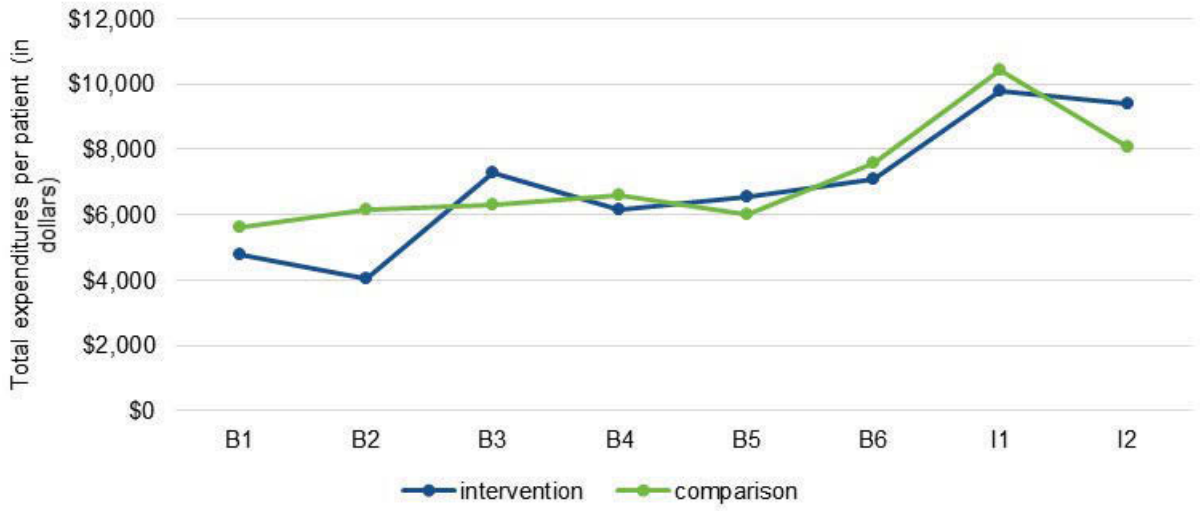
Below, we present first the outcomes over time (Section C.2). We conducted statistical significance tests to determine differences in mean outcome at each time point between intervention and comparison groups at $\alpha = .10$. In Section C.3, we present the results of the impact analyses. During our iterative modeling process for each outcome, we noted that resulting estimates were sensitive to model specification (that is, results differ depending on the variables included in the model). After each iteration, we examined model fit and performance and retained only those variables with the best predictive value. From the candidate models, we selected the one that provided the best combination of fit and parsimony to present in this report.

2. Description of trends in outcomes

Figures VI.3 through VI.5 display the average per patient outcomes (expenditures, hospitalizations, and ED visits, respectively), by group, for each six-month period from three years before to one year after the HCIA program period began. In general, outcome patterns over time were similar for intervention and comparison groups.

- Figure VI.3 shows that, for both groups, total expenditures increased over time.

Figure VI.3. Total expenditures per patient per 6-month period



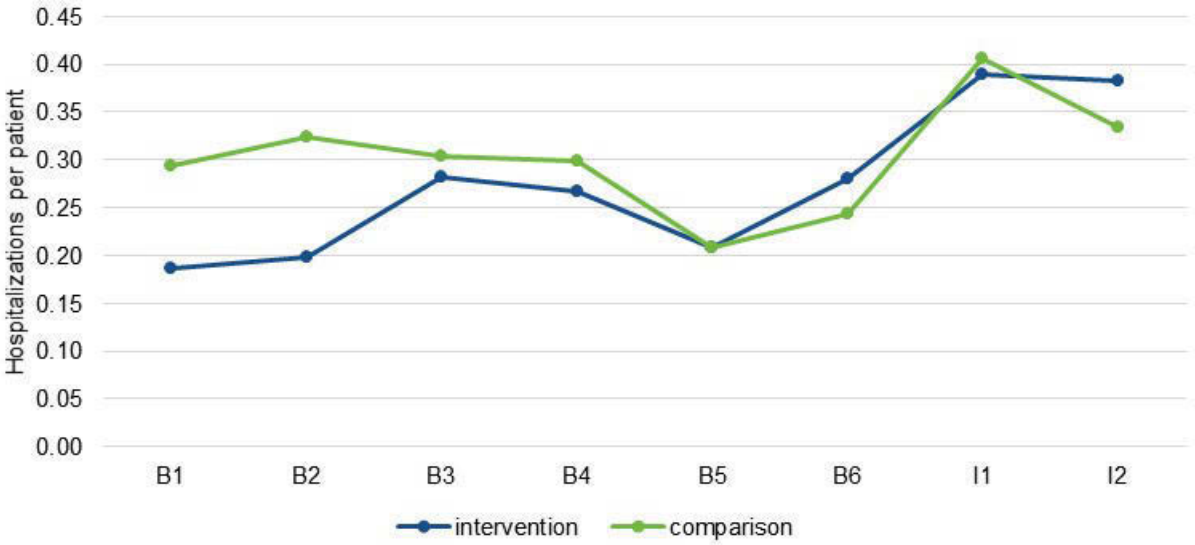
Source: Mathematica analysis of FFS Medicare claims data for the period from March 2010–June 2015.

Note: Each marker represents a period of 6 months. The baseline period (prefix B) encompassed the three years before the intervention began, and the intervention period (prefix I) encompassed the first year after the intervention began.

- Figure VI.4 reveals that hospitalization rates were a bit higher for the comparison group during the early baseline period, but were similar for the two groups beginning a year before HCIA implementation through the first year of the intervention period.

[30] Too few rural participants were enrolled in Medicare at each 6-month time point of the evaluation period to support the analyses.

Figure VI.4. Hospitalizations per patient per 6-month period

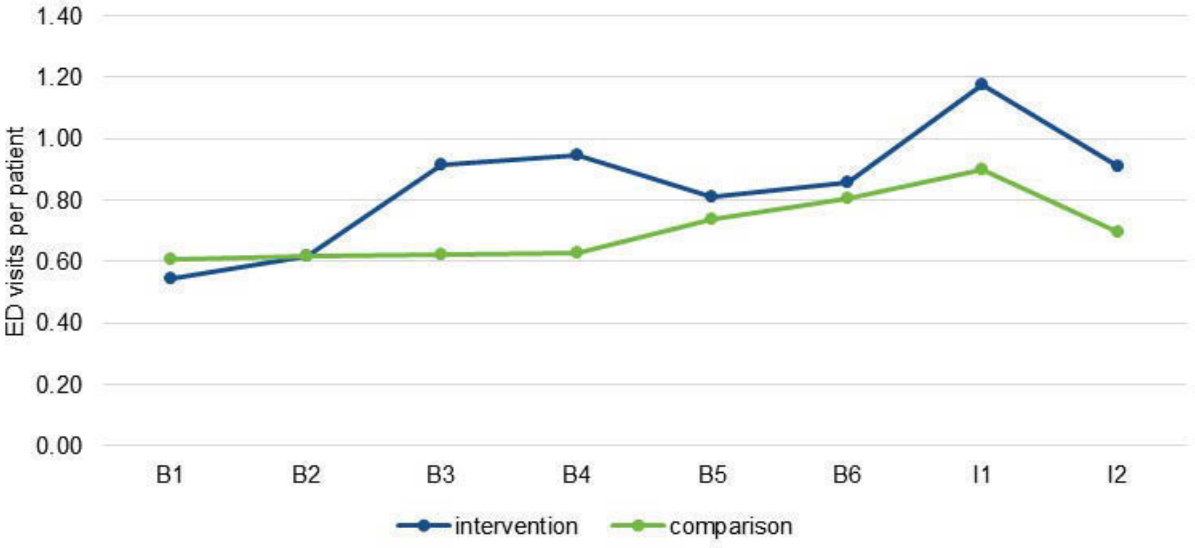


Source: Mathematica analysis of FFS Medicare claims data for the period from March 2010–June 2015

Note: Each marker represents a period of 6 months. The baseline period (prefix B) encompassed the three years before the intervention began, and the intervention period (prefix I) encompassed the first year after the intervention began.

- Figure VI.5 shows that the ED visit rate for the intervention group was generally higher than the comparison group's, but varied over time. The relation between the two groups did not change noticeably after HCIA began.

Figure VI.5. ED visits per patient per 6-month period



Source: Mathematica analysis of FFS Medicare claims data for the period from March 2010–June 2015.

Note: Each marker represents a period of 6 months. The baseline period (prefix B) encompassed the three years before the intervention began, and the intervention period (prefix I) encompassed the first year after the intervention began.

3. Results of impact analyses for CMMI's core measures

Findings from our impact analysis of the core measures are as follows:

- HLN's HCIA program was not associated with significant changes in ED visits (p = 0.72), total expenditures (p = 0.58), or number of hospitalizations (p = 0.84).
- The intervention group had only 35 readmissions over four years, far below the threshold of analysis (11 per quarter) used in previous CMS reports. Because there were so few readmissions, analysis of change over time would not be meaningful, so we did not conduct an impact analysis of readmission rates.

Table VI.3. Medicare impacts and total savings attributable to the HLN program

	All intervention group members	
	Change	90% confidence interval
Aggregate results		
Total Medicare expenditures (in thousands of dollars)	$260	[-512 to 1,031]
Hospitalizations	-101	[-949 to 747]
ED visits	20	[-71 to 112]
Per beneficiary month		
Expenditures	$128	[-252 to 508]
Hospitalizations	-0.05	[-0.47 to 0.37]
ED visits	0.01	[-0.04 to 0.05]
Number of participants	232	
Mean number of intervention months per participant	8.8	
Approximate proportion of intervention population represented in analysis[a]	17.5%	
Intervention period	March 2013 to June 2015	

Source: Mathematica analysis of FFS Medicare claims data, March 2013–June 2015.

Note: We derived impact estimates from regression models that controlled for age (linear and squared), gender, dual Medicare-Medicaid enrollment status, whether 12 months of baseline data were available, behavioral health diagnoses, disability status, service use in the first month the participant began the program, urbanicity, county poverty level, county unemployment level, HCC score, and indicators of chronic conditions not accounted for in HCC score. We derived the impact estimates in Stata using the margins command to compare the difference between the intervention and baseline period means for the intervention and comparison groups and accounting for the nonlinear modeling approach. The confidence intervals for total expenditures, hospitalizations, and ED visits were derived based on bootstrap methods and were adjusted for multiple testing based on the generalized Tukey method. A confidence interval that includes zero means that the observed changes are not statistically significantly different from zero.

[a] We calculated the approximate proportion of the intervention population represented in the analysis by dividing the number of participants (232) in the analysis by the number of individuals who participated in HLN program between March 2013 and June 2015 (1,326).

4. Analytic limitations

We did not find statistically significant associations between the HLN telepsychiatry program and health care utilization and cost outcomes. This may be because of the challenges we encountered during our modeling efforts. Limitations of the impact analyses include the following:

- The analysis population was limited to Medicare FFS enrollees, who represent only 17.5 percent of all participants served by the HLN program. Our findings may not generalize to the total HLN population or to differently insured participants covered by other insurance (for example, Medicare Advantage, Medicaid only, CHIP, or commercial insurance). Compared to the total participant population, the analytic sample was also older and had a higher proportion of females. We did not have access to urbanicity and disability status for the total HLN population,[31] so were unable to determine how the analytic sample compared to the larger HLN population on these variables.
- Because we included only Medicare-covered services provided at HLN sites, the analyses may underestimate service use and costs because they exclude services that were not covered by Medicare or those received from providers who were not part of the program.
- The model results were sensitive to model specification. This may be due to the relatively small number of Medicare FFS enrollees who were enrolled in the HLN program and included in these analyses (n = 232). Model results may fluctuate according to how variables are defined (for example, linear or binary) and according to the number and intercorrelation of variables entered into the model. For example, when "state" was not included as a covariate, impacts on core outcomes still were not statistically significant. (See the Appendix A for additional details on model specification.)
- Because outliers may have an undue influence on results, we removed them from the analysis.

To ensure the final model was not overfit, we subjected it to a test of model overspecification, and the results supported the assumption that each predictor variable contributed unique information. Nevertheless, including different covariates in the model yielded different results, so results should be interpreted cautiously.

5. Qualitative findings on the perceived effects of the program

Although quantitative analyses did not reveal statistically significant associations between the HLN program and core measure outcomes for Medicare participants, the respondents we interviewed on site visits described HLN's effect on its participants' health outcomes, quality of life, and access to care.

Access to mental health services. One of HLN's goals was to provide access to mental health services to patients whose communities lacked these services. Although we were not able to conduct quantitative analyses of HLN impacts among rural participants specifically, interview respondents told us about the ways the program helped rural participants by giving them access to mental health services. As one PCP staff member explained, "Because we are in a very rural area, no psychiatric services [are available] within 130–140 miles, and then the wait list was six months out." Another staff member added, "Oftentimes, 10 people can get scheduled within two weeks [through HLN], which is much preferred to the three to six months that it can take to find other psychiatry resources in the area." Several respondents pointed out that many rural

[31] Disability status and counties of residence for the Medicare population were available from Medicare enrollment files.

participants would have had to go without mental health services if they had not been able to enroll in HLN.

Mental health. Most respondents reported that participants thought highly of HLN's services, which enhanced participants' mental well-being and quality of life. PCP site coordinators told us about progress they observed in participants' mental health, including fewer symptoms of anxiety and depression and increased self-esteem and self-confidence. Participants themselves also had a greater sense of well-being. For example, one individual in the participant focus group described feeling "like a totally different person" since beginning therapy. Another stated, "I'm feeling happier all around and just doing a lot better."

Physical health. PCP site coordinators and HLN care navigators reported that HLN led to improved physical health outcomes for some participants. These respondents shared examples of improvements in participants' health behaviors after they started therapy, including better exercise habits and healthy eating. Program participants in our focus group also reported making progress towards physical health goals such as losing or gaining weight. In addition, respondents reported that participants benefited from the medication management and psychopharmacology services offered by HLN psychiatrists. One PCP staff member said, "In some cases, people are actually needing [fewer] medications of other kinds and having [fewer] somatic complaints, because they were making some headway regarding their mental health." Another respondent said participants were experiencing fewer side effects from their medications as a result of HLN's help managing their psychiatric medication regimen.

Satisfaction with care. Many participants and PCP staff were pleased with the HLN model of care. PCP staff and HLN behavioral health specialists told us they were initially skeptical about the effectiveness of delivering psychiatric services by videoconference, but they said participants responded positively to HLN's format and were comfortable using the technology to interact with HLN providers. Most participants and PCP staff also praised the high quality of the HLN specialists, saying the participants easily developed rapport and trust with them.

Accessibility. Participants appreciated the convenience of having a behavioral health appointment in their PCP's office. Some participants who lived in metropolitan areas and had access to mental health services outside HLN expressed a preference for HLN for this reason. One respondent was happy that her appointments were always on time, saying, "I can trust that when I get there, the appointment's going to start and I don't have to wait around for 15 or 20 minutes while they finish up another patient." In addition, several PCP staff said no-show rates were lower and adherence to treatment was greater in the HLN program than for traditional mental health services. They believed this was because the HLN appointments took place at PCP offices. One PCP said, "It seems that there's more follow-through with the patients when they come to a familiar area."

D. Findings about the workforce

Site visit interviews and staff focus groups included questions on workforce development, deployment, and training. We also conducted workforce surveys in spring 2014 and 2015 to gather information about staff burnout and stress, job satisfaction, and perceptions of training and support.

The care navigator role was seen as essential to the success of the program. HLN administrators and behavioral health specialists described care navigators as an integral part of the care team. These individuals liaised with the participants, PCP staff, and HLN specialists. They were responsible for ensuring that participants had timely and convenient access to services and were receiving their medications, attending follow-up visits, and accessing HLN's online health educational materials as needed. In addition to serving as the initial contact point for the participants, the care navigators also performed administrative tasks such as scheduling, billing, and obtaining PCP records. All care navigators were trained in HLN's IT platform and used it for the scheduling, billing, and provider support aspects of their roles. One behavioral health specialist said, "I think the care navigators have been key. They end up with a lot more contact with the primary care providers and the patients, and they are very skilled in terms of customer service and the technical part. The care navigators are both technicians—they facilitate the technology … [and handle the] customer service [aspect of] talking to the nurses, doctors, and the patient's family members. They are really the hub of it all." Feedback from some respondents indicated that the care navigators might also have enhanced participants' well-being by reinforcing positive health behavior and encouraging adherence to treatment regimens.

Staff turnover was initially high, but improved over the course of the program. In 2014, the staff we interviewed reported moderately high rates of turnover at all staff levels. A significant cause of the turnover in behavioral health specialists appeared to be caused by HLN staff leaving to take higher-paying positions with other organizations. In addition, delays in the provider credentialing process delayed the start dates of behavioral health specialists, often for long periods, and sometimes resulted in them leaving to pursue other opportunities. In 2015, as the program continued to mature, program leaders reported lower rates of turnover. One program leader explained,

> We had some turnover within HealthLink … at different levels. I think a lot of that is a function of being a start-up. At the beginning I don't know that we were 100 percent clear what the ideal care navigator would look like. People were hired in that position, but not all worked out. Through that kind of experience, you look at the ones who are doing well and say, "This is what they have in common." We've experienced that both in leadership and operational positions. I think we are getting more sophisticated about how we recruit the staff and how we train.

PCP staff and providers reported that the program improved care coordination and communication between providers. Providers said the HLN program promoted better communication and care coordination between PCPs and mental health providers and helped them more holistically address patient health care needs. During site visits, respondents reported that HLN fostered communication between the various stakeholders involved in a participant's care by obtaining consent from participants to share behavioral health consult notes (including the participant's medical history, diagnosis, and recommended treatment) with PCPs after each visit. This helped the program deliver better care coordination, because the PCP remained aware of the participant's status and the suggested course of treatment. HLN also encouraged PCP staff to reach out to the specialist or care navigator with any questions. PCPs found this level of communication helpful as they worked to address the mental health needs of their patients. According to one PCP, "Trying to get specialty knowledge from psychiatrists who are not with

the HealthLinkNow is very difficult … so communication on medications and doses and those changes [has] been very helpful."

The sharing of information between providers was not bi-directional. Although PCP staff and providers had easy access to HLN behavioral health consult notes for their patients, the HLN behavioral health specialists did not have this same access to participants' medical records. The electronic medical record systems of HLN and the PCPs were separate. Therefore, if a behavioral health specialist thought certain information might help them in their work with a participant, they had to specifically request the information from the PCP. Some behavioral health specialists said this lack of a shared medical record was a limitation and potentially a barrier to full collaboration between providers.

E. Program sustainability and spread

At the time of our site visit in May 2015, HLN leaders did not yet have a clear plan for sustaining the program with rural and frontier communities. In spring 2016, HLN leaders declined our invitation to conduct a follow-up discussion about sustainability plans for the program. The final quarterly report submitted by HLN indicated that they began the process of closing out telepsychiatry services in 82 of their rural and remote primary care clinics, offering clinics the option to contract with HLN for continued services. HLN also continued to provide services on a contract basis at other sites. Program leaders we interviewed reported that they planned to sustain HLN's business model of providing telepsychiatry services on a referral basis and were pursuing opportunities to expand the program into new settings where the demand for psychiatric evaluations outpaced supply. This includes contracting with entities like accountable care organizations and providing remote psychiatric assessments in pain clinics and correctional facilities. HLN also planned to continue the practice of directly recruiting patients through its website and other forums.

During the site visits, we observed several issues that have implications for the sustainability of programs like HLN's:

Securing patient referrals. HLN's business model depended on partnering with PCPs to obtain participant referrals and to offer remote services to participants in PCP offices. Initially, HLN's enrollment strategy involved approaching clinics and hospitals in rural areas on an individual basis, but recruitment of PCP and hospital sites proceeded slowly and yielded fewer referrals than expected. Although HLN's initial approach was appropriate for rural communities in which face-to-face contact was key to building business relationships, this approach was time-consuming and not consistently successful. In response, HLN found it necessary to expand the program's reach to Washington State. The awardee also worked to recruit larger health systems and to improve their marketing materials. The sustainability of programs like HLN will depend on keeping sites engaged and securing the participation of entities with large patient populations, not just smaller clinics and providers.

Keeping PCPs engaged. Effective collaboration with PCP sites was key to HLN's program implementation, and it is a necessary component of a sustainable program. In some cases, lack of engagement and buy-in at sites hampered HLN's ability to recruit patients and to provide telepsychiatry services. HLN staff reported that the support of both clinical and administrative leaders was critical. In particular, they emphasized the importance of having a "champion" at the

primary care clinic or hospital who was willing to guide the project, talk to providers about making referrals, and help address any challenges that arose. One administrative staff member said, "There's some cases where it's three to six months before we finally get a facility launched and then even once it's launched, a lot of times as we go back and check in, we find that the doctors don't even know about the service … there's a communication gap at the facility. So I think when we really find that champion, that's where we see the most success." Some HLN staff members thought there was a cultural component to some of the resistance. As one behavioral health specialist said, "Providers don't refer to psychotherapy that much—they don't think of it as an option. There is emphasis on medications ... It's obvious that there is a need for psychotherapy services, but it's devalued."

Credentialing providers. The lack of standardized regulations and guidelines for hospital and clinic credentialing significantly hampered project implementation and has implications for the sustainability and spread of HLN's approach. In order to provide telepsychiatry services at a given facility, each HLN behavioral specialist had to be credentialed. The purpose of credentialing is to verify that providers are adequately qualified and licensed to deliver patient care. This credentialing process is a requirement of health facilities and systems nationwide and is mandated by various health care oversight entities including CMS, the Joint Commission, and the National Committee on Quality Assurance. Most of the health care organizations HLN worked with credentialed providers themselves. This process varied from one facility to the next, but often included validating the provider's medical licenses, education, and insurance coverage. Providers were also required to submit a significant amount of paperwork. HLN attempted to streamline the process of credentialing by developing guidelines and protocols for staff and training them on this subject. HLN leaders also encouraged staff to be persistent in following up with organizations and maintaining clear documentation of each step in the credentialing process. Despite these efforts, obtaining credentials for HLN specialists remained a long, expensive, and complicated process. Credentialing challenges resulted in months-long implementation delays in many facilities and higher expenses for HLN. The awardee's challenges with credentialing had a negative impact on timely and cost-effective implementation in new facilities and could also affect the feasibility of spreading the model.

Enhancing provider and payer awareness of telemedicine. Another challenge HLN encountered was a general lack of public knowledge about telemedicine. Staff members found it necessary to introduce the model and educate providers and payers about telepsychiatry without an existing program or model for reference. For instance, care navigators reported that they sometimes found themselves educating insurance companies about why it was important to cover telepsychiatry for their beneficiaries. While respondents noted that insurance companies were becoming more aware of and knowledgeable about telepsychiatry, some of the companies still did not reimburse HLN for these services in all cases.

Considering the policy and legislative context. Developments in state and federal policies on telemedicine can significantly affect the continued success of a program like HLN. Over the course of the program, the three states in which HLN implemented telepsychiatry passed legislation in support of telemedicine. For instance, parity regulations in Montana required commercial payers to cover telemedicine and to reimburse these services at a level equal to that of reimbursement for in-person services. However, respondents noted that even though some state policies were supportive of telepsychiatry and favorable to the growth of the HLN model, gray areas and exceptions often kept HLN from being reimbursed for its services. For example,

one large insurance provider initially refused to reimburse HLN services in Montana unless the provider was physically based in the state.

F. Lessons learned

Although we found no statistically significant impacts on total expenditures, hospitalizations, or ED visits, our analyses were limited to Medicare FFS enrollees in HLN sites and may not be representative of all participants served by HLN's program. The HLN program did reveal some possible strategies for successfully harnessing telemedicine to improve patient access to behavioral health care services.

Programs that provide telemedicine on a referral basis must devote significant resources to recruiting and engaging providers, especially to reach rural and underserved populations. Site recruitment was a time and labor-intensive process, and challenges with enlisting provider sites to participate resulted in lower than expected referral rates. HLN leaders expended significant effort developing relationships with potential sites, including having state-based project managers visit individual facilities to introduce the project. This approach was particularly problematic in Montana and Wyoming due to the size and topography of these states. Furthermore, some sites preferred to take a wait-and-see approach before they were willing to commit resources, requiring the HLN team to make multiple recruiting visits.

HLN's experience has implications for the supports necessary for programs of this kind to maintain sustainable patient volume. HLN staff noted that dedicating more resources to business development and marketing would have been helpful. Furthermore, they emphasized the need for sustained engagement to maintain strong partnerships with local provider groups and health networks. The continued sustainability of models like HLN is likely to rely heavily on the participation of referral sites and their ability to provide an adequate volume of referrals. This means program leaders will have to dedicate enough resources to provider recruitment, marketing, and relationship building.

Obtaining appropriate credentials for HLN specialists to provide telemedicine services at participating sites proved to be a lengthy, complicated, and expensive process. The credentialing process typically took 90–120 days; in some cases, it took more than six months. Each facility required HLN providers to be credentialed according to its unique bylaws and credentialing practices, and requirements varied significantly from one facility to the next. Furthermore, many of the clinics required the approval of their boards to work with HLN, and board meetings took place infrequently. All of these factors caused significant implementation delays.

In an attempt to streamline the process, HLN implemented a proxy credentialing process consistent with CMS Final Rule 76 FR 25550. This allowed proxy agreements to be created to credential and confer privileges to telemedicine providers, allowing them to provide services to patients in a state other than the state the provider is in. Staff cited this proxy credentialing procedure, which significantly streamlined the credentialing process for hospitals and clinics that were part of the network, as favorable to program implementation. However, staff at some individual facilities were unaware of this policy and believed they still had to go through their own credentialing processes. HLN staff suggested that having a national credentialing body could resolve many of these issues.

HLN staff reported that many of the providers and payers they worked with were unaware of the protocols and regulations relevant to providing mental health care via telemedicine:

- HLN staff encountered resistance to change in the provider communities in Washington, Montana, and Wyoming. Many providers, particularly those in rural areas, were unfamiliar with the concept of telepsychiatry, and educating health clinic staff about telepsychiatry was therefore crucial to the program. Some providers also feared that the innovation would disrupt the flow of their practice, and they were hesitant to implement it. As a result, HLN staff had to schedule multiple in-person meetings to develop trust in the program.
- HLN staff also reported that many payers had limited experience with telemedicine and often had misconceptions about what telepsychiatry involved and whether or not it was a covered benefit. This created difficulties in receiving reimbursement for telemedicine services.

HLN staff highlighted the impact of the program on changing the mindsets of providers, hospitals, and health plans and educating them about telepsychiatry. They observed that this was itself a significant positive result of the program. Ongoing educational efforts will contribute to the effectiveness of future telepsychiatry programs.

We did not observe any significant changes in measured outcomes for the HLN participants included in our impact analyses. The fact that the results did not show increases in expenditures, hospitalizations, or ED visits might suggest that results of the program were comparable to usual care. Notably, some of the outcomes included in our analysis do not reflect HLN's target outcomes and the intent of the program. Of the three core CMS outcome measures assessed, HLN sought to achieve only a reduction of hospitalizations over the project period. Other outcomes that HLN targeted were improved patient and provider satisfaction and a reduction in the number of missed school days for juvenile participants. We were not able to quantitatively assess HLN's success in achieving these outcomes. However, results from qualitative interviews indicated that providers and patients were satisfied with the quality of the mental health care provided by HLN.

Behavioral conditions among HLN participants ranged from mild mental health and somatic conditions to severe psychotic disorders. In addition, participants were geographically dispersed across the three intervention states. Therefore, outcomes could differ by diagnosis or geographic distribution. For example, the value of HLN services may be greater for patients residing in rural areas where access to mental health providers is more limited. However, due to data limitations, we were unable to perform subgroup analyses based on rural residency or on duration or severity of mental health diagnosis. Future research using data from programs similar to HLN that serve greater numbers of these groups would enable subgroup analyses to help determine whether programs of this nature are more helpful for particular populations.

VII. INSTITUTE FOR CLINICAL SYSTEMS IMPROVEMENT

Findings from Mathematica's Evaluation of the Institute for Clinical Systems Improvement (ICSI) HCIA Program

- Many participants in ICSI's Care of Mental, Physical and Substance Use Syndromes (COMPASS) program showed significant improvements on at least one of several measures of health status.
- Our findings provide some preliminary evidence that the COMPASS program may have contributed to lowering rates of hospitalization and emergency department visits for participants who were enrolled in Medicare fee-for-service (FFS) systems. Future research may address this by including a comparison or control group.
- Strong organizational leadership and physician buy-in at the partner and site levels were crucial for implementing the COMPASS program successfully.
- Many of ICSI's partners have sustained aspects of the COMPASS program, but the FFS payment structure has been an obstacle to sustainability.

A. Introduction

The Institute for Clinical Systems Improvement (ICSI), a regional healthcare improvement collaborative in Minnesota, implemented a collaborative care management model called Care of Mental, Physical and Substance Use Syndromes (COMPASS) with partners in eight states: California, Colorado, Florida, Massachusetts, Michigan, Minnesota, Pennsylvania, and Washington. ICSI and its partners, known collectively as the COMPASS consortium, tested national dissemination of the model. This project was designed to improve care and lower costs for 2,700 direct participants.[32] Participants were adults with depression[33] and uncontrolled diabetes or cardiovascular disease who were enrolled in Medicare or Medicaid.

Each COMPASS participant had a personalized care plan and a care team (including a primary care provider and a care manager) that incorporated consulting specialists (for example, a consulting psychiatrist and a specialty physician). The care team regularly conducted systematic case reviews of participants' progress and outcomes. During these reviews, consulting psychiatrists and physicians discussed with the care team participants' current treatment plan, medication management, and progress. The COMPASS program also included an electronic health record (EHR), which informed the case reviews and helped providers track care plans and outcomes.

In this model, the participant received no direct behavioral or mental health care from a behavioral health specialist. Consulting specialists took part only in the systematic case reviews; primary care providers and care managers executed any decisions made as a result of those

[32] Direct participants received care from a clinic or care manager funded by the Health Care Innovation Awards (HCIA). Indirect participants received COMPASS care, even though the clinic or care manager was not funded by HCIA dollars.

[33] An enrollee with a score greater than 9 on the widely used Patient Health Questionnaire-9 (PHQ-9) was categorized as having suboptimally managed depression.

reviews. The care manager functioned not only as a liaison between the care team and the participant, but also worked to reduce barriers to care by addressing participants' social or physical needs and by educating them about their health.

ICSI's eight partners included health plans, independent physician practice groups, large integrated health care systems, federally qualified health centers, and regional health care collaboratives. These partners varied substantially in their location and size, the characteristics of their participant populations, and their experience with the COMPASS program. Most clinical partners implemented the program at multiple sites. In addition, ICSI had four clinical settings that participated in the consortium. In its final report to CMMI, ICSI noted that 171 clinical sites participated in the program overall. ICSI also had technical partners that assisted in implementing and evaluating the intervention (for example, creating the EHR and analyzing EHR data)—including, the Advancing Innovative Mental Health Center (AIMS Center) and the HealthPartners Institute for Education and Research (HPIER).

The findings in this report are based on quantitative data received by September 2015 and qualitative data collected before June 1, 2015, as well as enrollment data reported throughout the award period.

For our quantitative analysis of the COMPASS program's effects, we used two primary approaches. First, for a subset of COMPASS partners, we conducted pre-post analyses of CMMI's four standard outcome measures (total expenditures, hospitalizations, readmissions, and emergency department [ED] visits) for selected direct Medicare participants. These analyses were based on Medicare claims and administrative data.

Second, for direct and indirect participants who had poor health status at baseline, we conducted pre-post analyses of key health status measures such as blood pressure and blood glucose levels. These analyses used EHR data provided by the same subset of partners who provided the data for the Medicare analysis and included participants regardless of insurance status.

In addition to conducting overall pre-post analyses, we also examined trends over time for the subset of COMPASS partners. This analysis was designed to shed light on potential variation across sites.

We draw on the following data sources for this chapter:

- **Enrollment data** submitted by ICSI to the reporting website maintained by CMMI's technical assistance contractor (the Lewin Group) for the HCIA Round 1 initiative.
- **Medicare and Medicaid IDs** submitted to Mathematica by COMPASS consortium partners (ICSI, the Mayo Clinic, the Pittsburg Regional Healthy Initiative, and the Michigan Center for Clinical Systems Improvement) through June 30, 2015. We used these IDs to extract data on Medicare program enrollment, utilization, and expenditures. These partners also submitted extracts of their EHRs to Mathematica.

- **Qualitative data**, including telephone interviews, focus groups, and in-person site visits in April and May of 2014 and 2015. Mathematica also conducted in-depth interviews with key awardee staff, members of the workforce, and other stakeholders during these visits.

In this chapter, we discuss the qualitative findings related to the administrative context, present the methods and findings from the pre-post analyses of the Medicare and EHR data, and discuss the qualitative findings related to sustainability and workforce factors. We close with a summary of lessons learned and conclusions. This report provides final evaluation results for ICSI.

1. Overview of administrative context

This section describes the administrative context of the COMPASS program, based on information gathered during site visits and focus groups with COMPASS staff members. We also discuss the organizational features and strategies of the COMPASS program, and place the COMPASS program in the demonstration life cycle.

ICSI implemented the COMPASS program across multiple states in a wide variety of medical settings. Partner organizations faced some difficulty in adapting their systems to implement the COMPASS program, which presented challenges to initial implementation and recruitment. ICSI leaders addressed these challenges by actively supporting their partners through the implementation process and responding effectively to their issues and concerns. By the end of the award, ICSI had met its enrollment target. ICSI was largely successful in implementing the COMPASS program throughout its partner organizations.

ICSI leaders and COMPASS staff members credited several factors as key to successfully implementing the COMPASS program:

- **Experienced staff.** Clinical staff at many COMPASS sites had experience with other programs that featured mental health integration, primary care redesign, and care coordination. In many cases, they viewed the COMPASS program as a natural extension of their ongoing activities and therefore easy to incorporate into normal processes of care. For example, staff from one partner organization, a federally qualified health center, already had social workers in clinical sites. The transition to the COMPASS program was easier for them because they were used to incorporating the care manager's role into their care teams.
- **Established community partnerships.** Some COMPASS sites leveraged existing partnerships—including, relationships with local YMCAs, community colleges, and fire departments—to better implement the COMPASS program in their communities. For example, one site included a social work intern from a local university in its team to help the nurse care coordinators meet participants' needs for social services.
- **Collaborative approach.** The structure of the COMPASS program—a lead coordinating organization (ICSI) and several implementation sites—posed administrative and legal challenges (for example, the need for partner and medical group subcontracts, business associates' agreements for common use of the COMPASS registry's EHR data, and licensing of the registry at participating sites). However, ICSI worked to ensure that partner sites viewed the COMPASS program as a collaborative effort rather than the work of one organization. This approach was especially important because of the diverse organizations

involved in the program. Staff reported that the collaborative approach fostered partner buy-in, transparency, accountability, and the sharing of experiences and lessons learned.

- **Supportive leaders.** Staff reported that the support of both administrative and clinical leaders was important to the successful implementation of the COMPASS program. In particular, enthusiastic and influential physician champions and committed mid-level staff were critical. One partner staff member explained:

  > We had challenges initially getting both leadership and their mid-level operational managers to buy in and really commit to the implementation, because you need the executive leaders to be allocating resources and really supporting their mid-level managers in order to operate. And you also needed the mid-level managers to be on board and really be tracking on the implementation and making sure that processes and workflows are getting established. So we had a lot of clinics where one or the other would be invested leadership or would think it's great, but management didn't see how it could happen. . . . Or vice versa, management felt very passionate about it, but leadership felt like they had too many other priorities as an organization to pursue it. And so really doing the communication and outreach and working to get the buy-in from both parties . . . is really critical to successful implementation and ongoing program sustainability at an organization.

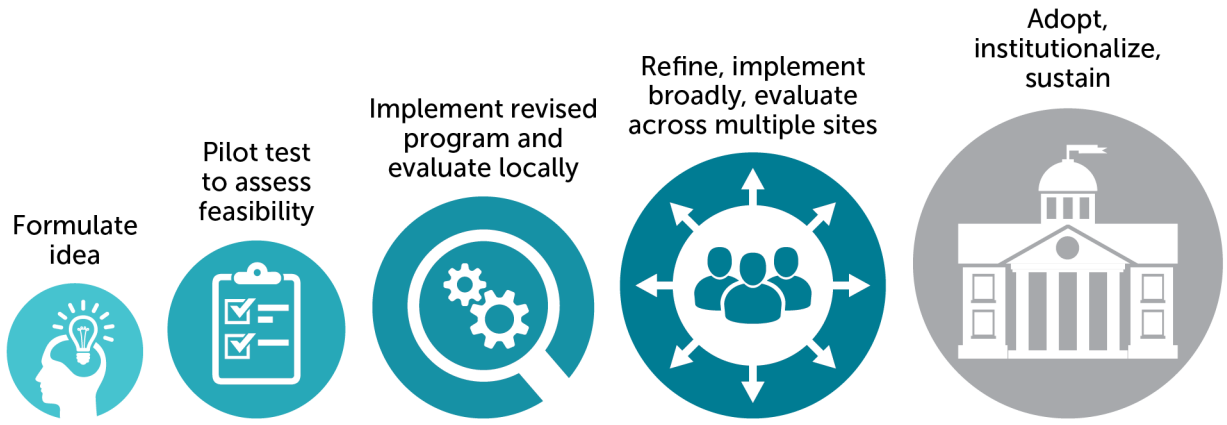
- **Composition of the case review team.** Respondents also emphasized the importance of having the right mix of individuals on the systematic case review team by selecting appropriate physician consultants and creating a multidisciplinary team with pharmacists and social workers to provide more insight and input in specialized areas. In addition, having a strong physician in a leadership role in the systematic case reviews—one who could champion the project—was useful for gaining and maintaining buy-in within sites.

- **Effective communication strategies.** Organizations that developed effective ways of communicating recommendations from the systematic case reviews to other members of the clinical practice were more successful in getting physicians to accept the program. Respondents reported that when care managers communicated recommendations (instead of having physicians or pharmacists do it), they met with varying levels of success. Existing relationships and practice cultures affected whether care managers were able to influence the group.

2. Progression through the phases of innovation

Before receiving HCIA funds, ICSI had already implemented a four-year statewide initiative (referred to as DIAMOND) that used a similar collaborative care management model for tracking and treating depression. The COMPASS program revised that model by adding in physical health conditions (uncontrolled diabetes and cardiovascular disease) and working with a broad range of partner organizations and clinic types. Given this history, the COMPASS program moved through four of the five phases of an innovation (Figure VII.1): ICSI used previous work to refine the model and implement it broadly across multiple sites.

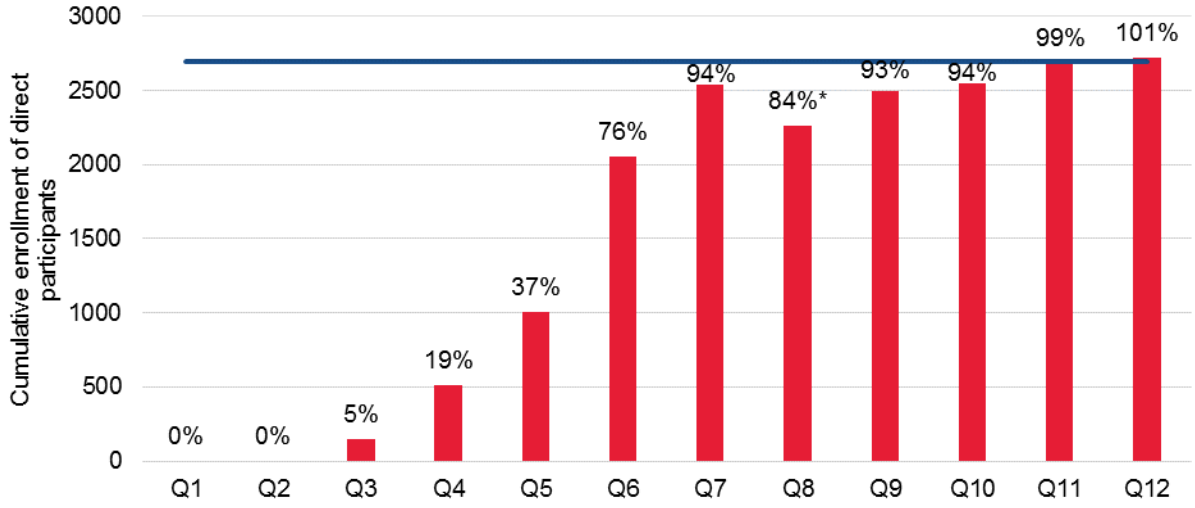
Figure VII.1. Phases of program innovation: ICSI



3. Enrollment

By the end of the 12th quarter (June 30, 2015), ICSI had enrolled 2,726 direct participants, slightly exceeding its original enrollment target of 2,700 (Figure VII.2). In the first two quarters, ICSI was initiating and piloting the project. Enrollment began in the third quarter with 148 direct participants, which increased to a total of 509 direct participants in quarter 4. In quarters 5 through 12, enrollment levels of direct participants fluctuated but rose over the award period overall, with ICSI reaching its enrollment target in quarter 12.

Figure VII.2. Percent of direct participant target enrollment, Q1–Q12



Source: Awardee's enrollment data reported to the website maintained by CMMI's technical assistance contractors (the Lewin Group).

Note: The blue horizontal line represents ICSI's target enrollment of 2,700 unique direct participants.

*In Q8, some participants dropped out of the intervention, which reduced the overall enrollment numbers (personal communication with ICSI leaders, June 15, 2015).

4. Participant demographics

Each COMPASS program site targeted enrollees who had active depression and uncontrolled diabetes, cardiovascular disease, or both. ICSI enrolled two subgroups of participants: direct and indirect. Table VII.1 shows the demographics of the total population (both direct and indirect) served by the COMPASS program.

Table VII.1. Demographic characteristics and insurance status of ICSI direct and indirect participants

	Number of participants	Percent of participants
Total	3,992	100
Age (years)[a]		
Younger than 18	0	0.0
18–34	146	3.7
35–44	382	9.6
45–54	874	21.9
55–64	1,192	29.9
65 or older	1,393	34.9
Gender		
Male	1,432	35.9
Female	2,535	63.5
Unknown	25	0.6
Insurance coverage		
Medicaid, non-dual	754	18.9
Medicare, non-dual	1,886	47.2
Dual	180	4.5
Private insurance	1,087	27.2
Other (includes other insurance, unknown, and uninsured)	85	2.1
State		
California	884	22.1
Colorado	323	8.1
Florida	91	2.3
Massachusetts	96	2.4
Michigan	463	11.6
Minnesota	1,212	30.4
Pennsylvania	744	18.6
Washington	179	4.5

Source: ICSI-generated data.

[a]Age groups do not sum to total due to individuals with unknown age.

B. Methods

In this section, we describe our methods for conducting quantitative analyses and collecting qualitative data. We begin by describing our methods for conducting pre-post analyses of four core outcome measures for Medicare fee-for-service (FFS) beneficiaries who participated in the COMPASS program: (1) total expenditures, (2) hospitalizations, (3) readmissions, and (4) ED visits. We then discuss our methods for conducting pre-post analyses of EHR data and the limitations of the quantitative analyses. We conclude with a discussion of qualitative data collection methods.

1. Medicare pre-post analysis

By using Medicare claims data accessed through the Virtual Research Data Center (VRDC), we conducted pre-post analyses of Medicare-covered service utilization and expenditures for ICSI program participants. The Medicare claims data available through the VRDC are limited to FFS Medicare enrollees. Therefore, we identified four ICSI partners (with a total of 18 different clinic sites) with relatively large numbers of Medicare FFS participants for inclusion in this analysis.[34]

By using Medicare identification numbers provided by the partners, we extracted Medicare claims from the VRDC for COMPASS program participants. We limited the analytic sample for the pre-post analysis to participants who met the following criteria in the baseline period—that is, the year before they enrolled in the COMPASS program—based on the Medicare administrative data: (1) enrolled in Medicare FFS Parts A and B, with Medicare as primary payer for at least six months; (2) had a diagnosis of either depression, cardiovascular disease, or diabetes; and (3) resided in one of four states (Michigan, Minnesota, Pennsylvania, or Wisconsin[35]). The residence criterion was based on the location of the partners we selected for inclusion in this analysis. The final analytic sample included 481 direct participants. Appendix A includes a table showing the demographic characteristics of this sample and provides more information on the identification of this sample.

We determined the reference period for the pre-post analysis relative to program enrollment for each participant. Enrollment dates varied across the participants in our sample, from February 2013 through June 2015. The baseline period (that is, the pre-period) included data for up to 24 months prior to enrollment. The follow-up period (that is, the post-period) included data for the month of enrollment and up to 24 additional months. We created periods representing 6-month time frames across the 24-month baseline period and the 24-month intervention period.

We estimated a regression model for each outcome, including indicators for each six-month time period. For each outcome, the model included the following control variables: (1) age; (2) gender; (3) race (white, black, and other); (4) original reason for entitlement (retirement age versus other); (5) dual Medicare-Medicaid enrollment status; (6) indicator for length of enrollment in the COMPASS program; and (7) hierarchical condition categories (HCC) score. The model also included indicators for the diagnoses of various chronic conditions—including, acute myocardial infarction; Alzheimer's disease; anemia; arthritis; asthma; atrial fibrillation; cancer (breast, colon, endometrial, lung, and prostate); cataract; chronic kidney disease; chronic obstructive pulmonary disease; congestive heart failure; depression; diabetes; glaucoma; hip/pelvis fracture; hyperlipidemia; hyperplasia; hypertension; hypothyroid; ischemic heart disease; osteoporosis; and stroke. We used the estimates from this regression model to calculate the regression-adjusted means presented in the figures in Section C.1.

[34] Appendix A includes information on our selection of ICSI's partners and the partners involved in this analysis.

[35] Although ICSI did not have a clinical site in Wisconsin, some participants lived close to the Minnesota-Wisconsin border and were served by a Mayo Clinic in Minnesota.

2. EHR data analysis

The same COMPASS consortium partners who provided the Medicare IDs submitted COMPASS registry data (EHR data) on all participants (both direct and indirect). Each partner submitted a full extract of its EHR data through September 1, 2015. Data extracts consisted of multiple files that documented care plans, diagnoses, care management notes, medication history, and hospitalizations. We reviewed and de-duplicated the data. Appendix A shows the demographic data for the EHR sample. Characteristics of the population for which EHR data were provided differed only slightly from the overall population that ICSI served.

We examined the following health status measures at baseline and at six-month and twelve-month follow-ups: (1) body mass index (BMI), (2) diabetes control, (3) blood pressure (BP) control, (4) low-density lipoprotein (LDL) level, and (5) depression severity. We constructed these measures based on the data recorded at the time of enrollment. We created both continuous and categorical variables for each health status measure. Table VII.2 provides further details.

Table VII.2. Health status measures

Outcome	Measure used	Categories	Guidelines used
Depression	PHQ-9	• Minimal Depression (PHQ-9 score 1-4) • Mild Depression (PHQ-9 score 5-9) • Moderate Depression (PHQ-9 score 10-14) • Moderately Severe Depression (PHQ-9 score 15-19) • Severe Depression (PHQ-9 score 20-27)	U.S. Preventive Services Task Force[a]
BMI	Weight and height	• Underweight = BMI < 18.5 • Normal weight = BMI 18.5–24.9 • Overweight = BMI 25–29.9 • Obese = BMI > 30	NCQA[b]
BP control	Systolic and diastolic blood pressure (mmHG)	• Hypertension = Enrollees ages 18 to 59 whose BP was > 140/90 mmHg • Hypertension = Enrollees ages 60 to 85 with a diagnosis of diabetes whose BP was > 140/90 mmHg • Hypertension = Enrollees ages 60 to 85 without a diagnosis of diabetes whose BP was > 150/90 mmHg	NCQA[c]
LDL	Blood LDL cholesterol level (mg/dL)	• Optimal = < 100 mg/dL • Near optimal = 100–129 mg/dL • Borderline high = 130–159 mg/dL • High = 160–189 mg/dL • Very high = > 190 mg/dL	NHLBI[d]
Diabetes control	Hemoglobin (Hb) A1c percentage	• Controlled = HbA1c < 8.0 percent • Borderline control = HbA1c >= 8.0 percent and <= 9.0 percent • Poor control = HbA1c > 9.0 percent	HEDIS[e]

[a]Information about the scoring of the PHQ-9 can be found at http://www.uspreventiveservicestaskforce.org/Page/Document/UpdateSummaryFinal/depression-in-adults-screening.

[b]See http://www.nhlbi.nih.gov/health/educational/lose_wt/BMI/bmi_dis.htm.

[c]See http://www.ncqa.org/report-cards/health-plans/state-of-health-care-quality/2016-table-of-contents/controlling-high-blood-pressure.

[d]See http://www.nhlbi.nih.gov/health/resources/heart/heart-cholesterol-hbc-what-html.

[e]See http://public.optimahealth.com/lists/optimaformslibrary/qi-hedis-definitions.pdf.

HEDIS = Healthcare Effectiveness Data and Information Set; NCQA = National Committee for Quality Assurance; NHLBI = National Heart, Lung, and Blood Institute; PHQ-9 = Patient Health Questionnaire-9

Our analysis focused on participants whose health status measures at baseline indicated a need for improvement. For each health status measure, we conducted pre-post analyses focused on detecting improvements at 6 and 12 months after enrollment. We conducted additional analyses to examine potential outcome differences across the four sites and between Medicare participants (included in the analysis mentioned above) and non-Medicare participants not included in the analysis but present in the EHR data. We also used the EHR data to conduct an analysis of the effectiveness of the systematic case review process.

3. Limitations of quantitative analyses

One of the most important limitations of the quantitative analyses is the absence of a comparison group. As a result, we have no counterfactual evidence—that is, information about changes in the outcome variables that might have occurred if participants had not been enrolled in the COMPASS program. Due to the lack of a comparison group, all findings from this analysis should be considered exploratory in nature.

As described in our second annual report, we made considerable effort to construct a comparison group that would provide adequate counterfactual evidence. However, many of ICSI's enrollment criteria were based on data not observable in Medicare claims or administrative files. Hence, we were severely limited in the variables that we could have used to generate a virtual comparison group through matching procedures. Specifically, we could have matched on some variables such as age and gender, but not on others such as the Patient Health Questionnaire-9 (PHQ-9) and many of the clinical indices used to enroll program participants. As a result, we would not have known the extent to which the participant group and the matched comparison group might have differed on these important indices.

Missing data also limited the extent to which we could analyze the EHR data and participant health status outcomes. We discuss this issue further in Section C.3.

4. Qualitative methods

During site visits in spring 2014, we conducted in-depth interviews with key staff of the awardee, including care managers, physicians, consultants, supervisors, ICSI administrative staff, and stakeholders such as evaluation team members. In total, we conducted 20 interviews with 34 individuals. In spring 2015, we conducted a second round of site visit interviews with the same types of team members. The majority of respondents were the same in both years, but some differed due to staff turnover or changes in staff roles. In 2015, we conducted a total of 21 interviews with 32 individuals. In addition, we convened a focus group with seven additional care managers at the partner sites during the 2015 site visit. These care managers did not participate in the individual interviews we conducted during the second site visit. Both site visits took place at ICSI's central office.

We also conducted phone interviews with leaders and clinical staff at all participating partners during both site visits. In May 2016, we conducted one follow-up telephone interview with ICSI leaders to discuss sustainability efforts since the end of the award in June 2015.

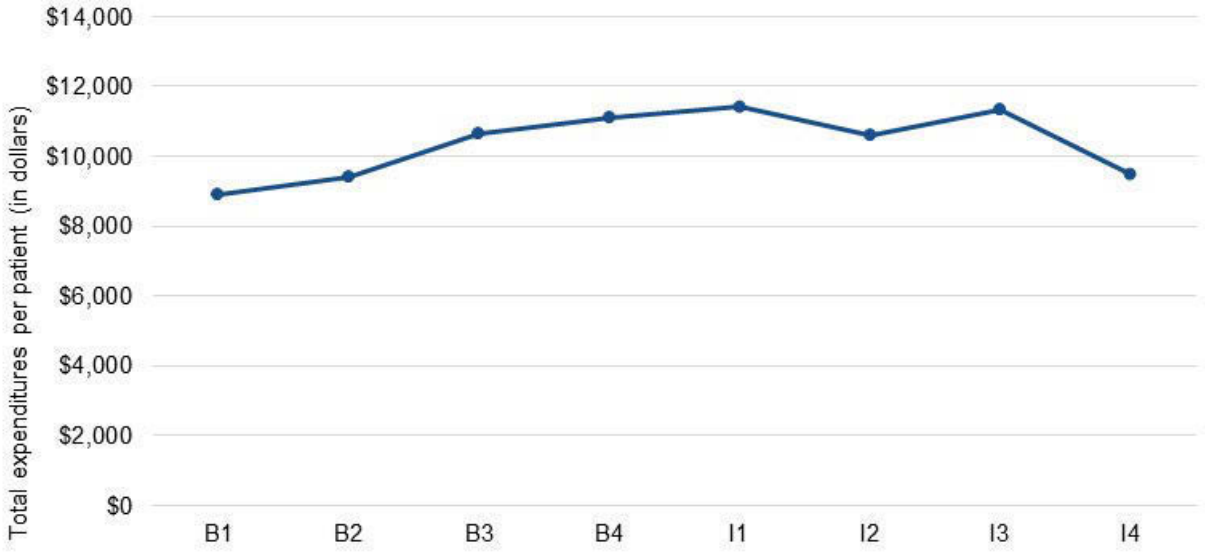
C. Summative findings

In this section, we present our findings from pre-post analyses of four core outcome measures among Medicare FFS beneficiaries who participated in the COMPASS program. We then discuss findings from the pre-post analysis of EHR data. We conclude with a discussion of our findings from the analysis of qualitative data on the perceptions of program effects.

1. Medicare FFS findings

Expenditures. Pre-post analyses showed no significant differences between the average expenditure during the baseline period and expenditure levels during any of the post-intervention periods. As shown in Figure VII.3, from the beginning of the baseline period (B1) to the end of the first intervention period (I1), expenditures were rising. During the second intervention period (I2), expenditures fell somewhat. They rose during the third intervention period (I3), and then decreased again during the fourth intervention period (I4).

Figure VII.3. Total expenditures per participant per six-month period

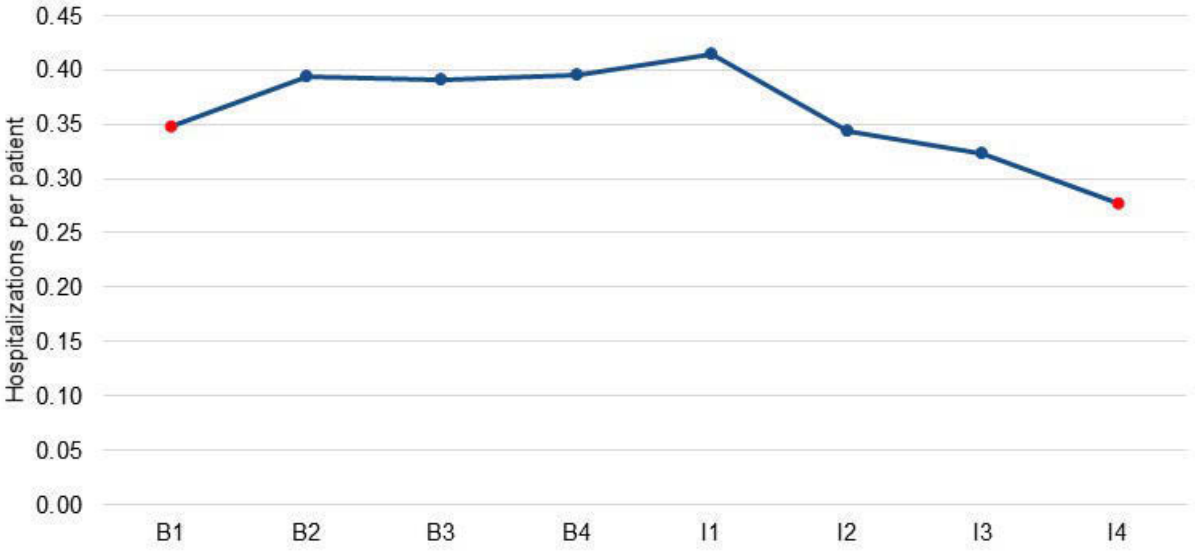


Source: Mathematica analysis of FFS Medicare claims data for the baseline and program periods, February 2011 to June 2015.

Note: Regression-adjusted means for the intervention population (N = 481) were based on characteristics at enrollment (I1). The regression model controlled for age; gender; race (white, black, and other); original reason for entitlement (retirement age versus other); length of enrollment; dual eligibility status; HCC score; and diagnoses of various chronic conditions. The number of observations per time point varies based on the number of participants enrolled in Medicare at any point in time.

Hospitalizations. Pre-post analyses indicated a significant decrease in the hospitalization rate at the fourth intervention period compared with the average baseline rate (Figure VII.4). The hospitalization rate trended upward from the beginning of the baseline period through the first intervention period and then fell consistently after that point. At the fourth intervention period, the average hospitalization rate was 0.28 per person—significantly different from the baseline average of 0.38 per person ($p < 0.05$).

Figure VII.4. Hospitalizations per participant per six-month period



Source: Mathematica analysis of FFS Medicare claims data for the baseline and program periods, February 2011 to June 2015.

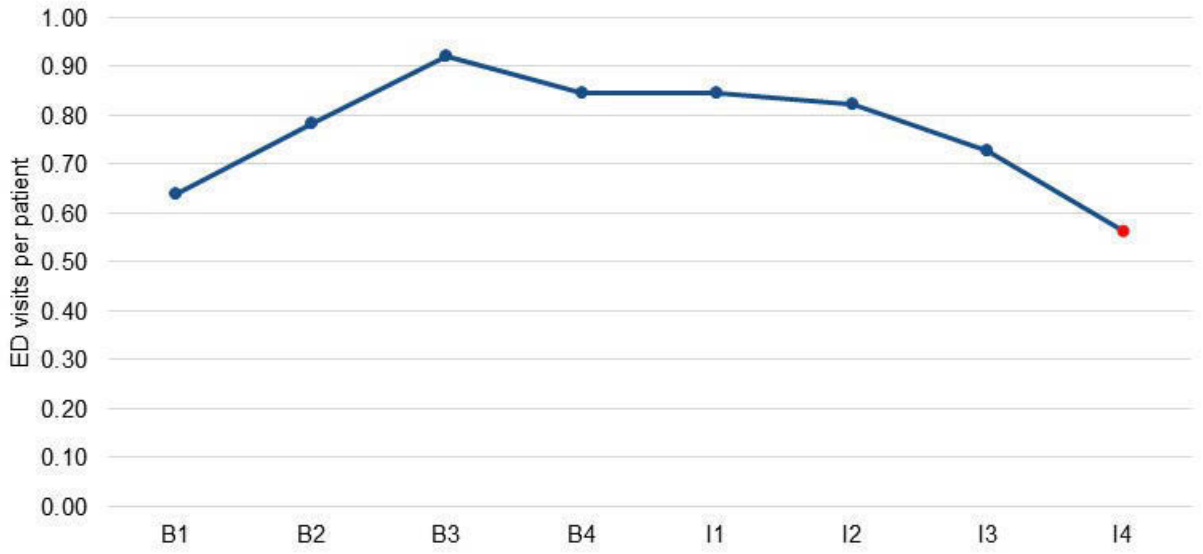
Notes: Regression-adjusted means for the intervention population (N = 481) were based on characteristics at enrollment (I1). The regression model controlled for age; gender; race (white, black, and other); original reason for entitlement (retirement age versus other); length of enrollment; dual eligibility status; HCC score; and diagnoses of various chronic conditions. The number of observations per time point varies based on the number of participants enrolled in Medicare at any point in time.

The red dot indicates that the intervention period mean was significantly different from the average across the baseline periods at the 90 percent confidence level.

Readmissions. Participants experienced too few readmissions to analyze for this report.

ED visits. Pre-post analyses revealed a significant decrease in the ED visit rate at the fourth intervention period compared with the average baseline rate (Figure VII.5). The ED visit rate climbed steadily through the baseline period until it began to decline in the fourth baseline period. The rate continued to fall throughout the intervention period. At the fourth intervention period, the ED visit rate was 0.28 per person. This was significantly lower than the baseline average of 0.79 ($p < 0.01$).

Figure VII.5. ED visits per participant per six-month period



Source: Mathematica analysis of FFS Medicare claims data for the baseline and program periods, February 2011 to June 2015.

Notes: Regression-adjusted means for the intervention population (N = 481) were based on characteristics at enrollment (I1). The regression model controlled for age; gender; race (white, black, and other); original reason for entitlement (retirement age versus other); length of enrollment; dual eligibility status; HCC score; and diagnoses of various chronic conditions. The number of observations per time point varies based on the number of participants enrolled in Medicare at any point in time.

The red dot indicates that the intervention period mean was significantly different from the average across the baseline periods at the 90 percent confidence level.

2. Summary of Medicare FFS analysis

Our pre-post analysis of total expenditures for COMPASS participants enrolled in FFS Medicare provides some indication that expenditures may have begun to decrease after the first six months of exposure to the COMPASS intervention, relative to a trend of increasing expenditures that started at the beginning of the baseline period. This decline may have been driven by observed declines in hospitalizations and ED visits. The expenditure levels observed in the intervention period, however, were not significantly lower than the baseline average. Patterns of expenditures, hospitalizations, and ED visits varied across the four sites. This may suggest variability in the populations at each site or variations in implementation of the COMPASS intervention. More research is needed to fully parse out potential site differences.

These results should be should be interpreted with caution. Moreover, without a comparison group, we cannot determine whether the observed changes in the outcome variables might have occurred if participants had not been enrolled in the COMPASS program. At this point, the analyses provide a very preliminary suggestion that the COMPASS program may have eventually resulted in a statistically significant decrease in overall expenditures for the Medicare beneficiaries.

Sample sizes also vary across the time points. All participants eligible for our analysis are included in the first intervention period (I1) estimate. However, since our analysis only includes

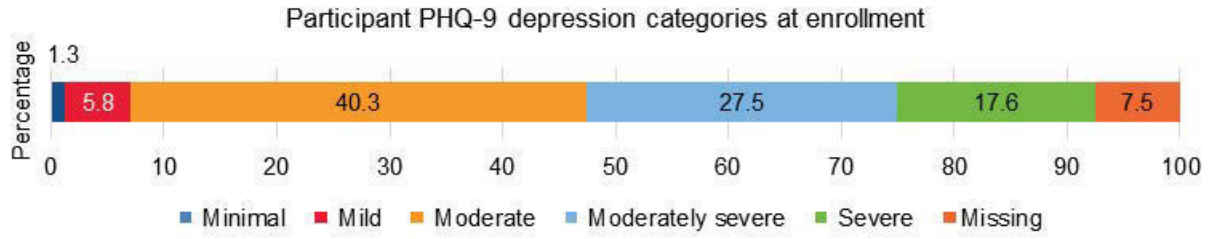
data through June 2015, we only observe those participants who enrolled between February and June 2013 for the full 24-month analysis period or through the fourth intervention period (I4). If these early enrollees differ from later enrollees on characteristics that are not controlled for in the regression model, such as severity of depression, this could result in differences between the I4 estimate and the estimates from earlier intervention periods.

3. EHR findings

We analyzed EHR data to determine how the COMPASS intervention affected various health outcomes at 6 and 12 months after enrolling in the program. In addition, we examined outcomes by site—between those included in our Medicare FFS analysis and those who were not (non-Medicare population—as well as the utility of the systematic case review.

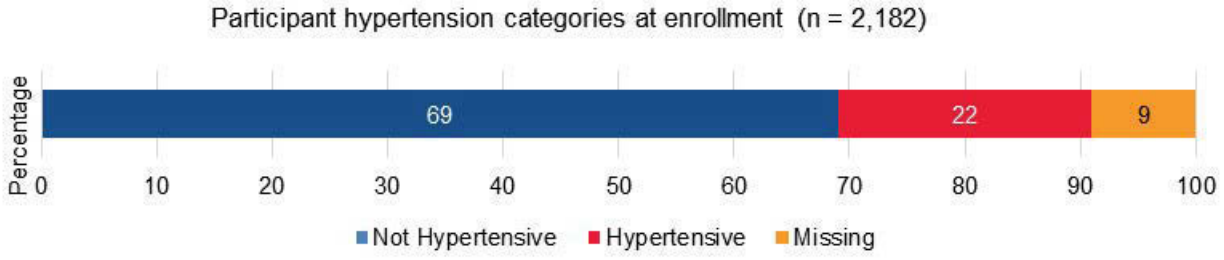
Baseline health status. The four partner sites submitted EHR data on 2,182 participants. Most participants had valid data at enrollment for measures of depression, blood pressure, and diabetes control, but measures of BMI and LDL suffered from substantial missing data. Consistent with enrollment criteria, approximately 85 percent of participants had moderate to severe depression (as determined by provision of valid PHQ-9 scores). Furthermore, for each additional health status measure, values were suboptimal for a substantial portion of the population. For example, of those with valid data, just under 50 percent had poor or borderline control of their diabetes, with roughly the same percentage being overweight or obese according to their BMI. Lower rates of hypertension (22 percent) were observed. Figures VII.6 through VII.10 show the distribution of participants across health status categories at enrollment.

Figure VII.6. Depression categories at enrollment



Source: ICSI EHR data at enrollment.

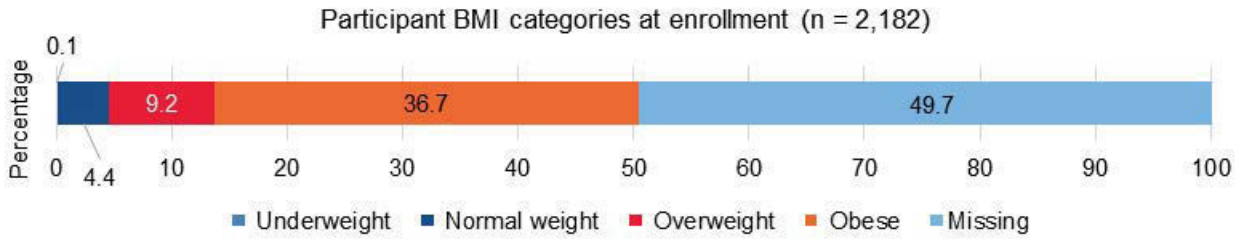
Figure VII.7. Hypertension categories at enrollment



Source: ICSI EHR data at enrollment.

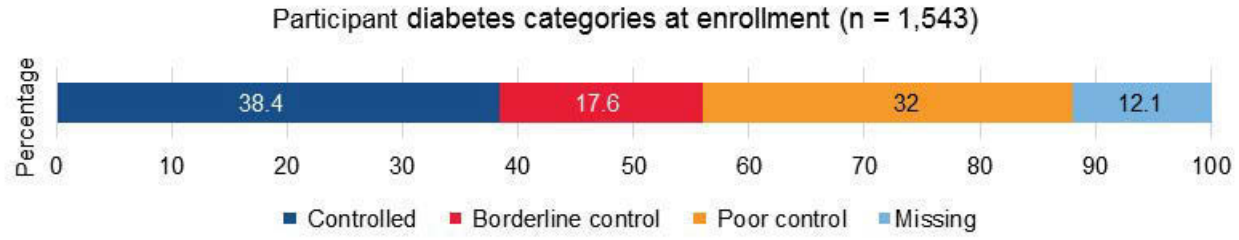
Note: Hypertension is determined by blood pressure values. See Table VII.2 for details.

Figure VII.8. BMI categories at enrollment



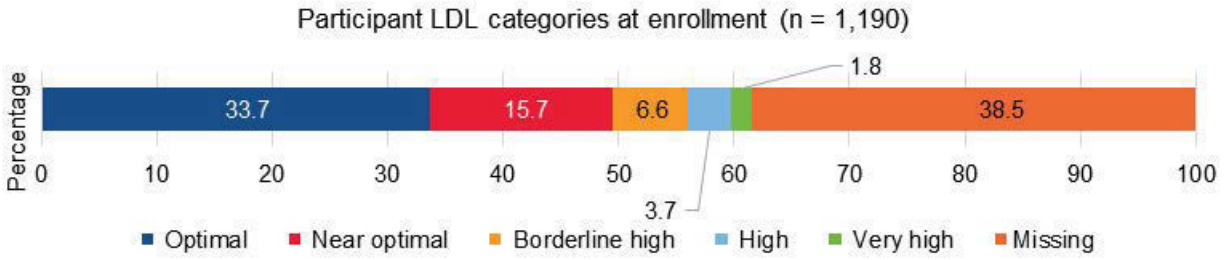
Source: ICSI EHR data at enrollment.

Figure VII.9. Diabetes control categories for participants with a diagnosis of diabetes at enrollment



Source: ICSI EHR data at enrolment.
Note: This figure only includes participants with a diagnosis of diabetes noted in their EHR record.

Figure VII.10. LDL categories for participants with a diagnosis of heart disease or congestive heart failure at enrollment

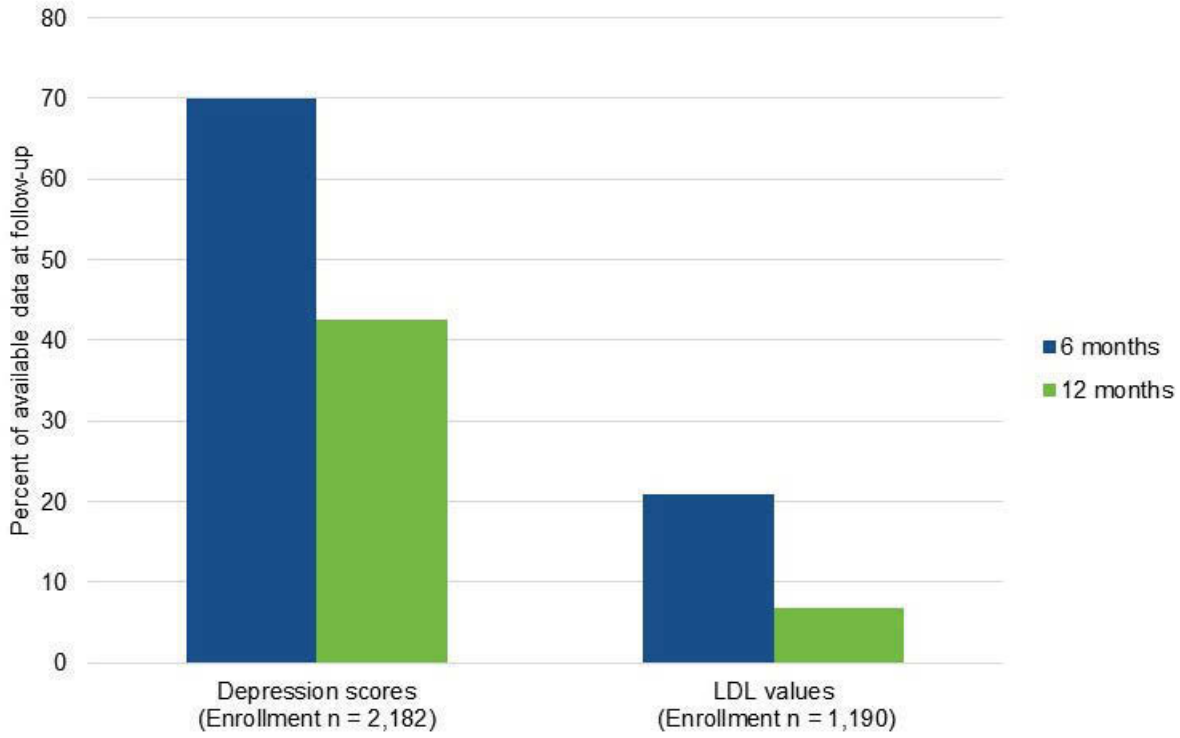


Source: ICSI EHR data at enrollment.
Note: This figure only includes participants with a diagnosis of heart disease or congestive heart failure noted in their EHR record.

Missing data. A substantial share of program participants did not have information recorded in the EHR for each health status measure at enrollment. The level of missing data for each health status measure is indicated in orange (Figures VIII.6 through VIII.10). The lowest rates of missing data were observed for the depression measure (7.5 percent), blood pressure (9 percent), and diabetes (12 percent). BMI and LDL had the highest rates of missing data at baseline (49.7 percent and 38.5 percent, respectively).

Data for the 6- and 12-month time frames were also missing for a substantial portion of the population. For example, only two sites recorded statin prescriptions[36] (for treatment of cholesterol risk groups) at both time points. Updating current medications was not required for the intervention. Another example is seen in depression and LDL values at 6 and 12 months. Depression measures were provided for 69.9 percent of the 2,182 participants at 6 months. This decreased at 12 months to 42.5 percent. LDL values had a large reduction in available data over time. LDL values were available for only 20.9 percent and 6.8 percent of the 1,190 participants at 6 and 12 months, respectively. Based on this, LDL and BMI were dropped as an outcome variables. Figure VII.11 shows the rates of available data for these two variables (the full chart is shown in Appendix A).

Figure VII.11. Percent of available data at 6 months and 12 months for depression and LDL outcomes



Source: ICSI EHR data.

Outcomes at 6 and 12 months. We examined 6- and 12-month outcomes for participants with suboptimal health status (that is, moderate to severe depression, hypertension, or uncontrolled diabetes) at enrollment. We did not conduct analyses of LDL and BMI because of missing data and small sample size, as this could limit generalizability and introduce potential bias. For this analysis, patients were only selected if they had baseline, 6-month, and 12-month data.

[36] Statin prescriptions are the standard treatment for those with or at risk for high cholesterol.

We assessed change in the health status measure at 6 and 12 months post-intervention using continuous measures (that is, change in total PHQ-9 score, HbA1c levels, and BP). Results showed significant decreases for depression, blood pressure, and diabetes ($p < 0.001$). The mean PHQ-9 depression score was 15.6 (moderately severe depression) at enrollment, which decreased to 9.1 and then to 8.3 (both indicating mild depression) at 6 and 12 months, respectively. The mean systolic blood pressure was 152.1 mmHG (hypertensive) at enrollment, which decreased to levels of 135.3 mmHG and 131.7 mmHG at 6 and 12 months, respectively (not hypertensive is less than 140 mmHG). The mean HbA1c level at enrollment was 10.8 percent (poor control), which decreased to 9.2 percent and then to 9 percent (borderline control) at 6 and 12 months, respectively. Table VII.3 shows these results.

Table VII.3. Health status measure values at enrollment, 6 months, and 12 months for ICSI participants with suboptimal health status at enrollment

		Enrollment		6 months		12 months	
Categories	N	Mean	STD	Mean	STD	Mean	STD
Moderate to severe depression at enrollment (measured in PHQ-9 score)	870	15.6	4.3	9.1*	5.9	8.3*	6
Hypertensive at enrollment (measured in systolic blood pressure)	202	152.1	15.1	135.3*	17.6	131.7*	15.7
Diabetes poor control at enrollment (measured in HbA1c levels)	183	10.8	1.4	9.2*	1.8	9*	1.9

Source: ICSI EHR data.

*Paired t-test conducted to examine differences between enrollment and follow-up periods; significantly different from baseline at $p < 0.001$.

STD = standard deviation

We conducted an additional analysis to examine the proportion of participants who achieved optimal health status during the follow-up period. The purpose of this analysis was to determine the proportion of participants who achieved a recommended level or improvement for each outcome at follow-up, as defined by the following:[37]

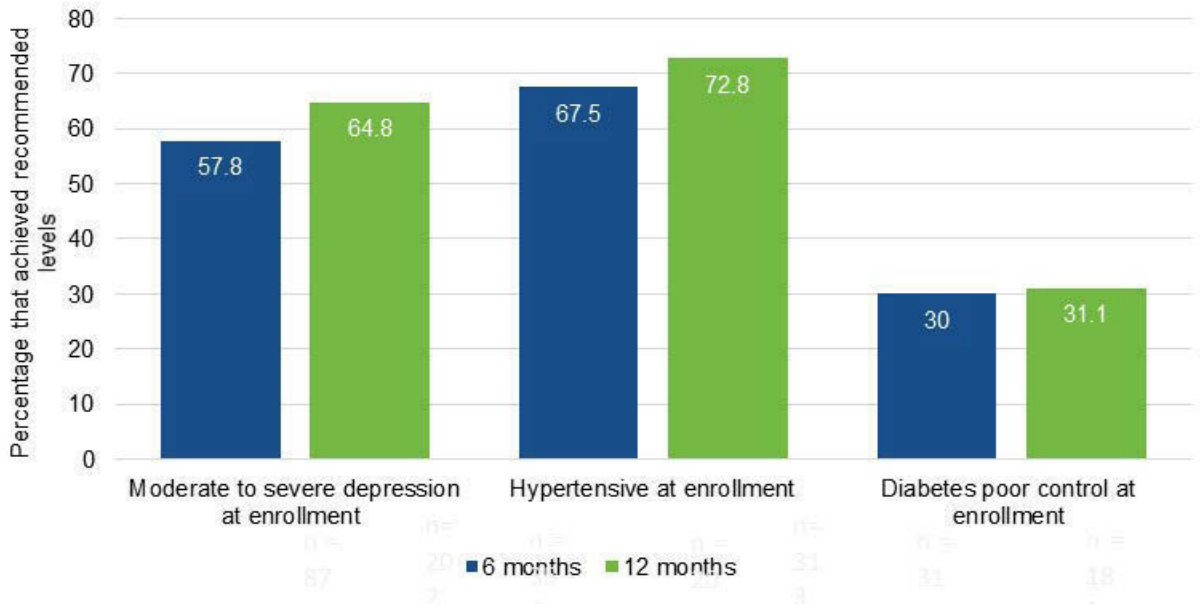
- **Depression.** No or mild depression as indicated by the PHQ-9 score (scores 1 to 9)
- **Hypertension.** No hypertensive blood pressure readings, as defined by blood pressure less than 140/90 mmHG and age range (for example, participants age 18 to 59 whose blood pressure was less than 140/90 mmHG)
- **Diabetes.** Controlled HbA1c (less than or equal to 8 percent)

Figure VII.12 shows the percent of participants with suboptimal values at enrollment who achieved recommended levels or improvements at follow-up. Over 50 percent of those with suboptimal depression or hypertension had achieved a recommended level at 6 months. Only slightly more had achieved recommended levels by 12 months (fewer than a 10 percent increase from the percentage at 6 months). Thirty percent of those with poorly controlled diabetes

[37] More detailed information can be found in Table VII.2.

achieved recommend HbA1c levels at 6 months, with only a small further increase at 12 months (31.1 percent).

Figure VII.12. Percent of ICSI participants with suboptimal health status at baseline who achieved recommended levels at 6- and 12-month follow-ups



Source: ICSI EHR data.

Note: Sample sizes represent the number of participants whose symptoms were uncontrolled at enrollment.

Outcomes by site. We examined outcomes by site (shown in Appendix A). Outcomes at 6 months were similar across the sites. Sites also had similar patterns of improvement at 12 months. Diabetes control showed a larger increase at 12 months for ICSI (73.8 percent) compared to the other sites (which ranged from 18.2 percent to 38.7 percent). This may indicate variations in the use of the COMPASS EHR. It could be that ICSI sites were more diligent at entering HbA1c information into the EHR (as these data were included with statin prescriptions). In addition, sample size varied by site. This variation may also reduce the accuracy of these results. These results should be interpreted with caution.

Outcomes by Medicare analysis inclusion. We conducted another outcomes analysis to assess whether there were substantial differences on health status measure outcomes between participants with and without Medicare FFS insurance coverage (Appendix A shows these results). Some differences were observed between the two groups. For example, participants without Medicare FFS coverage had a higher rate of improvement in hypertension at 6 months than participants with Medicare FFS coverage (65.3 percent versus 68.1 percent). Furthermore, the share of Medicare participants with improvements declined at 12 months from 65.3 percent to 59.5 percent, while the share for non-Medicare participants increased from 68.1 percent to 76.3 percent. It is unclear why this difference occurred; it may be due to sample sizes at the 12-month time frame. With the exception of this difference in 12-month improvement rate for hypertension, the analysis suggested that those included in our Medicare FFS analysis had similar outcomes to those who were not included.

Development of new conditions. Part of the COMPASS program was the systematic case review. This process helped to monitor participant progress and modify treatment or care plans. We examined the percentage of participants with optimal values on the health measures (for example, not hypertensive) at enrollment who had values indicating uncontrolled conditions at 6 months. We then calculated the portion of this subgroup whose conditions were controlled again at 12 months. Of the over 600 participants who were not hypertensive at enrollment, 82 (13.1 percent) became hypertensive at 6 months. The hypertension was controlled at 12 months for 65 of these 82 participants (79.3 percent). Likewise 17.0 percent of participants whose diabetes was controlled at baseline had uncontrolled status at 6 months. Among this 17 percent diabetes was controlled at 12 months for only 41.9 percent.[38] Based on these results, the development of new symptoms was common, but new symptoms were often controlled by the time of the 12-month follow-up assessment. The systematic case reviews and the responsiveness of the care teams may have contributed to identifying and controlling these new conditions.

4. Summary of EHR analysis

According to our analysis of EHR data, many participants who began the COMPASS program with suboptimal health status improved during their participation. Over 50 percent of those participants who were hypertensive or moderately to severely depressed at enrollment improved at 6 and 12 months. A smaller proportion (20 percent to 40 percent) of those who had uncontrolled diabetes improved at 12 months. The HbA1c health status measure appeared to be the hardest to improve at 12 months. These results were similar among Medicare FFS and non-Medicare subgroups. The substantial proportion of participants for whom outcomes did not improve in 12 months may speak to limitations of treatment options or characteristics of the population that make medical management particularly challenging (for example, comorbid conditions).

Among participants who initially had optimal indications for health status, only a modest portion developed new conditions (for example, controlled hypertension at enrollment, now hypertensive at follow-up) at 6 months. Of those who developed new conditions, many improved at 12 months. The systematic case reviews and the responsiveness of the care teams may have contributed to these outcomes.

These results are fairly similar to other research using similar interventions. For example, Katon and colleagues[39] used a collaborative care model to treat depression and chronic illness (diabetes and cardiovascular disease). They observed a 0.58 percent difference in HbA1c levels at 12 months (baseline = 8.14 percent, 12 months = 7.33 percent). Our results for ICSI show higher baseline HbA1c levels and a 1 percent reduction in levels at 12 months. They also observed a 5.1 mmHG difference in systolic blood pressure at 12 months (baseline = 135.7, 12 months = 131.0). As for HbA1c, our analysis revealed higher baseline values and a larger reduction (20.4 mmHG) in systolic blood pressure at 12 months.

[38] Only hypertension and diabetes are discussed due to small sample sizes in the other health status measures.

[39] Katon, Wayne J., Elizabeth H.B. Lin, Michael Von Korff, Paul Ciechanowski, Evette J. Ludman, Bessie Young, Do Peterson, Carolyn M. Rutter, Mary McGregor, and David McCulloch. "Collaborative Care for Patients with Depression and Chronic Illnesses." *New England Journal of Medicine,* vol. 363, no. 27, 2010, pp. 2611–2620.

Katon and colleagues used a different measure of depression than ICSI, so a direct comparison of depression values was not appropriate; however, computation of effect sizes was possible. These authors observed an effect size of 1.42 in terms of depression reduction. The effect size for our analysis of ICSI PHQ-9 depression scores was 1.40. Although the interventions were similar, the sample sizes and study designs were not. For example, Katon and colleagues utilized an experimental design and only focused on one health care system in Washington State with 14 primary care clinics and 214 participants. Based on these differences, comparisons should be interpreted with caution.

Our analysis had two primary limitations. First, data were missing for a substantial share of participants (ranging from 7.5 to 49.7 percent, depending upon the health status measure). Missing data reflect attrition from the program and a lack of reporting, which may have been due to the burden that the COMPASS EHR placed on the staff (double entry in two systems) and a lack of standardization of data entry requirements by the COMPASS program. Outcomes for individuals with missing data may differ from the findings reported here.

The rates of missing data suggest that the program's EHR system was not consistently used by participating sites to collect health status data at each visit. During our qualitative site visits, some partner staff members discussed the burden that the COMPASS EHR placed on them. These partner sites often had another EHR that they were required to use at their institution. These staff members reported that the double entry in both systems often took too much time so that sometimes information was only recorded in their institutional EHR.

ICSI leaders reported that they required quality control on some measures but did not require other measures to be updated. For example, LDL levels were not required to be entered. ICSI leaders cited changes during the project in the national recommendations for checking LDL. The national recommendation no longer included standards for LDL testing intervals. Instead, the national guidelines focused on prescribing and adherence to statin medications. As a result, ICSI no longer required these data to be entered. However, only two sites documented statin prescriptions in the ICSI EHR. Blood pressure, on the other hand, was one of the data elements for which ICSI performed quality control checks in response to high levels of missing data for this measure in the initial stage of the program's implementation.

These results may also suggest variations in the types of health status metrics collected during office visits. For example, one plausible explanation for the high levels of missing data regarding BMI is that height or weight were not captured at every visit, whereas blood pressure may have been a standard metric that was recorded at every visit.

Finally, the missing data may also corroborate a finding from the qualitative study. Partner staff (in particular care managers) anecdotally noticed that participants with more severe symptoms and greater social barriers (for example, transportation, isolation, and so on) were more difficult to engage. The attrition seen over time may reflect these difficult-to-reach participants self-selecting out of the program, contributing to the missing data.

A second limitation of the study is the absence of a comparison group. As a result, we have no information about changes in the outcome variables that might have occurred if participants had not been enrolled in the COMPASS program. For example, health outcomes might have

improved in the absence of the program, with standard care. Due to the lack of a comparison group, all findings from this analysis should be interpreted with caution. Further research (for example, random assignment or quasi-experimental design) is needed to fully examine outcomes for the COMPASS program.

5. Qualitative findings on perceived effects of the program

Findings from qualitative interviews support some of the results from the analyses of EHR data. Qualitative data also provide insights into other possible program effects that COMPASS staff anecdotally observed and discussed during site visits and interviews. Program staff, particularly care managers, mentioned several ways that COMPASS may have improved participant health outcomes:

- **Investment in care and care competence.** Most staff members reported that participants became more interested and invested in their own care and gained knowledge about their health. For example, staff reported that COMPASS participants did the following:
  - Became more compliant with care plans
  - Increased self-care activities, including improving their diet and working to get enough sleep
  - Increased self-monitoring of blood pressure and HbA1c levels

Staff believed that COMPASS participants also became more engaged in their treatment. One physician noted, "They've become more engaged with their primary [care providers], although most of the engagement has been with the care managers and the care managers let the primary [care providers] know what's going on. But [participants have] become more engaged and compliant with testing and laboratory work [and] with medication adherence."

This anecdotal improvement in participants' investment in and competence regarding their own care may explain some of the quantitative findings described in the previous section. Increased compliance, self-care, and monitoring could directly or indirectly lead to improvements in the health status variables. For example, increased compliance with a blood pressure medication regimen coupled with more frequent self-monitoring of blood pressure could contribute to the reduction in hypertension found through the quantitative analysis.

However, staff perceptions of changes in diet and self-monitoring of HbA1c levels were not reflected in our quantitative findings. Compared to other health status measures (for example, hypertension), blood glucose did not reach recommended levels for as many patients at follow-up. This measure may be harder to change in a year. If staff reports of participants' investment and competence in controlling their weight and diabetes are accurate, improvements may be visible over a longer follow-up period. Quantitative analysis of changes in BMI was not possible due to missing data.

- **Social outcomes.** Many staff members noted that social functioning among participants improved. For many participants, this included starting or returning to work. Some physicians and care managers reported that participants were less socially isolated, spending

more time with family and friends. These observations may also contribute to the improvement in depression symptoms at follow-up.

- **Physical health.** Anecdotally, care managers and physicians noted some changes in the physical health of participants. These included decreases in HbA1c levels, improvements in blood pressure, and lower cholesterol. A few staff members noted a reduction in some participants' weight and body fat. Our quantitative findings supported staff members' observations about blood pressure. Over 50 percent of those participants who were hypertensive at enrollment improved during the follow-up period. Due to missing data, it was not possible to corroborate staff reports of BMI and LDL reductions. Although some participants' HbA1c levels did improve, the degree of improvement was not as large as the other health status measures (that is, depression and blood pressure). One staff member reported, "Their blood pressures in general have come down. We've noticed a drop in A1c—although not marked because that really takes a long time to change. Their LDLs—marked drops."
- **Mental health.** Staff also reported decreases in some participants' PHQ-9 scores, indicating that their depression had improved. One staff member said one of her participant's PHQ-9 scores decreased markedly while she was involved in COMPASS. This respondent stated that the participant's PHQ-9 score "dropped about 10 points over that six- to eight-month period." Our analysis of EHR data supported this qualitative finding: over 60 percent of those with moderate to severe depression at enrollment improved at follow-up. Analysis of participants' PHQ-9 scores revealed an average drop of 7.5 points 12 months after enrollment.

Qualitative data also suggest improvements may have occurred in outcomes not captured in the EHR (for example, social history). Staff members we interviewed were careful to note that not all COMPASS participants experienced improvements, adding that improvements can take a long time. However, they pointed out that positive changes in one domain may lead to positive changes in another. For example, one care manager pointed out that minor improvements in depression may lead to improvements in social functioning and self-care, which in turn may lead to improvements in physical health.

Care managers also described a roller coaster–like effect of the program. They said some participants were initially engaged but their commitment and adherence to their care plan diminished with time. Then, about one year after enrollment, care managers again observed participant improvement in some of the domains. One care manager described the roller coaster in the following way:

> We have several [COMPASS] participants . . . that we've been working with for at least a year and really . . . no progress, nothing happening. And then month 13, 14, 15, you see the participant start to turn around and start to take an interest and actually start to follow up on all this teaching you've been doing. And all of this working you've been doing with them over the past year now responds. Now all of a sudden, they're starting to check their blood sugars, they're starting to take their medications, they're coming into appointments, they're following up with resources that you have given them over and over again, and it finally clicks for

> some reason. And [you realize] this participant is going to do okay because they're finally getting it.

D. Findings about the workforce

We synthesized quantitative and qualitative data to examine how ICSI trained and deployed its workforce. We drew on two main sources of data:

- **Quantitative data from our workforce survey** conducted in the spring of 2014 and 2015 to gather information about staff burnout and stress, job satisfaction, and perceptions of training and support
- **Qualitative data from staff interviews and focus groups** conducted during site visits to awardees in the spring of 2014 and 2015

The survey respondents and the interviewees on our site visits were drawn from the same workforce groups. However, due to staff turnover and the limited number of staff we could meet with during site visits, not every staff member participated in both data collection efforts.

Satisfaction. Many staff members said that they had a sense of personal accomplishment because they saw improvements in their participants, they gained and shared knowledge from other members of the care team, and they played a part in changing their organization's culture. The following common themes related to job satisfaction emerged from both site visits:

- **Staff perceived the COMPASS program as contributing to better health and quality of life for participants.** Many staff members described improvements that they observed in the COMPASS participants. Staff noted changes in the physical, mental, and social health of participants. This gave staff a sense of accomplishment and satisfaction. Some staff members even said that participants referred family and friends to the COMPASS program, underscoring the value of the program. Even small changes were satisfying for staff members.
- **The COMPASS program helped physicians use their time with participants more efficiently.** Many physicians and care managers reported that the COMPASS program reduced workloads for primary care providers. These care managers and physicians noted that the care team structure allowed care managers, who may be better equipped to handle overlapping psychiatric and medical issues, to spend a significant amount of time with participants with more complex conditions. In turn, primary care providers could focus on the specific medical needs of participants in the limited time they had to spend with them. One care manager said, "I think it has lightened the load of the primary care providers that have a limited amount of time to spend with some pretty complex medical and psychiatric issues, and [don't] have the time or maybe even the training to do both pieces."
- **Systematic case reviews made knowledge transfer and growth among the care team easier.** Part of the COMPASS model was a systematic case review involving the care manager, the primary care provider, and several consultants (usually a psychiatrist and another physician, although this varied from one partner to the next). The purpose was to review the treatment plans of each participant and track progress and outcomes. Many staff members discussed the knowledge that was gained and shared through systematic case

reviews as an added benefit of the COMPASS program. Many explained that this knowledge transfer allowed them to grow personally and professionally. One staff member summarized the growth this way:

> I think they were surprised by the knowledge transfer that was occurring through the consultant, especially through the psychiatric consultant, and how valuable that was. I don't think that was an expected gain. I think they were thinking they were going to tap into the psychiatrist for some direction, but . . . there was also some knowledge transfer, in that the clinicians that were involved in that are [now] more comfortable to manage [the participant's] depression on their own.

Support. Qualitative analysis of site visit interviews and findings from the workforce survey suggest that most members of ICSI's workforce were generally satisfied with their jobs and felt adequately supported at work. About 19 percent of survey respondents said that they did not have enough support in 2015 (towards the end of the award period), but only a few staff discussed this subject during site visit interviews. These staff members elaborated that they did not feel supported by their partner organizations' managers and senior leaders.

During the first site visit, this lack of support was brought up by only a few of the people we interviewed. Staff at two partner organizations said that they struggled to find support from their mid-level managers and the senior leaders at the partner level. These staff attributed this to a lack of buy-in from senior leaders or a lack of support for the organizational change involved with the COMPASS program. By the time of our second site visit, some staff members reported that manager turnover and changes in leadership had alleviated this situation. At two of the partner organizations, the attitudes about lack of support changed over time. Staff reported that some managers and leaders left the organizations and were replaced with individuals who believed in and supported the COMPASS model. One staff member noted the importance of having a good mentor and coach: "Without that, it's really difficult for the care manager. And one of our systems went through a number of changes with their leadership. They've recently put in a supervisor to do that mentoring, and they're seeing in that short period . . . the benefits of that." As noted above, supportive leadership was identified as key to successfully implementing and sustaining the COMPASS program.

Barriers to job satisfaction, job performance, and implementation. Although most of the staff we interviewed were satisfied with their jobs, they were dissatisfied with some aspects of the program. This dissatisfaction mostly centered on the program's EHR, the resources available to staff, and participants' level of commitment to the program. Many care managers thought the EHR often required extra work. This generally took the form of double entry into the COMPASS EHR and the partner organization's own system. One care manager said, "There wasn't a way for an electronic medical record to speak to the [COMPASS] registry [EHR]. They don't talk to each other so that caused entry work like you wouldn't believe. So that takes a lot of time and energy . . . away from the participant."

Some care managers also expressed dissatisfaction with the level of participant engagement and the resources that they had to help engage participants in care plans. Care managers

frequently pointed out that the complex social needs of their participants (that is, housing, transportation, and so on) often prevented them from engaging with the program. One care manager noted, "We didn't realize . . . how [many] psychosocial [needs] we'd be dealing with, and I think we found very early that we did not have enough social workers to manage the load." Some care managers said they would like more training on how to address these needs, how to locate and use available community resources, and how to engage participants.

In 2015, most staff thought the program succeeded in developing effective care teams, which was a notable change from the findings in our 2014 interviews. In 2014, many staff members at all levels reported fully integrated and successful care teams, but some did not. The staff who did not report fully integrated and successful care teams attributed this to the significant cultural shift involved in implementing the COMPASS program, which requires physician buy-in. The COMPASS program employs a nonhierarchical relationship among care team members as well as with the participant. This meant a shift in the culture at some partner organizations. One staff member described it as follows: "There is a nonhierarchical relationship, where the participant and care manager, caregiver, whatever role, is there in a true partnership. And we initially trained teams in behavioral activation, motivational interviewing, but that [is a] shift of paradigm for anybody in health care, to try and get on that equal footing with the participant." Some staff members reported that physicians in particular had trouble embracing the COMPASS paradigm. As a result, they did not buy into the program. One staff member said, "This was a challenge because they are an independent practice association and so they're very disparate. And the docs don't necessarily see themselves as part of [the partner organization]. They see themselves as the owner of their private clinic. And so [their attitude is], why would I want to engage in this work?"

By the time of our 2015 site visits, however, many staff reported a change in organizational culture and a shift in physicians' acceptance of the team-based model. Care managers thought this happened because physicians saw improvements in their participants and benefited from the knowledge sharing in the COMPASS process. As a result, the care teams became more effective, particularly in regard to the physician–care manager relationship. One care manager recalled a story about a primary care provider changing his view on the program after learning more about a participant who had been in his care for the past 20 years. The care manager reported:

> He'd been caring for this person, been struggling with their diabetes, I think was the big thing. . . . And this care manager came to him and said, "This participant can't read." In 20 years, he didn't know that. And so that was very powerful to him, that the care manager could spend more time, was providing value . . . and it totally increased his trust in her. And so then it became much easier for him to trust her with other things.

Training. Both new and existing staff received the same training. The primary focus of the training was to (1) explain the COMPASS model and use of the EHR; (2) describe the COMPASS workflow; and (3) teach care management and other clinical techniques (for example, motivational interviewing). Staff were trained to understand the program's goals, develop their knowledge of team-based care processes, and enhance their care management and

other clinical skills. ICSI tailored the training to each partner organization based on the staff members' skills and needs (as identified by readiness assessments).

Most staff found the trainings helpful. In response to the 2015 survey, the majority of staff (72 percent) indicated that the trainings they received in the past 12 months were either very or somewhat useful for their work; 11 percent reported that the trainings were not very useful. This pattern was similar to the one found in the 2014 survey results, and was consistent with feedback from staff we interviewed in both years.

Many staff identified the need for more training and education. Although most staff members found their training useful, many care managers suggested improvements, including additional trainings on behavioral and mental health issues (for example, the complex social needs of participants, the community resources available to participants, strategies for participant engagement, and treating depression and physical health issues). The COMPASS intervention does not involve any direct care from a behavioral or mental health specialist to treat depression or other comorbid disorders. Consulting psychiatrists assist in reviewing participants' treatment plans, medication management, and progress through the weekly systematic case reviews, but they do not provide direct therapeutic treatment (such as cognitive behavioral therapy) to participants. Staff requests for more training on mental health treatment may stem from the perception that participants' behavioral and mental health problems should be addressed more directly.

E. Program sustainability and spread

Our 2016 follow-up interview with ICSI leaders revealed that over 80 percent of the ICSI partners had retained some component of the COMPASS program. Only two partners chose not to continue any aspect of the COMPASS program. For one of these partners, not sustaining the COMPASS program was purely a financial issue. This was a smaller partner with one care manager and over 100 potential participants. The efforts to sustain the program would have outweighed the benefits. The other partner chose not to continue the COMPASS program in favor of another model of care.

Systematic case reviews and the care team approach were the most widely sustained components of the COMPASS program. In our latest interview, ICSI leaders reported that the majority of sites still utilized them and had expanded the systematic case review teams to include cardiologists and pharmacists. Many partners liked the access to consulting medical professionals (particularly psychiatrists) and the more structured COMPASS care management tools (for example, guidelines about care structure and follow-up), in addition to the rigorous care management approach. Furthermore, about half of the partners had expanded their participant eligibility conditions. Many of them included other specific chronic diseases, most commonly chronic obstructive pulmonary disease. Some also expanded the criteria to include participants with more complex needs. For example, one partner clinic was focusing on people who had at least four medical conditions or a recent discharge from the hospital.

ICSI leaders also discussed the spread of the COMPASS program. They reported that a physician from one of their partner organizations had started a business to teach others how to implement the COMPASS program. ICSI leaders further reported that some partners expanded the COMPASS model to other departments or clinics in their health care systems. ICSI leaders

also felt that the COMPASS program impacted the way that partner organizations treated depression. They perceived the COMPASS program as helping to shift organizational cultures to more of a complex care management—away from depression care that had been focused on keeping participants on antidepressants for the first few months of treatment. ICSI leaders felt that organizations changed their approach to one where depression care managers were focused on trying to keep participants engaged for a longer period of time and thinking about their medication adherence overall.

Despite these promising steps, ICSI leaders noted several continuing challenges regarding sustainability:

- **Funding.** Reimbursement for care management continued to be a challenge for the COMPASS program's sustainability. The current FFS environment does not often cover care management; therefore, partners have had to find different ways to fund this component. For example, one partner incorporated the program into its affordable care organization (ACO) infrastructure; another partner incorporated the program into its program for complex participants (which covers COMPASS-related expenses). This was also supported by our qualitative site visit data. Many of the partners noted that the traditional FFS model did not have payment systems for COMPASS services, particularly for care management time. This was a significant barrier to the program's sustainability.

  Senior leaders at ICSI discussed changing the FFS model to one in which the cost of care is considered more broadly in the context of a person's entire life course. For example, the cost of care may increase in a given time period in a person's life, but that care may lead to decreased costs later and potentially extend the individual's life. One leader explained:

  > We need to figure out . . . a way to pay for care management and for the time of . . . a psychiatrist to participate. And this is not something that we're really rewarded for today. We still get paid mostly for . . . volume-based stuff. But we hope we can continue to . . . grow our total cost of care, risk-based, ACO-type payment model, so we can try on more and more of these types of programs and not worry so much about the impact of financial [constraints].

  Senior leaders at ICSI conducted a training for partner organizations on the total cost of care model after the award ended in 2015.

- **Leadership change.** Leadership turnover has also reduced COMPASS sustainability. ICSI leaders perceived difficulties in getting new leaders up to speed about a program, particularly regarding findings from the award period.
- **Psychiatry access.** ICSI leaders noted a difficulty in access to psychiatrists. The COMPASS program requires regular access to psychiatrists through the systematic case review process. Several partners lost their psychiatrists. One partner dealt with this by contracting with a service that provides tele-psychiatry. Another partner had to shut down its COMPASS program for several months while it attempted to fill this role.

F. Lessons learned

ICSI implemented the COMPASS program to treat participants with depression and uncontrolled diabetes or cardiovascular disease across a diverse array of partner organizations. The evaluation provides several lessons learned for those wishing to implement similar programs:

1. Outcomes

The COMPASS program appears to have been successful in improving some health status outcomes for many of the participants**,** most notably depression and hypertension after 12 months of treatment. Diabetes showed some improvements, but the rate of improvement was not as high as those for the depression and hypertension health status measures. HbA1c levels after a year in the program were still uncontrolled for 60 percent of the sample. Outcomes were also comparable to prior research on a similar program.

The rates of hospitalizations and ED visits significantly decreased through the COMPASS program period, but expenditures did not. These results should be interpreted with caution because comparison groups were not available for the analyses, so we cannot determine the extent to which results differed from what would have happened in the absence of the intervention. Missing data also limited our analyses of the EHR data. To firmly establish the effectiveness of the COMPASS program in these sites, future evaluations should strive to remedy these methodological limitations.

2. Implementation

Organizational and physician buy-in and overall support at the partner level are critical to implementation success. The COMPASS program involved a considerable shift in organizational and physician culture in terms of participant care. Both leaders at ICSI and staff across all levels (care managers, physicians, and so on) at the partner organizations repeatedly listed this as the main ingredient for success. In the workforce survey, physician resistance to the COMPASS program was also one of the more common barriers to effective job performance that staff reported. ICSI leaders attempted to facilitate culture change through a collaborative approach with partners and through use of effective communication strategies, which they felt increased buy-in, accountability, and sharing of best practices and challenges.

The program EHR was seen as burdensome by some COMPASS staff members. This was evidenced by the level of missing data in the EHR, which hindered our analysis of the data. Having more consistent data entry guidelines and monitoring data entry may have solved this problem. Challenges may have also been addressed by exploring ways to integrate the partners' existing EHR systems to reduce the burden of duplicate data entry.

Participants with complex social needs were difficult to engage in the program. Attrition may have limited our analysis of the EHR data. More training and support may be needed to educate care managers about effective strategies for engaging these participants.

3. Sustainability and spread

Many of the partners have sustained at least some components of the COMPASS program. The systematic case reviews and care teams were the most widely sustained aspects of the program. Other components sustained included (1) the application of the COMPASS model to other departments or clinical sites within partners' health systems and (2) the application to populations. Future research may focus on evaluating their success.

The current FFS payment structure is a major challenge for sustainability because it does not facilitate reimbursement for care management. Programs that have sustained these components have needed to search for funding outside the traditional system. ICSI leaders suggested that changing the payment climate from FFS to total cost of care could help fund the COMPASS model. Having consistent leadership that supported the program was also essential for sustainability. This factor, coupled with organizational support, was key for finding additional funding to sustain the COMPASS program.

VIII. KITSAP MENTAL HEALTH SERVICES (KMHS)

Findings from Mathematica's Evaluation of the KMHS HCIA program

- We conducted a comprehensive evaluation of the KMHS Race to Health! program, including analyses of Medicare, Medicaid, and electronic health record (EHR) data. We also analyzed information we obtained from site visits, telephone interviews, focus groups, and a workforce survey.
- Quantitative analyses suggests that Race to Health! may have successfully cut down on emergency department (ED) visits, hospitalizations, and total expenditures for some KMHS Medicare patients. Race to Health! also seems to have helped reduce the number of hospitalizations for Medicaid patients; however, the program was associated with a higher number of ED visits for Medicaid patients. Quantitative findings also indicate the program was moderately successful in conducting physical health screening, targeting screening resources to patients with pressing needs, and improving patients' results on health status measures.
- KMHS staff attributed improved outcomes under Race to Health! in part to staff's increased focus on and understanding of physical health, as well as their use of patients' physical health data. In interviews and focus groups, respondents also talked about the new features of the program, which emphasized patients' wellness and self-management of health conditions, and credited these features with helping patients make important progress with their health.
- Race to Health! required a significant training effort, because most staff had no background in physical health care coordination or integration before the program began. Most of them said they felt supported in their roles, and believed the program's strong infrastructure provided a solid foundation for their work.
- As of April 2016, 10 months after HCIA funding ended, KMHS was sustaining all components of Race to Health! but expected to face challenges in sustaining some of these components in the future.

A. Introduction

Staff at Kitsap Mental Health Services, a community mental health center in Kitsap County, Washington, used HCIA funding to implement Race to Health! The program was designed to improve behavioral and physical health care and outcomes and, concurrently, reduce the cost of care for all adults and children receiving outpatient services at KMHS beginning January 1, 2013. HCIA funding for Race to Health! ended on June 30, 2015.

Race to Health! had two primary components:

- **Whole-health focus within KMHS.** Race to Health! was an organization-wide initiative to redesign KMHS' infrastructure and service delivery model and prepare staff to address clients' whole health (that is, mental and physical health and substance use). Before the award, KMHS had reorganized its staff into multidisciplinary care teams to better integrate substance abuse treatment into patients' care plans and improve coordination between KMHS staff and patients' physical health care providers. The HCIA funding made it possible for KMHS' care teams to receive training on substance use and physical health conditions and strategies for supporting clients' self-management of chronic diseases such as diabetes. In addition, KMHS expanded its electronic health records (EHR) system to include data on physical health, and hired new staff (medical assistants and healthy family

coordinators) to collect and monitor these data. Care teams used the data to better understand the full range of the clients' needs for health services and to improve coordination of primary care and behavioral health services. KMHS also used these data to identify clients who would benefit from more care coordination with key community stakeholders (for example, social service providers, health plans, law enforcement, and emergency medical service staff).

- **Integration of behavioral health and primary care in community settings.** As part of Race to Health! KMHS partnered with a community health clinic, Harrison Health Partners (HHP), with KMHS staff providing brief behavioral health interventions and referrals at four HHP primary care practices. KMHS also offered telephone and email psychiatric consultations to HHP's primary care providers and other primary care providers in the community.

Our analyses drew on data from the following sources:

- Enrollment data submitted by KMHS to the reporting website maintained by CMMI's technical assistance contractor (the Lewin Group) for the HCIA Round 1 initiative.
- Medicare claims covering the period from July 2010– June 2015.
- Medicaid claims covering the period from July 2011–June 2014.
- Health status data from KMHS's EHRs, covering the period from January 2014–June 2015.[40]
- A workforce survey that we conducted in the spring of 2014 and 2015 to gather information about staff burnout and stress, job satisfaction, and perceptions of training and support.
- Qualitative data from telephone interviews and in-person site visits in the spring of 2014 and 2015. During these visits, we conducted in-depth interviews with key awardee staff, members of the workforce, and other stakeholders.

In this chapter, we present the results of our evaluation to date. First, we review the administrative context of KMHS' program, discuss the stage the program has reached in the innovation life cycle, and describe the characteristics of the program's patients. In Section B, we discuss the methods we used to analyze the quantitative data and collect the qualitative data. In Section C, we present findings from our descriptive and impact analyses of the Medicare, Medicaid, and EHR data and findings from qualitative data about effects as perceived by staff and stakeholders. In Section D, we report key findings about the KMHS workforce based on qualitative interviews and the workforce survey, and in Section E, we report on KMHS' plans to sustain the program. We close with an overall summary of lessons learned in Section F. We will present additional analyses in an addendum to this report.

[40] Collection of health status data in KMHS's EHR ramped up over time. We decided to begin the analysis period of this data in January 2014 because we assumed the data from that point forward would be more complete and collected more consistently than earlier data.

1. Overview of administrative context

KMHS is a private, nonprofit community mental health center and the sole provider of public mental health services in Kitsap County, Washington. KMHS offers inpatient and outpatient mental health and substance abuse services to individuals at all stages of life. Under Race to Health!, KMHS implemented a bidirectional model of care, bringing physical health coordination and wellness programming into its outpatient program, and bringing behavioral health intervention and consultation into primary care settings in the community.

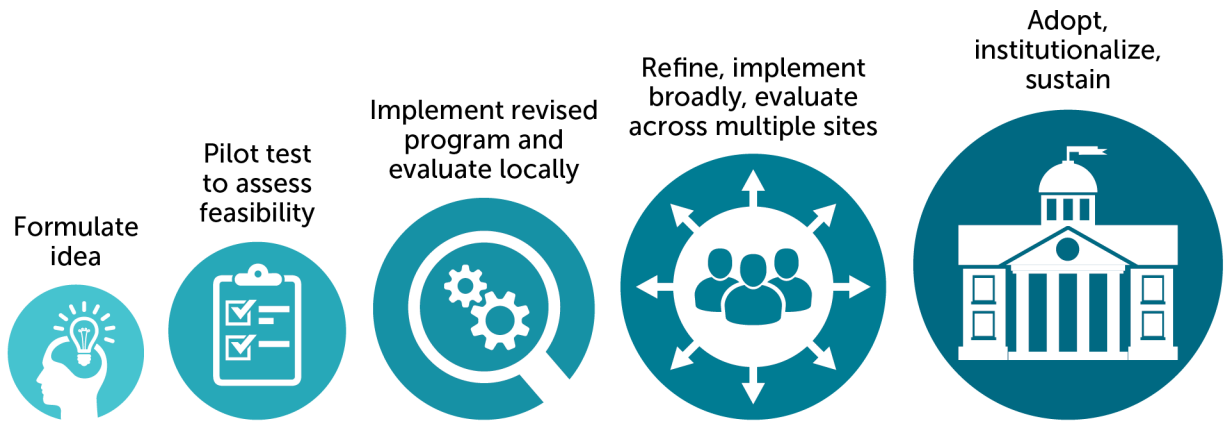
Race to Health! benefited from strong leadership buy-in within KMHS and from partner organizations. For example, KMHS's chief executive officer, administrative director, and other senior leaders led the development and implementation of the program. KMHS also benefited from strong internal staff capacity and well-established relationships with institutions in the community. For example, the agency used existing staff to implement some components of the program, both in administrative and frontline roles, and built on its existing relationship with HHP for the community-facing components.

2. Progression through phases of innovation

Race to Health! was an agency-wide transformation initiative that grew out of and built upon KMHS's previous work to promote coordinated and integrated care. Prompted by an enduring commitment to integrate mental health and substance abuse services and coordinate these services with primary care and other community providers, KMHS' leaders viewed the HCIA funding as an opportunity to carry forward specific plans to achieve long-term goals for service integration. As one program leader noted, "It was a model we wanted to do, whether we got [HCIA funding] or not."

KMHS leaders planned and prepared for the agency transformation before they received HCIA funding, then refined the program's goals and strategies during the award period. For example, before receiving the award, the agency restructured its existing staff into multidisciplinary care teams and later physically moved staff within the buildings to make collaboration easier. Because KMHS had already begun to introduce the concept of coordinated and integrated care before the HCIA, agency leaders and staff were well prepared to implement the program model after award, assess program functioning, continually adapt the model as part of a continuous improvement process, and ultimately institutionalize Race to Health! within the agency—Phases 1 through 5 in our conceptualization of innovation phases (Figure VIII.1).

Figure VIII.1. Phases of program innovation: KMHS



The awardee refined the Race to Health! program model throughout the award period by researching and selecting different evidence-based practices, curriculums, and program models and then tailoring their approach to the goals of the agency. For example, KMHS hired a healthy living program developer to identify and roll out wellness programming such as Living Well and the Stanford Chronic Disease Self-Management Program. Likewise, the agency's internal consultant on substance abuse helped KMHS identify and adapt the right screening and treatment for substance use disorder. The healthy living program developer worked with the healthy family coordinator to modify screening programs on wellness and substance use disorder for children's and families' care teams.

By design, KMHS continuously altered or improved the program components based on program leaders' observations and staff feedback. For example, program leaders learned lessons from staff's initial use of patients' physical health data to inform how they cared for the patients' mental health conditions. One program leader described the agency's early experiences with health data as "trying to drink from a fire hose. We flipped it on and we just got swamped." Over time, KMHS program staff initiated protocols and criteria to identify useful information on physical health and promote its use by care teams in ways that allowed "data to become information that can inform the treatment plan."

Program leaders worked throughout Race to Health!'s implementation to ensure the program's sustainability by participating in forums and networks outside the formal scope of the program. For example, senior KMHS staff participated in a network board that advises the state about its move toward integration of behavioral, physical, and substance use treatment. As discussed in Section F, KMHS institutionalized many aspects of the program after funding ended, and was working with partners to sustain Race to Health! activities in the community.

3. Enrollment

KMHS did not define a specific enrollment target for Race to Health! because the program was designed to reach everyone who used KMHS' outpatient services. For our evaluation, and to comport with KMHS' implementation procedures, we define the target population as all KMHS

patients who received face-to-face outpatient services between January 1, 2013, and June 30, 2015—a total of 6,662 patients.

Although all KMHS patients who received outpatient services were included in the target population, the awardee periodically identified subgroups of patients, known as cohorts, with more severe health conditions. KMHS staff chose the adult cohorts by using information from the state's PRISM data system and KMHS' EHR. PRISM is a web-based application that integrates data on Medicaid enrollees from multiple sources and provides risk assessment tools such as the chronic disability illness system, which assigns risk scores to Medicaid enrollees based on the severity of their health care needs. For the children's cohorts, staff asked providers for recommendations and then analyzed EHR data to search for comorbidities.

Staff made a special effort to ensure that the EHRs for members of the cohorts contained key data about their health status and use of health services. KMHS staff obtained this information by reaching out to patients' primary care providers and tapping into an Emergency Department Information Exchange data system that is available to providers in the state. KMHS used these more comprehensive EHR data to better understand the full range of patients' needs for health services and to improve coordination with primary care providers.

4. Patients' demographic characteristics

Because we had no single data source that covered the entire KMHS patient population and the program's targeted outcomes, we conducted three sets of quantitative analyses based on different data sources (Medicare, Medicaid, and EHR). These sets of analyses may produce different findings because they include different populations.

The characteristics of the patients in the three data sets differ substantially (Table VIII.1). The Medicare analysis includes individuals who were only enrolled in Medicare only (22.3 percent) and those concurrently enrolled in both Medicare and Medicaid (77.7 percent). The Medicaid analysis includes only individuals enrolled solely in Medicaid. The analysis of health status measures includes individuals in all three groups: Medicaid only (73.9 percent), Medicare only (2.4 percent), and dual enrollees (22.4 percent).

Children were not included in the analysis of Medicare and health status measures, but they make up 36.2 percent of the Medicaid analysis population. No individuals over age 64 are included in the Medicaid analysis, and only 4.5 percent of the individuals in the health status measure analysis are over 64. In contrast, 30.6 percent of the Medicare analysis population is over 64. The majority of Medicaid, Medicare, and EHR analysis populations are female (56.2, 54.4, and 57.0 percent, respectively).

Over three-quarters (77.9 percent) of the Medicare sample were eligible for Medicare because of a disability, but less than half (42.9 percent) of the Medicaid sample was eligible due to a disability. The Medicare and Medicaid groups also differ in the prevalence of specific mental health diagnoses (Table VIII.2). For example, 31 percent of the Medicare or dual enrollees have disorders related to schizophrenia, whereas among Medicaid enrollees, this figure is much lower (6 percent).

Table VIII.1. Demographic characteristics of groups included in analyses

	Medicare analysis		Medicaid analysis		EHR analysis (health status)	
	Number of patients	Percent of patients	Number of patients	Percent of patients	Number of patients	Percent of patients
Total population	846	100%	3,776	100%	2,640	100%
Medicaid, non-dual	0	0.0	3,776	100.0	1,950	73.9
Medicare, non-dual	189	22.3	0	0.0	64	2.4
Dual	657	77.7	0	0.0	592	22.4
Unknown	0	0.0	0	0.0	34	1.3
Age						
Less than 18	0.0	0.0	1,368	36.2	0	0
18–34	138	16.3	1,188	31.5	1,147	43.4
35–44	142	16.8	520	13.8	523	19.8
45–54	173	20.5	460	12.2	546	20.7
55–64	134	15.8	240	6.3	305	11.6
65 or older	259	30.6	n/a	n/a	119	4.5
Gender						
Female	460	54.4	2,121	56.2	1,504	57.0
Male	386	45.6	1,655	43.8	1,136	43.0
Medicaid/Medicare eligible based on disability						
Yes	659	77.9	1,620	42.9	n/a	n/a

Source: The Medicare analysis is based on Medicare administrative data for July 2010–June 2015. The Medicaid analysis is based on Mathematica analysis of Medicaid Analytic Extract (MAX) and Alpha-MAX data for Washington State for July 2011–June 2014. The analysis of health status measures is based on EHR data for January 2014–June 2015, provided by KMHS.

Note: The Medicare analysis is limited to individuals who were not enrolled in Medicare Advantage, had Medicare as a primary payer and were enrolled in parts A & B, and received mental health treatment at KMHS or a comparison facility. The Medicaid analysis is limited to individuals with full benefits who were enrolled in Medicaid (with Medicaid as the first payer, and not dually enrolled in Medicare) for at least 6 months after beginning mental health treatment at a KMHS or other facility between the beginning of the program period and June 2014. The Medicaid analysis also excluded S-CHIP enrollees and individuals with missing enrollment records. The EHR analysis is limited to KMHS patients who had at least one face-to-face visit in 2014.

Table VIII.2. Diagnoses of Medicare and Medicaid samples

	Medicare analysis		Medicaid analysis	
	Number of patients	Percent of patients	Number of patients	Percent of patients
Diabetes	n.a.	n.a.	226	6.0
Hypertension	n.a.	n.a.	427	11.3
Drug abuse diagnosis	n.a.	n.a.	280	7.4
Alcohol abuse diagnosis	n.a.	n.a.	282	7.5
Schizophrenic disorder	260	30.7	225	6.0
Bipolar disorder	175	20.7	483	12.8
Depressive disorder	186	22.0	841	22.3
Dementia	76	9.0	<11	<11
Other psychotic disorder	55	6.5	96	2.5
Anxiety, dissociative, and somatoform disorder	20	2.4	546	14.5
Adjustment reaction disorder	65	7.7	759	20.1
Other mental health diagnosis	37	4.4	1,112	29.5

Source: The Medicare analysis is based on Medicare administrative data for July 2010–June 2015. The Medicaid analysis is based on Mathematica analysis of MAX and Alpha-MAX data for Washington State for July 2011–June 2014.

Note: Psychiatric diagnoses were not available for the EHR group. The psychiatric diagnosis indicators (schizophrenic disorder through other mental health diagnosis) for the Medicare analysis were created using ICD-9 diagnosis codes found on any of the patient's psychiatric services claims in the month during the program period in which the patient was first attributed to a facility and the two months following. The diagnosis indicators for primary care and substance use (diabetes, hypertension, drugs, alcohol) for the Medicaid analysis were created using ICD-9 diagnosis codes found on any of the patient's claims in the 12 months before the month during the program period in which the patient first had a psychiatric services claim. The psychiatric diagnosis indicators (schizophrenic disorder through other mental health diagnosis) for the Medicaid analysis were created using ICD-9 diagnosis codes found on any of the patient's psychiatric services claims in the month during the program period in which the patient first had a psychiatric services claim and the two months following.

n.a. = not applicable

B. Methods

1. Quantitative methods

In this section, we describe the methods we used for each of the three sets of quantitative analyses.

Medicare impact analysis. For the Medicare sample, we were able to conduct a rigorous impact analysis using a difference-in-differences time series model with a matched comparison group. This analysis focused on four outcome measures:

1. Total expenditures
2. Hospitalizations
3. ED visits

4. Office visits[41]

Hospital readmissions were too infrequent to produce reliable impact estimates.

Office visits may serve as a measure of the extent to which KMHS altered their patients' use of preventive and well care services. KMHS specifically endeavored to reduce use of acute care services by monitoring patients' physical health more often, promoting use of preventive care services, and encouraging better self-care; any reductions in acute care may lead to corresponding reductions in expenditures.

We obtained Medicare data from the CMS Virtual Research Data Center. We used Medicare data covering the period from July 2009 through June 2015.[42] We included claims for anyone with a Medicare claim for an outpatient mental health visit at KMHS between July 2010 and June 2015 in the dataset.

The pool of potential comparison group members included anyone with a claim for an outpatient mental health visit at a comparison mental health facility or a facility serving clients with dementia in the state of Washington. We used the Substance Abuse and Mental Health Services Administration's Treatment Finder to identify 16 mental health facilities in Washington State with characteristics similar to those of KMHS. Because these 16 facilities served a limited number of patients with dementia and thus provided an insufficient pool of comparison patients to whom KMHS patients with dementia could be matched, we also identified facilities in Washington that had at least 100 beneficiaries with Medicare administrative claims for dementia and included all patients with dementia from these facilities in the potential comparison pool.

Both the intervention group and the potential comparison group were limited to patients who had Medicare as their primary payer, were enrolled in Medicare Parts A and B, and were not enrolled in Medicare Advantage.

We used propensity score matching to select the final comparison group. We matched up to five members of the comparison pool to each KMHS patient. With the matching algorithm, we sought to identify comparison group members who resembled the members of the intervention group on several key characteristics that are predictive of future Medicare service use and expenditures, including demographics, disability status, Hierarchical Condition Categories (HCC), dual Medicare/Medicaid enrollment status, and mental health diagnoses. The standardized differences between the KMHS patients and the comparison group were within 10 percent for all measures included in the matching analysis, indicating a strong match. Appendix A includes more details on the data processing and matching methods.

Medicaid impact analysis. We were also able to conduct a rigorous impact analysis using a difference-in-differences time series model with a matched comparison group for the Medicaid population. The analysis includes Medicaid beneficiaries enrolled in both fee-for-service and

[41] Office visits are evaluation and management services, including preventive services or well care provided to a new or established patient in a physician's office, nursing home, or patient home.

[42] Data for July 2009–June 2010 were only used to identify chronic conditions among patients participating in mental health treatment at KMHS or a comparison facility in July 2010–June 2011.

managed care. Because of data limitations, the analysis for the Medicaid population focused on two outcome measures: hospitalizations and ED visits. We were unable to calculate total expenditures for the Medicaid population because many of the KMHS Medicaid-enrolled patients are enrolled in Medicaid managed care, and data on expenditures were unavailable for this population. In addition, we did not use hospital readmissions in our analyses because they were so rare that our estimates would have been unreliable.

We obtained the Medicaid administrative data for the analyses from the CMS Virtual Research Data Center. We used MAX and Alpha-MAX data for Washington State for the period from July 2011 through June 2014. Our analyses were limited to this period because managed care reporting for 2009 and 2010 was not comparable to the reporting for 2011 to 2014. Also, when the file for this analysis was developed in October 2015, the Alpha-MAX data included only claims paid through September 2014. These data cover a more limited period than our Medicare analysis; as a result, results from our Medicaid analyses should be considered preliminary.

To identify KMHS patients for this analysis, KMHS provided a finder file drawn from their EHR data, from which we identified individuals who had an in-person visit recorded in the EHR on or after January 1, 2013; to be included in the analytic file, individuals had to also have a Medicaid mental health service claim during 2013 or 2014. Unlike for the Medicare analysis, we could not identify facilities using the Medicaid data to create the comparison group. Instead, the comparison pool was all individuals not in the treatment group who had a mental health service claim in the state's Medicaid data during 2013 or 2014 and did not have an in-person visit at KMHS in the EHR data after January 1, 2011. Individuals who were not eligible for full Medicaid benefits, who did not have Medicaid as the first payer, who were dually enrolled in Medicare, who were S-CHIP enrollees, who had missing enrollment records, or who had less than six months of enrollment in Medicaid following initiation of treatment at KMHS or another facility during the program period through June 2014 were excluded from the analysis. The comparison population in this analysis therefore represents a broad population of individuals receiving mental health treatment from all types of providers throughout the state.

As for the Medicare analysis, we used propensity score matching to match up to five members of the comparison pool to each KMHS patient. With the matching algorithm, we sought to identify comparison group members who resembled the members of the intervention group on several key characteristics that are predictive of future Medicaid service use and expenditures, including demographics, disability status, Chronic Illness and Disability Payment System (CDPS) conditions,[43] and, to the extent feasible, the mental health diagnoses listed on the

[43] We calculated CDPS scores based on a risk adjustment model developed by the University of California, San Diego (USCD), which some Medicaid programs use to adjust payments for beneficiaries who are disabled or on Temporary Assistance for Needy Families. Scores reflect the ratio of predicted health expenditures for a given beneficiary relative to average Medicaid per-person expenditures. Each beneficiary's CDPS scores are estimated based on diagnoses in the past 12 months of Medicaid claims data, as well as demographic characteristics. We created CDPS scores following UCSD's CDPS + MRx methodology. They were based on the conditions reported in Medicaid claims data in the 12-month period before the month the enrollee first had a claim corresponding to the conditions focused on in the program's goals. For individuals who had a mental health visit in January 2014, the score was calculated based on Medicaid data for January through December 2013.

person's claims. Appendix A includes more details on the data processing and matching methods.

Analysis of health status measures. KMHS' EHR data files had a great deal of information about the services that staff provided to patients during the study period. We selected three measures that could indicate the effectiveness of the program: (1) blood pressure screening, (2) body mass index (BMI) screening, and (3) metabolic screening.[44] We were able to construct these measures with the EHR data, and each had strong evidence for clinical importance. These analyses are focused only on members of the intervention group. We describe the specific methods for calculating these measures below.

Blood pressure screening. One of the program's goals was to screen for and help patients control their blood pressure. To analyze the program's impact on blood pressure, we followed the criteria of the National Committee for Quality Assurance (NCQA) measure for controlling blood pressure.[45] We calculated the proportion of individuals who had an initial screening during 2014, and used reported blood pressure measurements to identify individuals with hypertension. Of those with an initial screening indicating hypertension during the first six months of 2014, we calculated the proportion who had (1) a follow-up blood pressure screening within 12 months of the initial screening and (2) blood pressure that was under control at the most recent measurement within that 12-month period.

BMI screening. We included BMI in our analyses because Race to Health! included an increased focus on promoting a healthy diet, nutrition, and exercise. The National Institutes of Health defines overweight as having a BMI of 25 to 29.9, and obese as a BMI of 30 or greater; BMIs in the overweight, obesity, and extreme obesity categories are associated with increased disease risk.[46] For our analyses, we identified individuals who had a BMI screening in 2014, and assigned each individual to one of the BMI categories based on this initial screening. For individuals who were overweight or obese during an initial screening in the first six months of 2014, we calculated the proportion who had a follow-up BMI screening within 12 months and who had a 5 percent or higher reduction in BMI relative to their baseline. We also analyzed effects on a continuous measure of BMI. Among patients who were overweight or obese at their initial screening in 2014, we calculated the mean BMI at the initial and follow-up screenings.

Metabolic screening for non-diabetic patients on antipsychotics. Patients taking antipsychotic medications are at higher risk of developing diabetes. As part of Race to Health!, KMHS medical assistants conducted metabolic screenings of individuals taking antipsychotic medications, with the goal of detecting indications of diabetes early. To assess the reliability with which screenings were completed, we used the NCQA measure for the metabolic screening rate for patients on antipsychotics. For each six-month analytic time period from January 2014 to

[44] Results of the analyses of other health status measures are in Appendix A.

[45] The American Heart Association (2015) defines high blood pressure generally as 140 systolic mmHg or higher, or 90 diastolic mmHg or higher. More specific definitions are also used as necessary, such as for comparisons of elderly adults to younger ones, or of individuals with diabetes and those without (AHA 2015; James et al. 2014). NCQA operationalized these criteria to create its measure for controlling high blood pressure.

[46] See https://www.nhlbi.nih.gov/health/educational/lose_wt/BMI/bmi_dis.htm.

June 2015, we calculated the number and percentage of non-diabetic patients who also (1) had a diagnosis of schizophrenia or bipolar disorder and (2) had a prescription for an antipsychotic medication during the measurement period. We then calculated the number and percentage of those patients who had a metabolic screening.

2. Qualitative methods

We collected qualitative data during site visits to the awardee in the spring of 2014 and 2015. During our spring 2014 site visit, we conducted in-depth interviews with key awardee staff, including staff in each position on the care team, as well as Race to Health! program leaders and staff members who provided data and operations support. We also conducted interviews with the behavioral health provider and psychiatric consultant who delivered behavioral health consultation and services at community primary care practices. In addition, we held telephone interviews with two of HHP's primary care physicians and two physicians from other primary care practices in the community who received HCIA-funded consultations. In total, we conducted 21 interviews with 42 individuals.

In spring 2015, we conducted a second round of site visit interviews with many of the same types of team members (but not necessarily the same people, because of staff turnover or changes in staff roles). We conducted 18 interviews with 39 individuals. We also convened a focus group with 12 staff members from KMHS and from community stakeholder organizations that collaborated on Race to Health! Finally, in April 2016, we conducted one follow-up telephone interview with KMHS leaders to discuss how they had worked to sustain the program since the end of the award in June 2015.

C. Summative findings

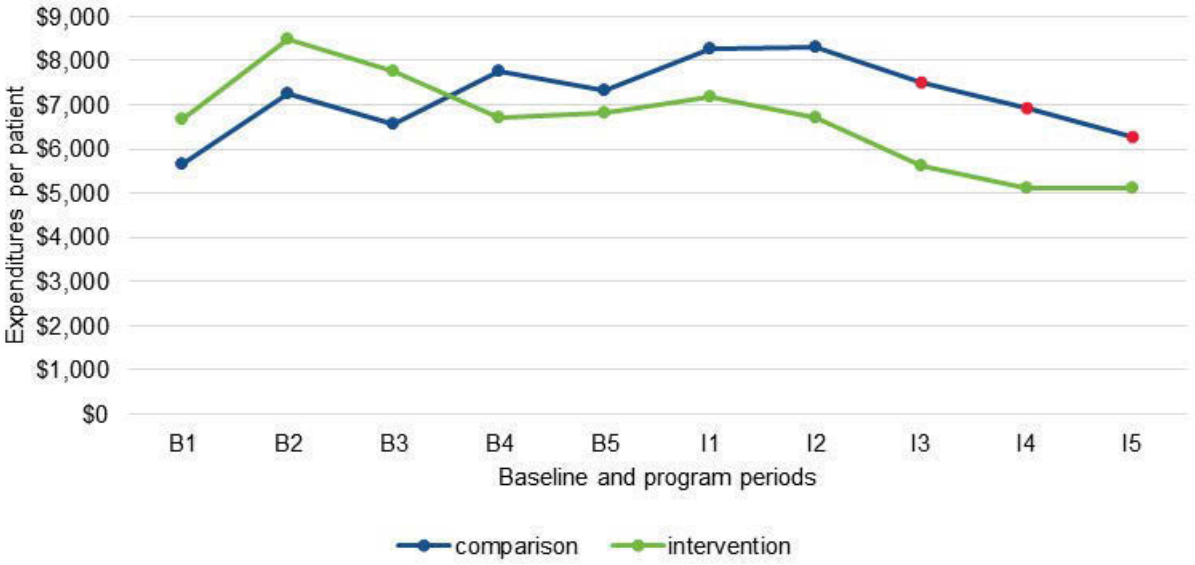
Below, we report the findings from our Medicare, Medicaid, and health status analyses.

1. Results of Medicare analyses

a. Descriptive analyses

Expenditures. During the baseline period, expenditures for the comparison group went up whereas expenditures for the intervention group went down (Figure VIII.2). Expenditures for both groups decreased after the intervention began, with expenditures for the comparison group remaining higher than those for the intervention group. From I3 to I5, the gap was significantly different from the average gap during the baseline period. This finding suggests that the intervention may have lowered expenditures for patients in the intervention group relative to the comparison group.

Figure VIII.2. Total Medicare expenditures per patient per 6-month period: Beginning of baseline to end of intervention

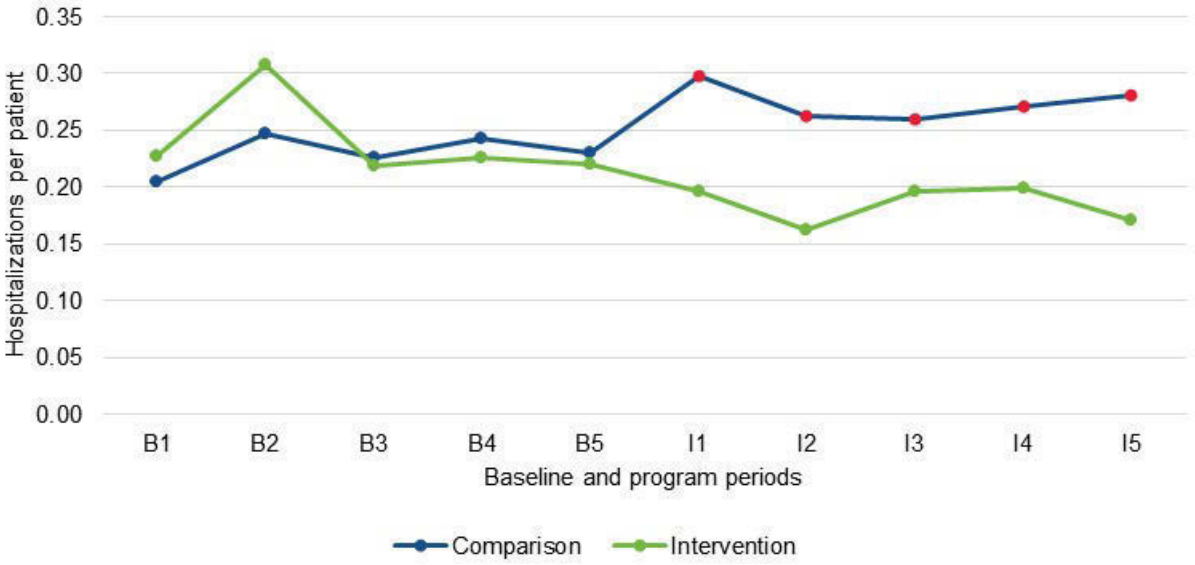


Source: Mathematica analysis of Medicare administrative data for July 2010–June 2015.

Note: Means are regression-adjusted. Red points indicate that the difference between the intervention and comparison group mean in the intervention period is significantly different from the average difference that occurred in the baseline period. B1, B2, etc., indicate each 6-month period of the analysis before the intervention began, and I1, I2, etc., indicate each 6-month period of the analysis during the intervention period; 823 intervention group members and 2,643 comparison group members were included in this analysis. Sample sizes varied from period to period, depending on data availability.

Hospitalizations. Other than at B2, the average difference in hospitalization rates between the intervention and comparison groups in the baseline period was negligible (about .01). Figure VIII.3 shows the extremely small gaps between the two lines for most of the baseline period. Presumably, in the absence of the intervention, the hospitalization rates of the two groups would have remained quite comparable over time. However, during the intervention period (from I1 to I5), hospitalization rates were consistently higher among the comparison group than the intervention group, and the between-group gap at each measurement point was significantly different from the average baseline difference of .01. This finding suggests that Race to Health! may have helped lower the hospitalization rates for the intervention group.

Figure VIII.3. Hospitalizations per Medicare patient per 6-month period: Beginning of baseline period to end of intervention period



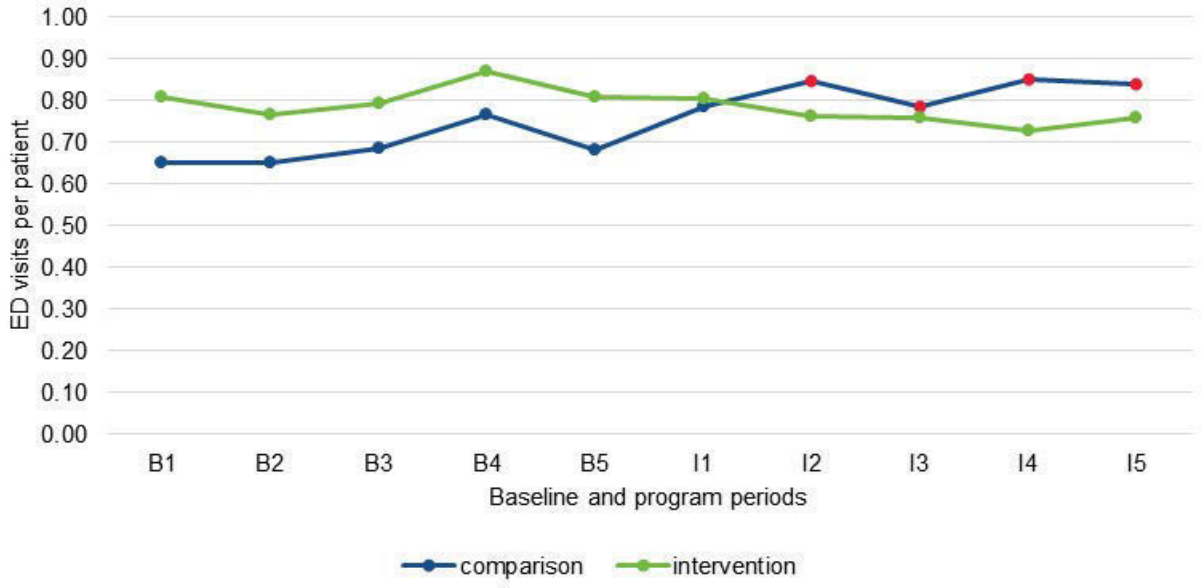
Source: Mathematica analysis of Medicare administrative data for July 2010–June 2015.

Note: Means are regression-adjusted. Red points indicate that the difference between the intervention and comparison group mean in the intervention period is significantly different from the average difference that occurred in the baseline period. B1, B2, etc., indicate each 6-month period of the analysis before the intervention began, and I1, I2, etc., indicate each 6-month period of the analysis during the intervention period; 823 intervention group members and 2,643 comparison group members were included in this analysis. Sample sizes varied from period to period, depending on data availability.

Emergency department visits. During the baseline period, the ED visit rate was consistently higher among the intervention group than the comparison group, but the difference between groups was quite small (only by about .12 visits per patient; Figure VIII.4). Presumably, this difference would have persisted in the absence of the program.

However, as Figure VIII.4 shows, the comparison group's rate began trending upward as the intervention period began, whereas the ED rate for the intervention group began trending downward. Even though the absolute value of the resulting gap was not very different, the direction was reversed compared with the baseline period: by the I2 measurement point, the intervention group's rate was lower than the comparison group's rate, and the gap between the two groups from I2 to I5 was significantly different from the average baseline gap, suggesting that the program reduced ED visits for the intervention group relative to the comparison group.

Figure VIII.4. Emergency department visits per Medicare patient per 6-month period: Beginning of baseline period to end of intervention period



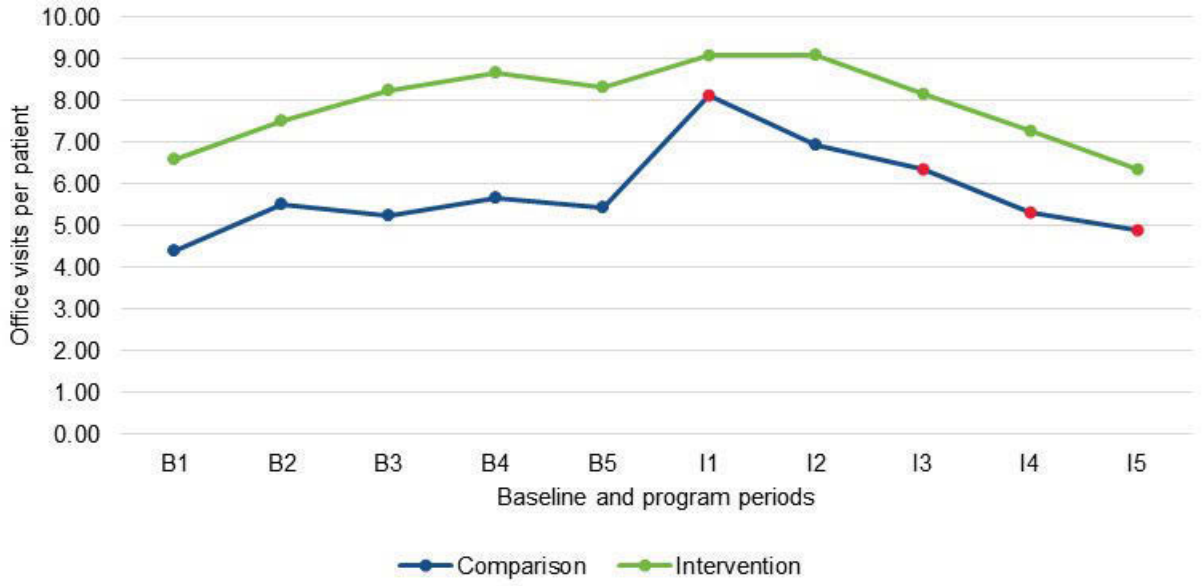
Source: Mathematica analysis of Medicare administrative data for July 2010–June 2015.

Note: Means are regression-adjusted. Red points indicate that the difference between the intervention and comparison group mean in the intervention period is significantly different from the average difference that occurred in the baseline period. B1, B2, etc., indicate each 6-month period of the analysis before the intervention began, and I1, I2, etc., indicate each 6-month period of the analysis during the intervention period; 823 intervention group members and 2,643 comparison group members were included in this analysis. Sample sizes varied from period to period, depending on data availability.

Office visits. As Figure VIII.5 illustrates, office visits generally trended upward for both the intervention and the comparison groups during the baseline period. On average, from B1 to B5, the mean number of visits for the intervention group exceeded the mean number of visits for the comparison group by 2.7 visits.

During the intervention period, the number of office visits remained higher for the intervention group, but the difference between the two groups became smaller. At I.1, the difference between the two groups was less than 1—significantly lower than the baseline average of 2.7. The differences also were significantly less at I3, I4, and I5. The number of office visits decreased throughout the intervention period for both groups. These figures suggest that KMHS' program may have contributed to reducing the number of office visits over time relative to the comparison group; however, because trends in the comparison group office visit rate changed substantially at the start of the intervention period, these results should be interpreted with caution.

Figure VIII.5. Office visits per Medicare patient per 6-month period: Beginning of baseline period to end of intervention period



Source: Mathematica analysis of Medicare administrative data for July 2010–June 2015.

Note: Means are regression-adjusted. Red points indicate that the difference between the intervention and comparison group mean in the intervention period is significantly different from the average difference that occurred in the baseline period. B1, B2, etc., indicate each 6-month period of the analysis before the intervention began, and I1, I2, etc., indicate each 6-month period of the analysis during the intervention period; 823 intervention group members and 2,643 comparison group members were included in this analysis. Sample sizes varied from period to period, depending on data availability.

b. Impact analyses

KMHS Medicare population. For the impact analysis, we conducted an interrupted time series analysis to assess the difference in pre- vs. post-period trends in outcomes for the intervention group relative to the control group. The results of this analysis also suggest that the program significantly reduced overall Medicare expenditures, hospitalizations, ED visits, and office visits for KMHS patients relative to the comparison group (Table VIII.3). Notable findings are as follows:

- During the study period, we estimated that Medicare expenditures decreased $266 per enrolled beneficiary month for intervention group patients relative to the comparison group (p-value < 0.01). Overall, we estimated total savings of $5,144,000 for the 13 percent of KMHS patients who were Medicare beneficiaries.
- There were fewer hospitalizations and fewer ED visits for patients relative to the comparison group by 0.02 and 0.03 per enrolled month, respectively (p-value <0.01 for both estimates).
- The mean number of office visits decreased significantly during the intervention period relative to the baseline period for both the intervention and comparison populations but the relative decrease was greater for the intervention group (p-value < 0.01).

Table VIII.3. Impacts and total savings attributable to Race to Health—KMHS Medicare population

	All intervention group members	
	Change	90% confidence interval
Aggregate results		
Total Medicare expenditures (in thousands of dollars)	-$5,144	[-$7,994 to -$2,293]
Hospitalizations	-297	[-434 to -159]
Emergency department visits	-546	[-838 to -254]
Office visits	-2,560	[-4,030 to -1,089]
Per beneficiary month		
Expenditures (in dollars)	-$266	[-$413 to -$118]
Hospitalizations	-0.02	[-0.02 to -0.01]
Emergency department visits	-0.03	[-0.04 to -0.01]
Office visits	-0.13	[-0.21 to -0.06]
Number of patients	846	
Mean number of intervention months per patient	23	
Approximate proportion of intervention population represented in analysis[a]	13%	
Intervention period	January 1, 2013, to June 30, 2015	

Source: Mathematica analysis of fee-for-service Medicare administrative data for baseline and program periods, January 2010–June 2015. Data for calendar year 2009 were used to develop indicators of baseline health status.

Note: Impact estimates were derived from regression models controlling for age (linear and squared), gender, race/ethnicity, cohort participation, dual eligibility status, whether 12 months of baseline data were available, behavioral health diagnoses, length of time in mental health treatment, disability status, and HCC condition indicators. We derived the impact estimates in Stata using the margins command to compare the difference between the intervention and baseline period means for the treatment and comparison groups accounting for the nonlinear modeling approach. The confidence intervals for total expenditures, hospitalizations, ED visits and office visits were derived based on bootstrap methods and were adjusted for multiple testing based on the generalized Tukey method. Readmissions were not included in the adjustment for multiple testing due to small sample size.

[a] We calculated the approximate proportion of intervention population represented in the analysis by dividing the number of patients (846) in the Medicare analysis by the number of individuals who participated in KMHS' program between January 2013 and June 2015 (6,662).

Analyses by cohort status. As noted, one component of Race to Health! involved periodically identifying and selecting groups of patients, known as cohorts, based on their physical comorbidities. Although all KMHS patients were exposed to the HCIA program because it was implemented throughout the agency's patient population, KMHS staff specifically targeted cohorts for physical health data collection because of their high level of need. We conducted impact analyses to determine whether the effects of the Race to Health! were more prominent among the cohort patients. Overall, we found no consistent evidence that the program had a greater or lesser benefit for the cohorts than for the overall patient population. (Details of these analyses appear in Appendix A.)

c. Analytic limitations of the Medicare analysis

The primary limitation of this analysis is its lack of generalizability, because it is limited to Medicare fee-for-service (FFS) enrollees. This is about 13 percent of the KMHS target population affected by the implementation of Race to Health! Individuals enrolled in Medicare

Advantage were excluded from this analysis. The choice to participate in Medicare Advantage is associated with particular health care needs,[47] and therefore the program may have different effects on Medicare Advantage beneficiaries than it does on beneficiaries in FFS Medicare. Out-of-pocket expenditures and services not covered by Medicare may have been affected by the program, but they were not addressed in this analysis. Overall, our findings are not generalizable to all KMHS patients and services.

2. Results from Medicaid analyses

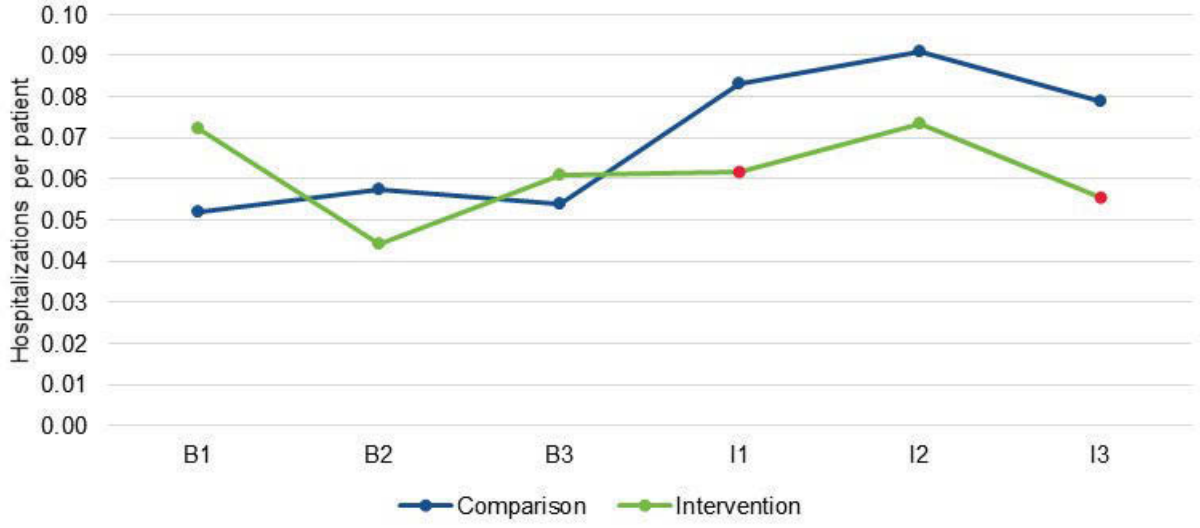
a. Descriptive analyses

Before developing impact estimates for the intervention and comparison group, we analyzed the trends in means for two outcome measures, hospitalizations and ED visits.

Hospitalizations. As Figure VIII.6 shows, there was a somewhat erratic trend in hospitalization rates for the intervention group during the baseline period (B1–B3); the trend for the comparison group was more stable. As a result, the average difference between the two groups was zero during the baseline period.

After the intervention period began, the comparison group's hospitalization rates trended sharply upward. The intervention group's rates also trended upward, but not as sharply; as a result, the gap between the lines widened. Overall, at I1 and I3, the difference between the two groups was significantly different compared with the average baseline difference. These findings suggest that Race to Health! may have helped retain lower hospitalization rates for the intervention group in the face of forces leading to dramatic increases in the comparison group.

Figure VIII.6. Hospitalizations per Medicaid patient per 6-month period: Beginning of baseline period to end of intervention period



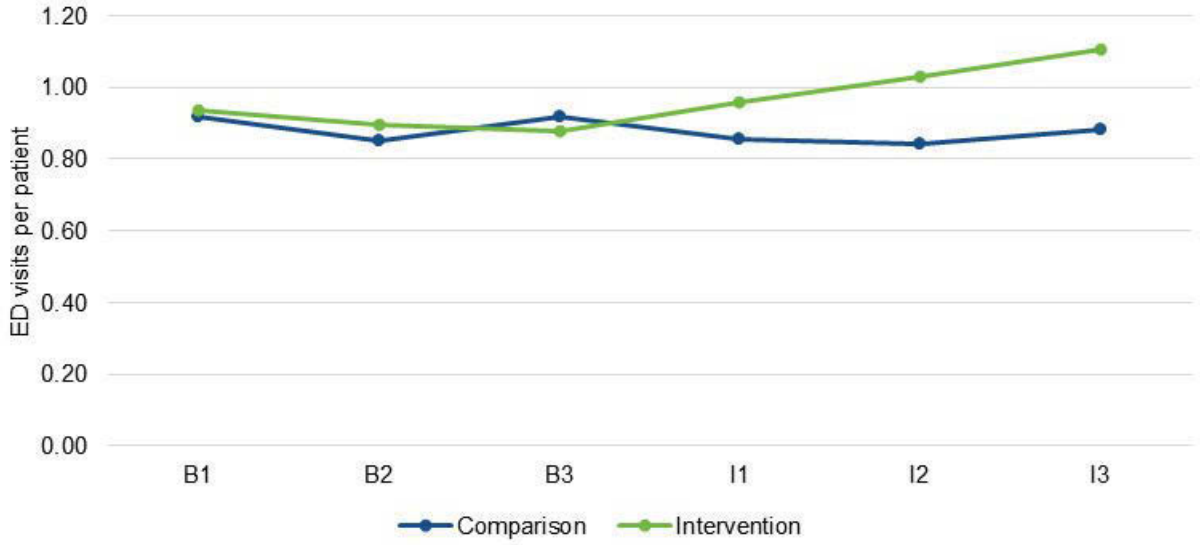
Source: Mathematica analysis of Medicaid administrative data for January 2011–June 2014.

[47] Biles, Brian, Giselle Casillas, and Stuart Guterman. "Variations in County-Level Costs Between Traditional Medicare and Medicare Advantage Have Implications for Premium Support." *Health Affairs*, vol. 34, no. 1, January 2015, pp. 56–63.

Note: Means are regression-adjusted. Red points indicate that the difference between the intervention and comparison group mean in the intervention period is significantly different from the average difference that occurred in the baseline period. B1, B2, etc., indicate each 6-month period of the analysis before the intervention began, and I1, I2, etc., indicate each 6-month period of the analysis during the intervention period; 3,753 intervention group members and 16,234 comparison group members were included in this analysis. Sample sizes varied from period to period, depending on data availability.

ED visits. As Figure VIII.7 illustrates, the ED visit rates for the intervention and comparison groups were similar during the baseline period. After the intervention started, the ED visit rate for the intervention group began an upward trend while the rate for the comparison group remained stable. Consequently, the difference between the two groups at all three intervention periods was significantly greater than the average baseline difference. These findings suggest that the KMHS program may have contributed to increased ED visit rates for Medicaid beneficiaries.

Figure VIII.7. Emergency department visits per Medicaid patient per 6-month period: Beginning of baseline period to end of intervention period



Source: Mathematica analysis of Medicaid administrative data for January 2011-June 2014.

Note: Means are regression-adjusted. Red points indicate the difference between the intervention and comparison group mean in the intervention period is significantly different from the average difference that occurred in the baseline period. B1, B2, etc., indicate each 6-month period of the analysis before the intervention began, and I1, I2, etc., indicate each 6-month period of the analysis during the intervention period. 3,753 intervention group members and 16,234 comparison group members were included in this analysis. Sample sizes varied from period to period, depending on data availability.

Cohort analyses. As for the Medicare analysis, we found little evidence that the program had different effects for Medicaid beneficiaries in the cohorts.

b. Impact analysis

Results of the interrupted time series analysis suggest that the program increased overall Medicaid ED visits for patients in the intervention group relative to the comparison group ($p < 0.01$) and decreased hospitalizations ($p < 0.05$) (Table VIII.4). These results are preliminary because we expect to obtain additional Medicaid data and conduct further analyses. We will

report the findings from these additional analyses in the addendum to this report, which we expect to submit to CMMI in the spring of 2017.

Table VIII.4. Medicaid impacts attributable to intervention

	All intervention group members		Cohort		Non-cohort	
	Change	90% confidence interval	Change	90% confidence interval	Change	90% confidence interval
Aggregate results						
Hospitalizations	-199	[-338 to -60]	-70	[-148 to 8]	-128	[-245 to -11]
ED visits	1,592	[1,074 to 2,109]	303	[16 to 591]	1,288	[857 to 1,720]
Per beneficiary month						
Hospitalizations	-0.003	[-0.006 to -0.001]	-0.007	[-0.014 to 0.001]	-0.003	[-0.005 to -0.000]
ED visits	0.026	[0.018 to 0.035]	0.029	[0.001 to 0.056]	0.026	[0.017 to 0.035]
Number of patients	3776		648		3128	
Mean number of intervention months per patient	16		16		16	
Approximate proportion of intervention population represented in analysis[a]	57%		n/a		n/a	
Intervention period	January 1, 2013, to June 30, 2014					

Source: Mathematica analysis of Medicaid administrative data for baseline and program periods January 2011–June 2014.

Note: Impact estimates were derived from regression models controlling for age (linear and squared), gender, race/ethnicity, cohort participation, dual eligibility status, whether 12 months of baseline data were available, behavioral health diagnoses, disability status, and CDPS scores. We derived the impact estimates in Stata using the margins command to compare the difference between the intervention and baseline period means for the treatment and comparison groups accounting for the nonlinear modeling approach. The confidence intervals for hospitalizations and emergency department (ED) visits were derived based on bootstrap methods and were adjusted for multiple testing based on the generalized Tukey method.

[a]We calculated the approximate proportion of intervention population represented in the analysis by dividing the number of patients (3,776) in the Medicaid analysis by the number of individuals who participated in KMHS' program between January 2013 and June 2015 (6,662).

c. Analytic limitations of the Medicaid analysis

The results presented in this section have important limitations:

- The results reflect the intervention period (IY1) from January 2013 (the program start date) to June 2014. HCIA funding for the program continued through June 2015. Consequently, as data on additional patients and the full period of the intervention become available, the findings for this population may be different.
- This analysis was limited to Medicaid enrollees and Medicaid-covered services. Therefore, the findings are not generalizable to all KMHS patients and services. The group of patients included in this analysis reflected 57 percent of the KMHS target population.

- We did not assess total expenditures in this analysis because most KMHS patients were enrolled in Medicaid managed care, and the encounter data analyzed for this population did not include reliable information on expenditures.
- We did not report results on hospital readmissions, because there were so few readmissions that results would have been unreliable.
- We were unable to identify specific mental health facilities in the Medicaid data available for this analysis. Thus, we were unable to limit the comparison population to individuals receiving mental health treatment at facilities similar to KMHS. The comparison population included in this analysis represents a broader population of individuals receiving mental health treatment at all types of providers throughout the state.
- Alpha-MAX data for January through June 2014 included claims paid through September 2014. Alpha-MAX data includes four quarters of paid claims, whereas traditional MAX files reflect seven quarters of payments. As a result, services provided through June 2014 will not be fully represented if payments for the services were not processed by September 2014. Differences in claim submission lags for KMHS patient providers relative to comparison patient providers may influence findings, particularly for the third post-intervention period.

3. Analysis of health status measures

As noted, we focused on three measures (blood pressure screening, BMI screening, and metabolic screening) to assess the program's effectiveness. We describe the results of our descriptive analyses for each measure below.

Blood pressure screening. We examined blood pressure screening rates for four groups of KMHS patients:

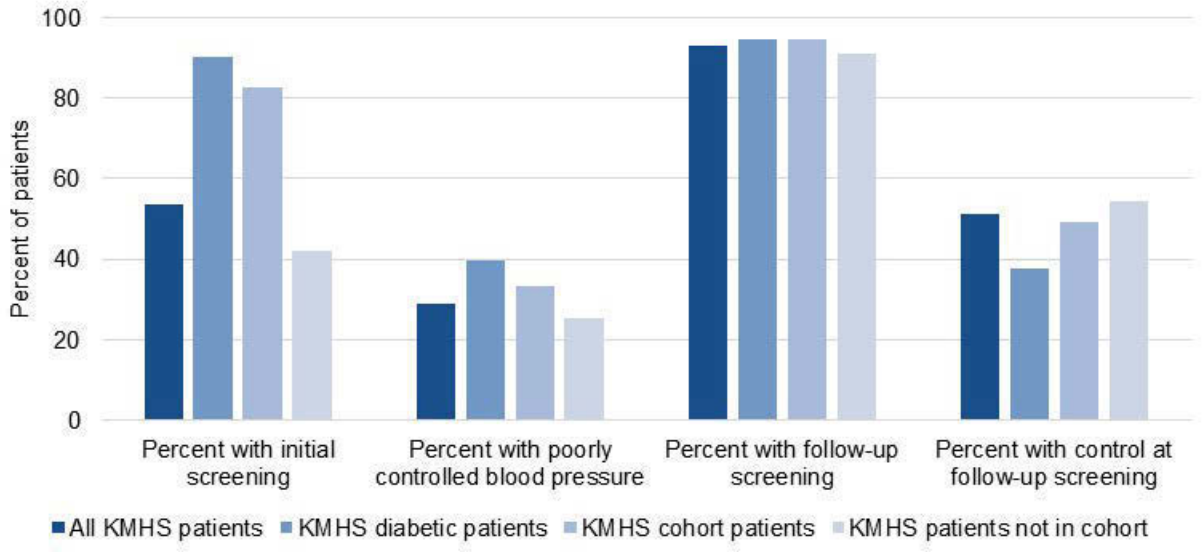
1. All patients included in the study
2. Patients with diabetes, because they are at risk for high blood pressure
3. Patients in the cohorts, because they received particular attention from KMHS for physical health data collection
4. Patients who were not in one of the cohorts, to serve as an internal comparison group for cohort patients

For each of these groups, we calculated the percentage of the group who (1) received an initial screening, (2) had poor blood pressure control at the initial screening, (3) needed and received follow-up screening, and (4) among those with poor control at initial screening, had adequate control at follow-up screening.

We found that over half of all KMHS patients had an initial blood pressure screening in 2014; of these, nearly one-third had poorly controlled blood pressure at their initial screening (Figure VIII.8). Initial screening rates were much higher for diabetic and cohort patients (90 percent and 83 percent, respectively), suggesting that KMHS staff were successful in screening patients at high risk. As expected, these groups also had higher rates of poorly controlled blood pressure at initial screening (40 percent and 33 percent, respectively).

Among those screened within the first six months of 2014 and who had poorly controlled blood pressure at their initial screening, follow-up screening rates were over 90 percent for all KMHS patients and each patient subgroup—suggesting that KMHS staff were successful in conducting follow-up screening for the high-risk patients. Of those with poorly controlled blood pressure at their initial screening, nearly half of all patients and half of the cohort patients had controlled blood pressure at the follow-up. Fewer diabetic patients had controlled blood pressure at follow-up (38 percent).

Figure VIII.8. Blood pressure screening, control, and follow-up



Source: Mathematica analysis of KMHS' EHR data for January 2014-June 2015.

Note: The denominator for "percent with initial screening" and "percent with poorly controlled blood pressure at initial screening" was limited to those meeting the criteria who were active patients in 2014 and were age 18 or older as of January 1, 2014. The denominator for "percent with follow-up screening" and "percent with control at follow-up screening" was further limited to those who had poorly controlled blood pressure at initial screening in the first six months of 2014.

BMI screening. For this analysis, we examined BMI screening rates for four groups of KMHS patients:

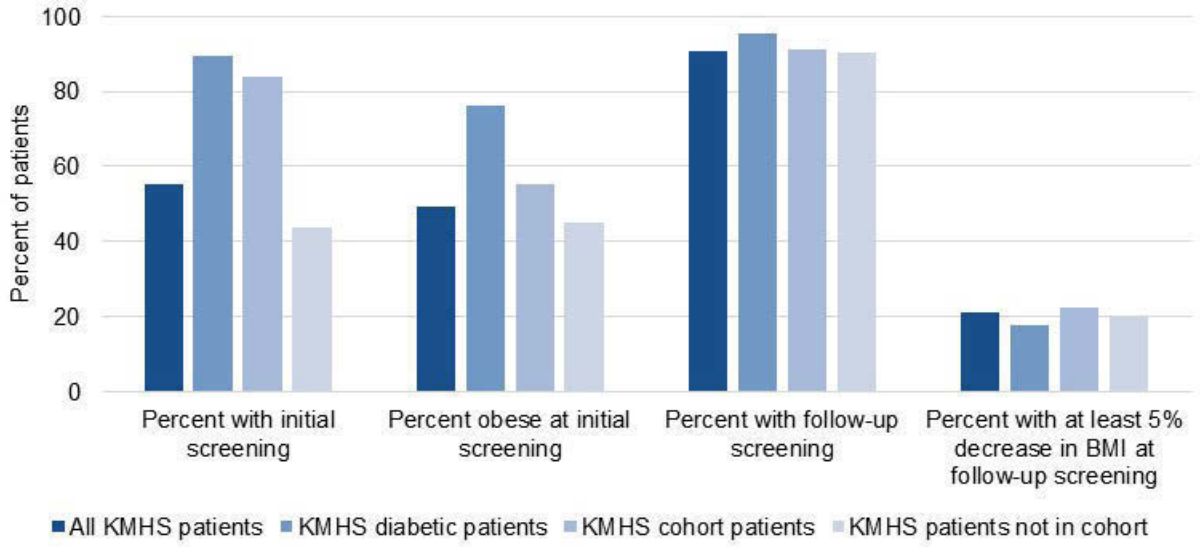
1. All patients included in the study
2. Patients with diabetes, because they are at risk for obesity
3. Patients in the cohorts, because they received particular attention from KMHS staff for physical health data collection
4. Patients who were not in the cohorts, because they serve as an internal comparison group for cohort patients

For each of these groups, we calculated the percentage who received an initial screening, were found to be obese at the initial screening, needed and received follow-up screening, and, among those who were overweight or obese at initial screening, had at least a 5-percent decrease in BMI at follow-up.

More than half of KMHS patients had an initial BMI screening in 2014 (Figure VII.9). As for blood pressure control, screening rates were much higher for patients with diabetes and those in the cohorts (89 percent and 84 percent, respectively), indicating that KMHS was successful in screening high-risk patients for obesity. Nearly half of all patients were obese at initial screening, and another one-quarter were overweight. Of diabetic patients, over three quarters were obese at initial screening, and another 16 percent were overweight.

Among those screened in the first six months of 2014, follow-up screening rates within 12 months were over 90 percent for those who were overweight or obese at initial screening. For less than a quarter of the patients overweight or obese at initial screening, BMI decreased by at least 5 percent at follow-up.

Figure VIII.9. BMI screening, rate of obesity, and follow-up



Source: Mathematica analysis of KMHS' EHR data for January 2014–June 2015.

Note: The denominator for "percent with initial screening" and "percent obese at initial screening" was limited to those meeting the criteria who were active patients in 2014 and were age 18 or older as of January 1, 2014. The denominator for "percent with follow-up screening" and "percent with at least 5% decrease in BMI at follow-up screening" was further limited to those who were overweight or obese at an initial screening in the first six months of 2014.

Metabolic screening. We examined metabolic screening among non-diabetic KMHS patients who were prescribed antipsychotic medications.[48] Consistent with the NCQA measurement definition, we limited our analytic population to patients who (1) were between the ages of 18 and 64 at start of measurement period, (2) had an active diagnosis of schizophrenia or bipolar disorder during the measurement period, and (3) did not have an active diagnosis of diabetes or receive a diabetes medication during the measurement period. Within this group, we determined the number of patients who received a metabolic screening on the day of or after

[48] Specifically, we included only patients who received a prescription for an antipsychotic medication between the first day of the measurement period and 30 days before the end of the measurement period. This ensured that the prescription was active during the measurement period and allowed time to receive the metabolic screening following receipt of the prescription.

their first antipsychotic prescription in the measurement period. We repeated this analysis for three six-month measurement periods: January to June 2014, July to December 2014, and January to June 2015.

In each six-month period, over 90 percent of the sample had a prescription for an antipsychotic medication (Table VIII.5). Among those, less than 18 percent had a metabolic screening during the measurement periods, indicating substantial room for improvement.

Table VIII.5. Metabolic screening for non-diabetic KMHS patients on an anti-psychotic medication

	January–June 2014	July–December 2014	January–June 2015
	Number (Percent)	Number (Percent)	Number (Percent)
Non-diabetic patients on antipsychotic medication	358 (92.0%)	374 (91.9%)	379 (92.4%)
Of above, those with metabolic screening	59 (16.5%)	28 (7.5%)[a]	65 (17.2%)[b]

Source: Mathematica analysis of KMHS' EHR data for January 2014–June 2015.

[a] Percent within all KMHS patients was statistically significantly different ($p < 0.05$) between January–June 2014 and July–December 2014.

[b] Percent within all KMHS patients was statistically significantly different ($p < 0.05$) between July–December 2014 and January–June 2015.

4. Qualitative findings on perceived effects

During interviews, respondents reported several key features of Race to Health! that may have influenced patients' outcomes. They believed the program's emphasis on training staff to focus on whole health led to better health outcomes for patients. Staff reported that their own improved awareness of physical health conditions enhanced their ability to discuss these issues with patients, advocate on patients' behalf, and help connect them to necessary medical care.

KMHS staff also said greater access to and use of physical health care data helped them treat patients differently, which may affect quantitative outcomes. For example, staff noted that ED visit alerts made them more aware of patients' ED visits and also allowed them to discuss the reasons for those visits. ED visit data also helped KMHS staff coordinate better with patients' other providers in the community. For example, KMHS staff used ED data to identify patients who would benefit from "collaborative care conferences," which brought together providers and other community stakeholders (emergency medical technicians, the police department, health plans) to coordinate a specific patient's care. Respondents also noted that KMHS staff gained a better understanding of their patients' medications because of the information medical assistants collected from primary care providers, which helped the agency's psychiatrists make more informed decisions about prescribing. For Medicare-enrolled patients, this heightened awareness of ED visits, exchanges between the clinic team and patients, and better coordination with community providers collectively may have helped reduce the patients' need for both ED and office visits (given that we found reductions on both measures). For Medicaid patients, the story is a bit less clear because we found that the program may have increased ED visits. We will re-examine this finding when we obtain and analyze additional Medicaid data.

Respondents noted that Race to Health! wellness activities, including health education and groups supporting self-management of chronic conditions, helped some patients adopt healthier behaviors that may ultimately result in better long-term health. For example, they reported that patients practiced more healthy behaviors, such as exercising and quitting smoking. Wellness programing also empowered patients to manage their own care and seek better relationships with physical and behavioral health providers.

D. Findings about the workforce

KMHS leaders and staff noted that most behavioral and physical health care staff had limited training in integrated care models or settings. As a result, staff required more intensive training after being hired. To address this challenge, through Race to Health!, KMHS developed a training academy to ensure all agency staff were quickly oriented to the agency's unique integrated care model and culture of whole health. Delivering training during care team meetings—and emphasizing that seeking consultation was part of care team members' roles—also helped ensure that training continued despite changing priorities and high caseloads.

In both qualitative interviews and workforce survey responses, KMHS staff reported that they felt well-supported in their work. During site visit interviews, KMHS staff highlighted the importance of having a strong infrastructure in place to support the organization's transition to a whole-health focus. In particular, respondents pointed to the program's new health information technology and staff roles (that is, medical assistants, healthy family coordinators, and the healthy living program developer) as critical to supporting existing staff as they took on new responsibilities and expanded their focus beyond clients' mental health.

E. Program sustainability and spread

As of April 2016, KMHS continued to sustain all components of its Race to Health! program, both within the agency and in the community, but expected to face challenges sustaining some of these components in the future.

When HCIA funding ended in June 2015, KMHS leaders committed agency funds to support the positions necessary to continue the program's operations, including multidisciplinary outpatient care teams for adults and children that focus on patients' whole health, chronic disease and wellness education programming within the agency, and behavioral health intervention and psychiatric consultation in community primary care practices. However, KMHS leaders reported in 2016 that Washington State has since made significant cuts to Medicaid funding for community mental health services, which resulted in "difficult choices" for sustainability. For example, the agency will likely no longer be able to continue funding some positions dedicated to program development and staff training. The awardee noted that its use of HCIA funding to develop new programming, train staff for new responsibilities, and adapt health IT infrastructure helped to "lay the foundation" for more lasting change within the agency. Despite losing some capacity in program development and staff training, program leaders expect to continue integrating treatment for mental health and substance use disorders with physical health care, in keeping with the program's goals. Staff also credited the strong leadership and commitment of senior KMHS executives with being a critical factor in the program's sustainability.

In addition, KMHS leaders noted they hope to sustain most Race to Health! activities in the community, including providing behavioral health intervention and referral services at HHP primary care practices. Although KMHS will no longer be able to fund the behavioral health professional position to deliver brief behavioral health intervention and consultation in primary care sites throughout the community, KMHS maintains a strong relationship with HHP, its partner organization, and is discussing the potential for HHP to commit funding for this position. In addition, KMHS expects to assign some of the psychiatric consultant's responsibilities to other KMHS psychiatric providers when the current consultant retires in fall 2016. These psychiatrists will take turns being on call to respond to consultation requests and questions from community primary care providers. However, KMHS leaders noted that the agency will no longer be able to benefit from having a single, designated person interacting with primary care providers in the community, serving as a role model, and providing on-site training.

Program leaders and staff also reported that separate regulations and funding streams for services, depending on whether they address physical health, behavioral health, or substance abuse, pose barriers to the program's sustainability. Several respondents noted challenges posed by 42 CFR Part 2, the federal regulation on sharing of patient information related to substance use. KMHS and its partners reported that the regulation created uncertainty about which types of information could be shared between providers.

Program leaders also voiced concern about the uncertain future of behavioral health service delivery and financing in Washington State, and noted the difficulty of developing an innovative service delivery model without having supportive financing and policy arrangements in place at both the state and federal level. For example, the state is moving toward adopting an integrated purchasing model for behavioral health services. KMHS faces uncertainty about the effect of this policy change on the funding streams that would support the program's components in the future.

F. Lessons learned

Overall, robust evidence from our evaluation suggests that Race to Health! may have reduced Medicare expenses for its patients, possibly by reducing the number of ED visits, hospitalization rates, and office visits.[49] These results may be related to several aspects of the program that KMHS staff underscored during our site visits, including the heightened focus on physical health care in patients' interactions with their mental health and substance abuse treatment providers, more access to and use of physical health data within KMHS and in coordination with other providers, and development and implementation of new groups to promote wellness and self-management of chronic disease. Our analysis of health status measures further suggests that in its implementation, the program successfully achieved many of its goals.

Qualitative findings revealed that programs like Race to Health! may require significant investment in training and technology infrastructure, which staff highlighted as an important foundation for their transition to this new model of care. Training may be particularly important

[49] For Medicaid patients, we found that the program may have decreased hospitalizations and increased ED visits. We consider this finding quite preliminary and well re-examine it when we obtain additional data.

to program implementation if, as with KMHS, few staff members have had exposure to integrated and coordinated care approaches.

Finally, our evaluation findings indicate that sustaining coordinated and integrated service delivery models like Race to Health! is possible with strong support from leadership, creative approaches to partnerships with community health and social service providers, and an up-front focus on building the infrastructure and staff capacity necessary to continue the program. However, such innovative service delivery models ultimately require innovative payment structures if they are to be sustained in the long term.

IX. MAIMONIDES MEDICAL CENTER[50]

Findings from Mathematica's Evaluation of MMC's HCIA Program

- Our quantitative impact analysis showed no statistically significant change in outcomes for the FFS Medicare population (a small proportion of their participant group).
- Our quantitative pre-post analysis showed a statistically significant increase in expenditures and decreases in hospitalizations and ED visits for the Medicaid population (the majority of their participants); however, without the use of a comparison group, we were unable to determine if the changes in outcomes were due to the program or other external factors.
- MMC's experience with similar projects, its preexisting network of familiar partners, and creative use of funding sources available through New York State's Medicaid agency all contributed to a successful implementation and sustainment of its program.
- MMC's electronic care coordination platform (CCP), a key program component, was not enough to fully meet the coordination needs of providers. HCIA-funded care management staff played a vital role in helping fulfill these needs. However, MMC continually improved the CCP over the course of the program, and made changes to other components so they fit better with the way providers ultimately used the CCP.

A. Introduction

Maimonides Medical Center (MMC), a tertiary care center in southwest Brooklyn, New York, used HCIA funding to implement a program designed to improve the care of individuals with serious mental illness (SMI). To implement the program, MMC partnered with members of the Brooklyn Care Coordination Consortium, a group of more than 20 social service agencies and medical institutions. MMC and its partners designed the program for individuals who lived or received care in selected zip codes in southwest Brooklyn and who had mood disorders (including depression and bipolar disorders), schizophrenia, or other psychotic disorders.

Program staff worked with participants' existing medical, mental health, and community service providers to create multidisciplinary care teams, who were supported by HCIA-funded care management staff. Members of the care team shared information through an electronic care coordination platform (CCP) built to give participants a virtual medical and mental health home.

Before the HCIA funding, New York State granted MMC status as a Medicaid health home; as a result, MMC ultimately expected to provide health home services to roughly 7,000 Medicaid enrollees. The HCIA award gave MMC (1) the capacity to provide care management to 500 individuals with Medicare, commercial insurance, or no insurance (that is, individuals who were not eligible to receive services through the Medicaid health home payment model) and (2) funding to establish the technology and training infrastructure necessary to provide virtual health homes to the entire target population of 7,500 participants (that is, the 7,000 Medicaid beneficiaries and 500 other participants).

[50] We thank New York State Department of Health (NYSDOH) for providing data to support the Medicaid analyses in this chapter. The findings and conclusions presented are those of Mathematica Policy Research alone and not those of NYSDOH.

For the program's Medicare beneficiaries we were able to identify a valid comparison group and therefore could assess the program's impact on key outcomes. However, for reasons described below, we were unable to identify a valid comparison group for the program's Medicaid participants; hence, we conducted a pre-post analysis for this group.

The findings in this chapter are based on quantitative and qualitative data collected or received through May 11, 2016, and enrollment data reported throughout the award period. We will present findings from additional analyses in an addendum to this report.

We drew on the following data sources:

- Enrollment data submitted by MMC to the reporting website maintained by CMMI's technical assistance contractor (the Lewin Group)
- Medicare claims and enrollment data extracted from CMS' Medicare data files (MMC provided identifiers for program participants)
- Medicaid claims and enrollment data extracted from New York State's Medicaid data files (MMC provided identifiers for program participants)
- Qualitative data, including telephone interviews and in-person site visits in spring 2014 and 2015. During the site visits, Mathematica conducted in-depth interviews with awardee leadership, members of the workforce, and other stakeholders. In addition, Mathematica convened focus groups with members of the workforce and with program participants in spring 2015. We also conducted a telephone interview in spring 2016 with MMC leaders to discuss sustainability of the program.

1. Overview of administrative context

MMC is a large medical center in southwest Brooklyn, New York. MMC executives led the HCIA program in partnership with a group of care management organizations and other social service organizations throughout the community. Internal staff and external partners consider the organization to be progressive and nimble, saying that MMC is always looking for ways to improve care by staying abreast of opportunities presented by state and local policies.

MMC leaders also saw the program as a continuation and augmentation of work they developed and tested under other funding sources. For example, MMC originally conceived and began to develop the CCP, a vital component of its HCIA program, under a Health Care Efficiency and Affordability Law grant from the state, but used HCIA funding to significantly expand and refine the CCP. Leaders, staff, and partners told us that MMC's long-standing commitment to improving care was a key facilitator of HCIA implementation.

Program leaders also said that the MMC was well served by its existing partnerships and standing in the community because of the HCIA program's reliance on partner organizations for program leadership and care management services. In addition, leaders told us that establishing a committee structure to develop and hone program components and procedures, and to help the program evolve and respond to the needs of partner organizations and their staff, kept partners engaged and invested in the success of the program. For example, the program's Health Information Technology Committee gave input on and suggested improvements to the CCP throughout the program's implementation—guidance that also benefited the partner

organizations. Program leaders also used strong internal self-monitoring to guide the committees as they evaluated performance and identified areas for improvement.

MMC developed and is sustaining its program in the midst of significant changes and reforms to the state's Medicaid program. In one such reform, the state amended its Medicaid waiver to use specialty managed care arrangements to meet the unique needs of Medicaid beneficiaries with mental illnesses. Most relevant to MMC, the state plans to enroll adult beneficiaries with serious behavioral health needs in Health and Recovery Plans (HARPs). HARPs will integrate physical, mental health, and substance use disorder services, and offer an expanded range of home- and community-based behavioral health services to enrollees. MMC developed a HARP pilot in collaboration with some of its core partners and Health First (an insurance company that is one of its major payers), and is using the pilot as part of its strategy for sustaining the HCIA program.

The state is using the federal Delivery System Reform Incentive Payment (DSRIP) waiver program, part of the federal 1115 waiver authority, as its primary mechanism for achieving Medicaid reforms. The DSRIP program is designed to restructure the health care delivery system in the state by reinvesting federal savings from the state's recent Medicaid redesign in the state's Medicaid program, with the primary goal of reducing avoidable hospital use by 25 percent over five years. The program requires a consortium of organizations to collaborate with the goal of improving the community's health through certain prescriptive projects in certain domains (such as integrating primary and behavioral health care) and the use of required metrics to ensure that desired outcomes are achieved. Up to $6.4 billion is available to providers across the state if they reach certain predefined results. MMC is implementing projects under this initiative and using DSRIP funding as part of its strategy to sustain the HCIA.

The other significant external factor that may have affected program implementation is the New York City's perpetual dearth of affordable housing for those with SMI. Almost all respondents named housing as a significant issue that affected their ability to better coordinate care and improve the health and well-being of clients. Program staff underscored the difficulties in addressing the mental and physical health needs of participants who lack stable housing, noting that "a lesson learned in general is that … we can't emphasize enough how big a part [housing plays] in engagement," and that the housing crisis made participants "so understandably very focused on [housing]. [Housing] is priority 1, 2 and 3, and then everything else is like 7, 8, 9." Respondents mentioned some state efforts to tie housing eligibility to health homes enrollment, but it is not clear how far along this policy initiative is.

2. Progress through phases of innovation

MMC's HCIA funding enabled the awardee to develop and expand initiatives created under previous funding opportunities. MMC tested the HCIA program in multiple sites from the outset and continuously refined it throughout the award period (Figure IX.1). By the end of the award, MMC was working actively to ensure that the program would be fully institutionalized and sustained—the last phase in our framework.

3. Enrollment

For this evaluation, MMC defined direct participants as people who were uninsured, had commercial insurance, or were enrolled in Medicare only. These direct participants received care

management and outreach services, and also benefitted from HCIA funding used for the improved information technology (IT) infrastructure and training to support the care management services. By the end of the program's 12th quarter, MMC enrolled 635 direct participants (Figure IX.2), well over its original projections of 500.

Figure IX.1. Phases of program innovation: MMC

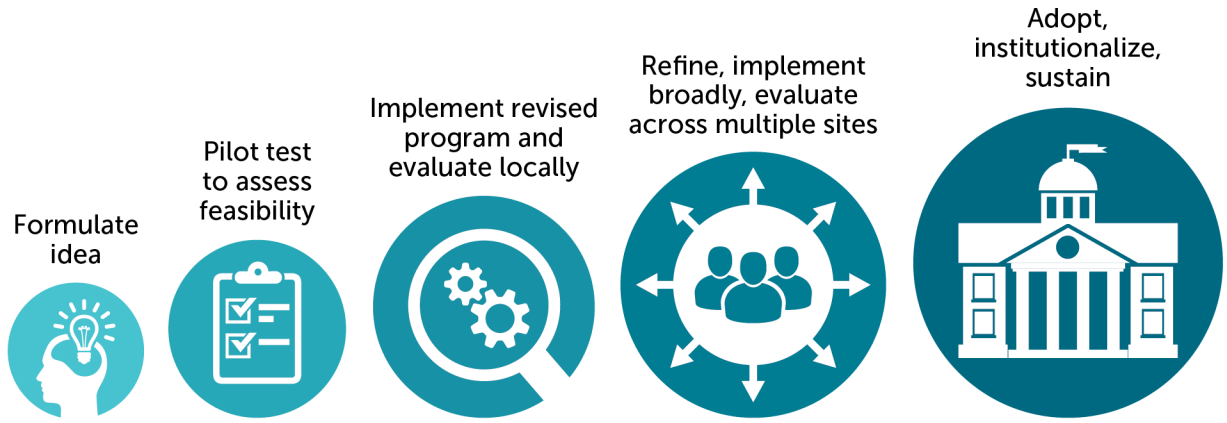
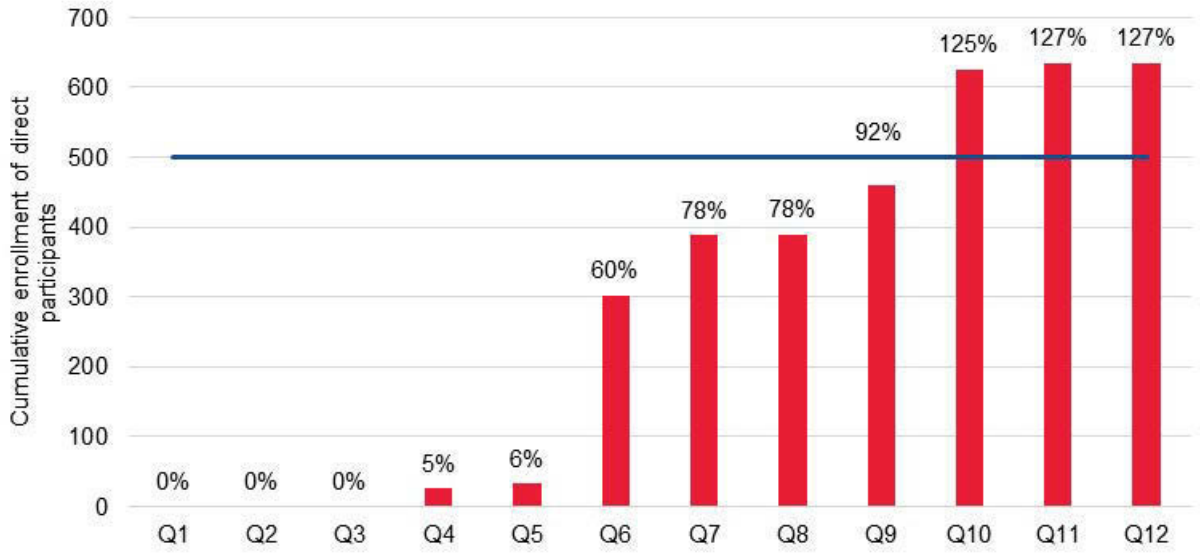


Figure IX.2. Percent of target enrollment achieved by quarter, Q1–Q12



Source: Quarterly reports submitted to the website maintained by CMMI's technical assistance contractor (the Lewin Group).

Note: The blue horizontal line represents MMC's enrollment target of 500 unique participants.

MMC considered enrollees with Medicaid, including dual Medicare-Medicaid enrollees, as indirect participants. Medicaid already covers health home services like the care management services that direct participants received through the HCIA funding. Therefore, Medicaid-enrolled participants' care management services were not funded by the award. Indirect

participants benefited only from the HCIA funding used for improved IT infrastructure and staff training. MMC reported indirect enrollment data by quarter in such a way that we were unable to calculate a total count of unique indirect participants.

The analytic populations used for our FFS Medicare impact analysis and Medicaid pre-post analysis are different from MMC's direct and indirect participant populations described above. Specifically, our analytic populations are defined differently because of data limitations and exclusions:

- Our FFS Medicare analytic population includes participants for whom Medicare was the primary payer, including dual Medicare-Medicaid enrollees. (This differs from MMC's definition of direct participants, which excludes dual enrollees and includes uninsured, commercially insured, and Medicare-only individuals.)
- Our Medicaid analytic population includes participants with Medicaid as their primary payer. (This differs from MMC's definition of an indirect participant, which includes dual enrollees as well as Medicaid-only enrollees.
- We excluded from our analyses participants who received care management services through Coordinated Behavioral Care (CBC), another Medicaid health home that MMC collaborated with beginning in 2014. We excluded 2,773 CBC or dual CBC/MMC participants from the analyses—694 from the Medicare analysis and 2,079 from the Medicaid analysis—because they primarily received services outside the MMC Medicaid health home program.

4. Demographic characteristics

Table IX.1 summarizes the demographic characteristics of our Medicare and Medicaid analytic populations. At the time of their enrollment, more than two-thirds (69 percent) of the MMC Medicare participant population were age 45 or older, 48 percent were between the ages of 45 and 64, and 21 percent were aged 65 or older. The majority of participants were originally eligible for Medicare because of their disability status (88 percent), and two-thirds were dual enrollees.

To participate in MMC's HCIA program, an individual must have had a diagnosis in at least one of the following five mental health condition categories: schizophrenia and related disorders, bipolar disorders, depressive disorders, certain psychotic disorders, and childhood emotional disturbance.[51] More than half of MMC's FFS Medicare participants had a diagnosis of schizophrenia and related disorders and/or a diagnosis of bipolar disorder, and more than one-third had a diagnosis of depressive disorder (Table IX.1).

The MMC Medicaid-only population was relatively young in comparison to the Medicare or dual eligible population; almost half (43 percent) were under the age of 45 (Table IX.1). Just over half (57 percent) of these participants had a disability, which is a smaller proportion than in the Medicare analytic population (88 percent). Unlike for the Medicare analytic population, we did not require that Medicaid participants in the analytic population have a targeted diagnosis

[51] For more information about the ICD codes used to identify relevant diagnoses, please see Appendix A.

listed in their claims, although more than three quarters (77 percent) did.[52] Almost half of all Medicaid participants had a diagnosis of bipolar disorder (45 percent), and slightly fewer had diagnoses of schizophrenia and related disorders and/or depressive disorders (35 percent and 37 percent, respectively). Comorbidities with physical (non-mental) illnesses were common in the Medicaid group. For example, based on the Chronic Illness and Disability Payment System[53] algorithm, 53 percent of these participants had cardiovascular conditions, 45 percent had AIDS, 39 percent had substance use disorders, 31 percent had pulmonary conditions, and 27 percent had diabetes.

Table IX.1. Demographic characteristics of MMC Medicare and Medicaid analytic populations

	Number of Medicare participants[a]	Percent of Medicare participants	Number of Medicaid participants[b]	Percent of Medicaid participants
Total	464	100.0	5,518	100.0
Age				
18–34	62	13.4	1,323	24.0
35–44	82	17.7	1,036	18.8
45–54	116	25.0	1,855	33.6
55–64	106	22.8	1,236	22.4
65 or older	98	21.1	68	1.2
Gender				
Female	220	47.4	2,883	52.2
Male	244	52.6	2,635	47.8
Eligibility status				
Disabled	410	88.4	2,912	52.8
Dual enrolled in Medicaid	315	67.9	-	-
Mental health diagnoses[c]				
Schizophrenia and related disorders	271	58.4	1,927	34.9
Bipolar disorders	239	51.5	2,461	44.6
Depressive disorders	170	36.6	2,066	37.4
Other psychotic disorders	< 11	<2.4	217	3.9
Childhood emotional disturbance[d]	< 11	<2.4	11	0.2

Source: Mathematica analysis of data for Medicaid beneficiaries obtained from New York State Medicaid enrollment and claims data for October 2012–June 2015 and for Medicare beneficiaries obtained from the Master Beneficiary Summary File, February 2013–June 2015.

[52] All Medicare participants in our analyses were required to have at least one diagnosis in their claims records. However, this requirement was not applied to the Medicaid population. Because we could not construct a valid comparison group for the Medicaid pre-post analysis, there was no need to use diagnosis as a matching variable (which was required for the Medicare impact analysis). For this reason, the proportion of the Medicaid population with the targeted mental health conditions diagnoses is lower than for the Medicare population. Overall, 4,247 (77 percent) of Medicaid participants had at least one targeted mental health diagnosis listed in their claims files.

[53] The Chronic Illness and Disability Payment System is a diagnostic classification system that Medicaid programs can use to make health-based capitated payments for Medicaid beneficiaries who are disabled or on Temporary Assistance for Needy Families.

Note: Coordinated Behavioral Care (CBC) and dual CBC/MMC enrolled participants were excluded from the Medicare and Medicaid analyses because they primarily received services outside the MMC health home in the analysis period.

[a] Medicare participants include Medicare-only and dual Medicare-Medicaid enrollees.

[b] Dual Medicare-Medicaid enrollees are excluded from the number of Medicaid participants.

[c] Participants can have more than one diagnosis.

[d] We excluded participants with diagnoses of other psychotic disorders and childhood emotional disturbance from our Medicare analysis because there were too few of them; participants with these diagnoses were retained in the Medicaid analysis, for which the analytic population was larger.

B. Methods

1. Quantitative methods

We analyzed the program's impact on four of CMMI's core outcome measures: total Medicare or Medicaid[54] expenditures, hospitalizations, readmissions,[55] and emergency department (ED) visits. These outcomes are appropriate to use for evaluating the MMC program because its improved care coordination and management strategies were expected to reduce participants' use of acute care services and thereby reduce expenditures.

By including both Medicare and Medicaid-covered participants, we can analyze the impact of the health home and care management services funded by Medicaid and the HCIA, as well as the improved technology infrastructure and training funded solely by the HCIA, on (1) Medicaid costs and service utilization for the Medicaid-only population and (2) Medicare costs and service utilization for Medicare-only and dual Medicare-Medicaid enrollees.[56] We begin by describing our methods for the Medicare population analysis.

Analysis of the program's impact on Medicare enrollees. To conduct the impact analysis for the FFS Medicare population, we constructed both an intervention group and a comparison group and used a difference-in-differences analytic model. We included two years of baseline data and two years of intervention period data, measured in six-month intervals, for the Medicare participants in the intervention and comparison groups.

We defined the intervention group as FFS Medicare beneficiaries who enrolled in the MMC program between February 2013 and June 2015, and who had evidence of schizophrenia and related disorders, bipolar disorders, and/or depressive disorders in their Medicare claims data.[57] For purposes of this analysis, enrollees had these conditions if they had at least one inpatient or two or more outpatient Medicare claims (not including prescription drugs) with the relevant

[54] Medicaid expenditures include both fee-for-service and managed care payments. When service-level payment information is not available for managed care-covered services, these payment amounts are estimated based on fee-for-service payment guidelines.

[55] We were unable to estimate readmissions for the Medicaid population.

[56] Due to data limitations, Medicaid costs and service use for dual Medicare-Medicaid enrollees are not included in the analyses, even though dual enrollees are included in the Medicare analytic population. Although Medicare is the primary payer, the exclusion of Medicaid costs for dual enrollees means that specialized services for people with serious mental illness covered under Medicaid options and waivers provided to dual enrollees are not reflected in the analyses.

[57] Participants with diagnoses of other psychotic disorders or childhood emotional disturbance were dropped from the analysis due to small sample size.

diagnosis in the two years before enrollment (see Appendix A for more information about the ICD codes used to determine diagnoses and develop categories). In addition, we required:

- At least six months of fee-for-service (FFS) Medicare data in the year before enrolling in the program
- Six months of continuous FFS Medicare data around the enrollment month
- Participant's physical location in the MMC service area in Brooklyn, New York during the month of enrollment

We selected a comparison group with propensity score matching methods. Specifically, we took the following steps:

- We developed a comparison group of Medicare enrollees with diagnoses of schizophrenia and related disorders, bipolar disorders, and/or depressive disorders residing in three comparison cities: Philadelphia, Pennsylvania; Pittsburgh, Pennsylvania; and Chicago, Illinois.
  - We selected these three sites based on a comprehensive analysis of the relevant demographic, socioeconomic, and health care characteristics of approximately 20 of the largest urban centers in the country that are also located in states that did not implement a Medicaid health home program.
  - We only considered states that did not implement a Medicaid health home program in order to create a comparison sample that included full dual Medicare-Medicaid enrollees who were not eligible for health homes.[58]
  - To identify cities that were most similar to Brooklyn on relevant measures, we compared locations by examining the following characteristics:
    - Total Medicare spending
    - Medicare enrollees' hospital discharge rates
    - Number of all physicians and primary care physicians per 100,000 residents[59]
    - City poverty rate
    - Median household income[60]

[58] All full dual Medicare-Medicaid enrollees in New York are eligible for health homes. Given the similarity of health homes to the HCIA program, including full dual enrollees from New York in the comparison group would bias our results toward the null.

[59] Source for total Medicare spending, hospital discharge rates, and number of participants per 100,000 residents: 2012 Hospital Service Area data. Source: The Dartmouth Atlas of Health Care, The Dartmouth Institute for Health Policy and Clinical Practice, 2016. Available at http://www.dartmouthatlas.org/data/region/. Accessed August 24, 2016.

[60] Source for city poverty rate and median household income: U.S. Census Bureau Community Facts, 20102014 American Community Survey 5-Year Estimates. Available at http://factfinder.census.gov/faces/nav/jsf/pages/community_facts.xhtml. Accessed August 24, 2016.

Chicago, Pittsburgh, and Philadelphia were the only three cities that resembled Brooklyn on all measures of interest.[61]

- We identified potential comparison group members as those who resided in zip codes within the three comparison sites (Chicago, Pittsburgh, and Philadelphia), met the requirements of having Medicare as primary payer and having adequate Medicare FFS data, and had evidence of schizophrenia and related disorders; bipolar disorders; and/or depressive disorders in Medicare claims data at any time in the two years before the start of the MMC program or through the last enrollment month included in these analyses (June 2015).
- For each potential comparison group member, we created a pseudo-enrollment month that reflects the month when the member likely would have enrolled in the program if he or she had been in the intervention group. For each potential comparison group member, we identified months with visits to a primary care provider within the intervention time period (February 2013-June 2015) and then randomly selected one of the months. We took this step to ensure that potential comparison members had some engagement with the health care system as measured by a primary care visit. The random selection process assigned weights to the potential enrollment months based on the proportion of intervention participants who enrolled in the same month.
- We used propensity scores to match the intervention and comparison groups to each other. Propensity score matching (PSM) is an appropriate technique when the total number of participants in the intervention group is relatively low, as is the case with the MMC Medicare population. Instead of exact matching on individual characteristics, PSM estimates the probability of comparison group members' having enrolled in the intervention group had it been available to them, based on relevant participant characteristics. We matched up to seven potential comparison group members to each member of the intervention group.
  - We first matched intervention and comparison group participants who had identical diagnoses of schizophrenia and related disorders, bipolar disorders, and/or depressive disorders, and had the same disability status.
  - We then fit propensity score models using other important baseline characteristics predictive of intervention group status, such as demographics, chronic conditions, dual enrollment, service use, and expenditures.

Medicaid enrollee pre-post analysis. We were unable to identify an appropriate comparison group for the Medicaid-enrolled participants. All Medicaid enrollees who met the eligibility criteria for the MMC program (that is, that had a serious mental illness along with a risk for adverse events or outcomes) were eligible for Medicaid health homes. Eligible individuals who did not enroll with MMC's health home were an inappropriate comparison group for one of two reasons. First, they may have enrolled with another health home thereby receiving a similar intervention as the MMC participants. Second, they may have been contacted and refused to enroll. Individuals who refused to enroll are likely to have unobservable characteristics related to their health care needs that are substantially different from individuals

[61] The greatest difference between Brooklyn and the comparison cities was in total Medicare spending, which we expected given the known difference in Medicare spending between New York and the comparison cities. Please see Appendix A for more information about the methods we used to assess the quality of the match between intervention and comparison group participants.

who choose to enroll. Differences between states in Medicaid coverage and eligibility guidelines prevented identification of a Medicaid population outside New York with coverage comparable to that of New York Medicaid enrollees. Therefore, we conducted a pre-post analysis without a comparison group for MMC Medicaid participants.

We defined the intervention group as MMC Medicaid participants whom we could match to the New York State Medicaid data we used for this analysis, were not Medicare-Medicaid dual enrollees, and were not CBC participants. We had four years of baseline data and two years of intervention data available for the analysis. We established six-month periods for analyzing the data.

2. Qualitative methods

During our spring 2014 site visits, Mathematica conducted in-depth interviews with care managers and their supervisors, care navigators, outreach specialists, program leaders at MMC and partner organizations, and other stakeholders, such as trainers. We also conducted a total of 13 group interviews with care manager supervisors, care managers, care navigators, three outreach specialists, psychiatrists, and primary care physicians.

In spring 2015, Mathematica conducted a second round of site visit interviews with individuals in many of the same positions (but not necessarily the same people, due to staff turnover or changes in staff roles); this included 13 interviews with 28 individuals. In addition, we conducted interviews with members of the workforce whom we did not speak with in 2014, including peer specialists. We also convened a focus group with 13 other care managers, care navigators, and outreach specialists during the second site visit. During both site visits, we assessed respondents' perceptions of program impacts, implementation barriers and challenges, workforce satisfaction, and relevant internal and external context. Finally, in spring 2016, we conducted a telephone interview with MMC leaders to discuss sustainability of the program.

C. Summative findings

In this section, we present the results of the impact analysis for MMC Medicare-enrolled participants for the four core outcome measures: total expenditures, hospitalizations, readmissions, and ED visits. These outcome measures include all Medicare-covered psychiatric and non-psychiatric services provided to the analysis population, with the corresponding expenditures.[62] We also present the results of the pre-post analysis for the Medicaid-covered MMC population for three of the four core outcome measures: total expenditures, hospitalizations, and ED visits. Finally, we discuss the analytic limitations of our quantitative analyses, and compare and contrast the quantitative and qualitative findings.

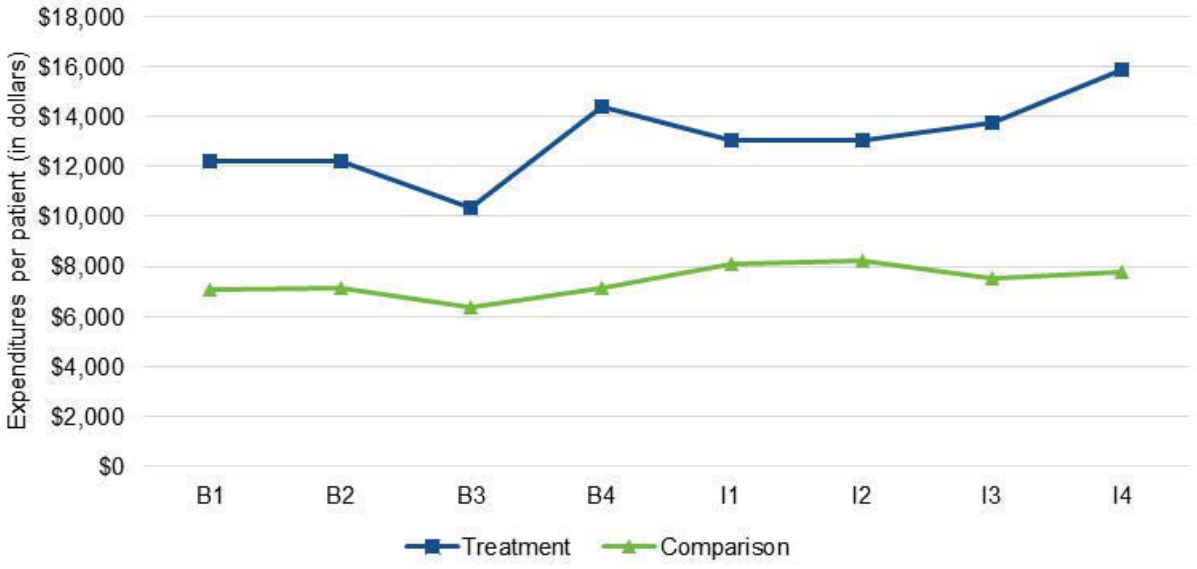
1. Impact estimates for CMMI's core measures: Medicare

Before developing the Medicare impact estimates, we analyzed the trends in means on the outcome measures for the intervention and comparison populations. We plotted the findings

[62] Medicaid costs and service utilization for dual Medicare-Medicaid enrollees are not included in the analyses, even though dual enrollees are included in the Medicare analytic population. Although Medicare is the primary payer, the exclusion of Medicaid costs for dual enrollees means that specialized services for people with serious mental illness covered under Medicaid options and waivers provided to dual enrollees are not reflected in the analyses.

from this analysis in four line graphs to provide a visual comparison of changes in outcomes over time (Figures IX.3-6) and examined the extent to which the gap between the intervention and comparison groups at each post-intervention time point was different from the average gap between these groups during baseline time periods.

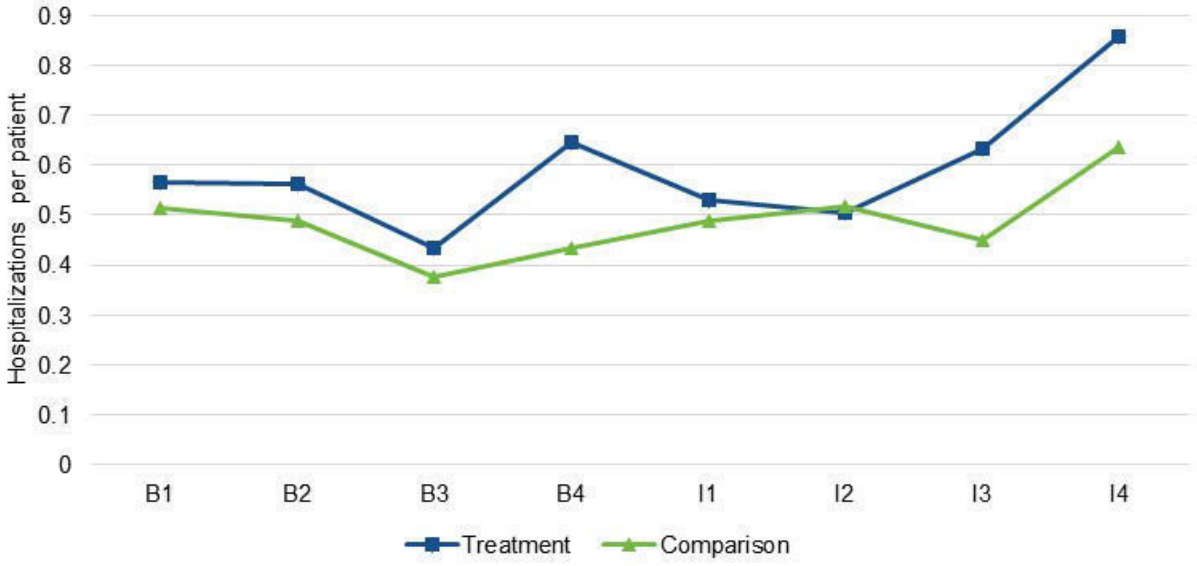
Figure IX.3. Total Medicare expenditures per patient per 6-month period: Two years prior and two years after start of intervention



Source: Master Beneficiary Summary File for baseline and program periods, February 2011–June 2015.

Note: Time periods are measured in six-month increments. I1 = months 1–6 of the MMC program, I2 = months 7–12, I3 = months 13–18, I4 = months 19–24. The baseline time periods are measured similarly: B1 = months 19–24 before the start of the program, B2 = months 13–18, B3 = months 7–12, and B4 = months 1–6.

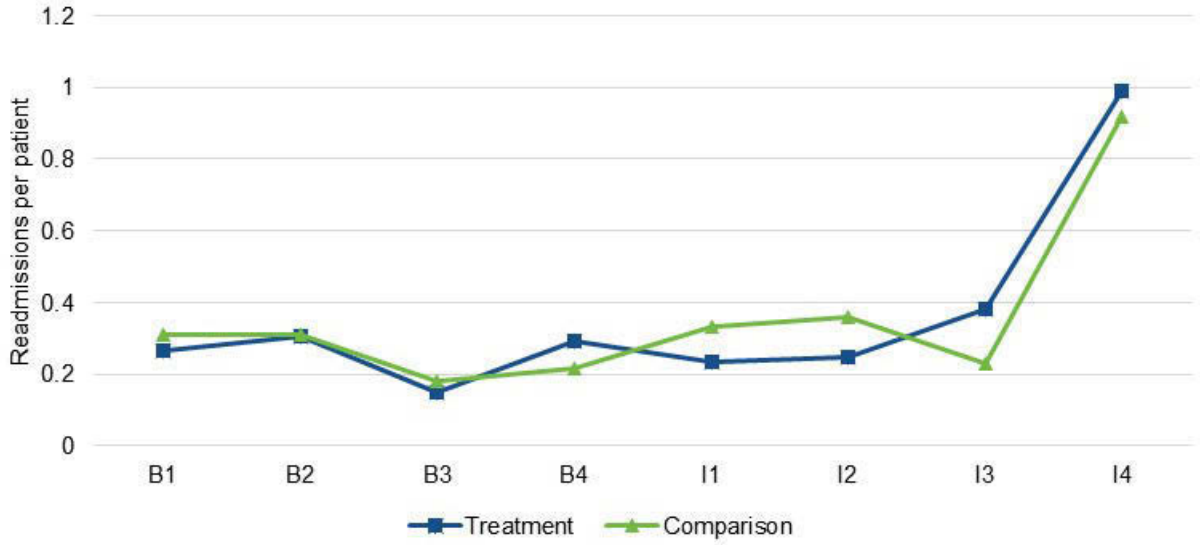
Figure IX.4. Hospitalizations per Medicare patient per 6-month period: Two years prior and two years after start of intervention



Source: Master Beneficiary Summary File for baseline and program periods, February 2011–June 2015.

Note: Time periods are measured in six-month increments. I1 = months 1–6 of the MMC program, I2 = months 7–12, I3 = months 13–18, I4 = months 19–24. The baseline time periods are measured similarly: B1 = months 19–24 before the start of the program, B2 = months 13–18, B3 = months 7–12, and B4 = months 1–6.

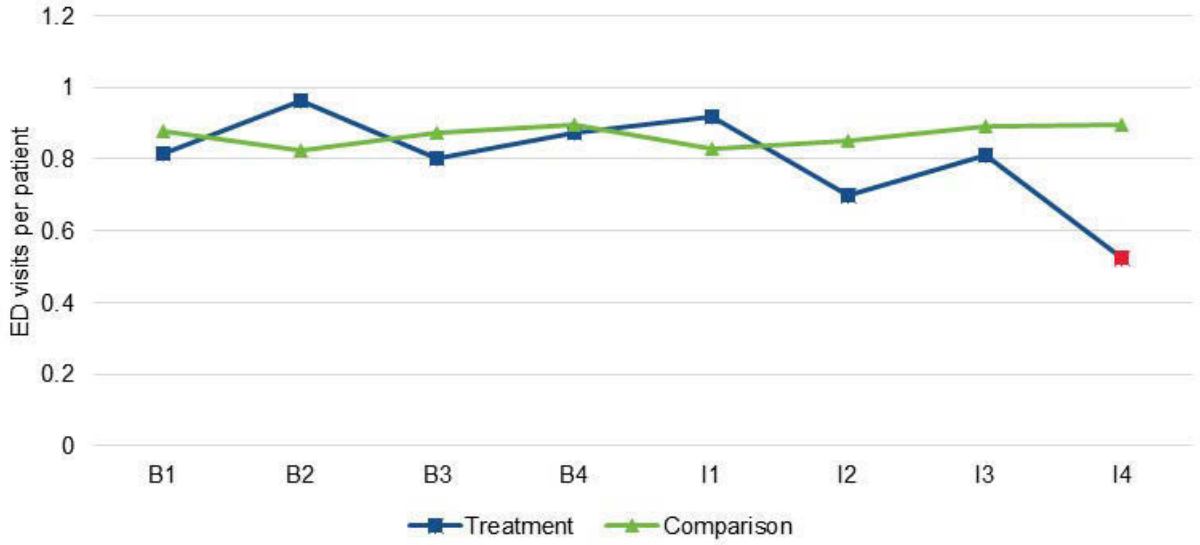
Figure IX.5. Readmissions per Medicare patient per 6-month period: Two years prior and two years after start of intervention



Source: Master Beneficiary Summary File for baseline and program periods, February 2011–June 2015.

Note: Time periods are measured in six-month increments. I1 = months 1–6 of the MMC program, I2 = months 7–12, I3 = months 13–18, I4 = months 19–24. The baseline time periods are measured similarly: B1 = months 19–24 before the start of the program, B2 = months 13–18, B3 = months 7–12, and B4 = months 1–6.

Figure IX.6. Emergency department visits per Medicare patient per 6-month period: Two years prior and two years after start of intervention



Source: Master Beneficiary Summary File for baseline and program periods, February 2011–June 2015.

Note: Time periods are measured in six-month increments. I1 = months 1–6 of the MMC program, I2 = months 7–12, I3 = months 13–18, I4 = months 19–24. The baseline time periods are measured similarly: B1 = months 19–24 before the start of the program, B2 = months 13–18, B3 = months 7–12, and B4 = months 1–6.

ED = emergency department.

As Figures IX.3-6 illustrate, the two lines follow similar paths during the baseline periods and, for the most part, during the post-intervention periods. The only statistically significant difference was for ED visits in the last intervention period (I4). We note the following for this set of figures:

- Expenditures were higher for the intervention group than the comparison group in every time period because Medicare spending in the comparison cities was, on average, lower than spending in Brooklyn.[63]
- In the final time period (I4), only a small number of participants had four full periods of program enrollment. The sharp increase in readmissions per person in the final time period may be because the characteristics of individuals enrolled for four periods were different from the characteristics in the group overall. Therefore, the apparently dramatic change in the figure is not likely to reflect a meaningful increase.
- The difference in ED visits for intervention and comparison group participants relative to the baseline average increased significantly in the final intervention time period ((I4), suggesting that the program may have cut down on participants' use of the emergency department. However, the number of participants whose data were available for analysis at this time period was small. Consequently, this finding may not be robust.

In addition to comparing the trends and differences in means over time, we also calculated the impact estimates for each outcome for the overall intervention period. Impact estimates can show if there was a statistically significant impact of the MMC program as a whole on the outcomes of interest for the participants. Our impact analysis for the Medicare participant population showed no statistically significant results (Table IX.2).

[63] Based on 2012 Hospital Service Area (HSA) data. Source: The Dartmouth Atlas of Health Care. 2016. The Dartmouth Institute for Health Policy and Clinical Practice. Available at http://www.dartmouthatlas.org/data/region/. Accessed August 24, 2016.

Table IX.2. Impact estimates for MMC Medicare enrollees

	All intervention group members	
	Change	90% CI[a]
Aggregate results		
Expenditures (in thousands)	-$26	[-$2,177, $2,125]
Hospitalizations	-39	[-158, 80]
Readmissions	-40	[-333, 252]
ED visits	-71	[-255, 114]
Per beneficiary per month[b]		
Expenditures	-$3.44	[-$289.8, $282.9]
Hospitalizations	-0.01	[-.02, .01]
Readmissions	-0.01	[-.04, .03]
ED visits	-0.01	[-.03, .02]
Number of participants	464	
Mean number of intervention months per participant	16.19	
Approximate proportion of intervention population represented in analysis[c]	8%	
Intervention period	February 2013–June 2015	

Source: Mathematica analysis of FFS Medicare claims data for baseline and program periods February 2011-June 2015.

Note: Regression model controlled for age, sex, race (white, black, Hispanic, unknown), enrollment date, SMI diagnoses (schizophrenia, bipolar, and/or depressive disorders), disability status, dual Medicare-Medicaid enrollment status, HCC condition indicators,[64] and geographic location.

[a] Because all the confidence intervals include zero, they show that none of the effects were statistically significant; therefore, any observed effects may be due to chance and not to the program.

[b] The per beneficiary per month unit of measurement is different from the time series graphs in Figure IX.4 above, which are per beneficiary per 6-month period. The differences in the direction of effect between these estimates and the time series graphs suggest that some of the apparent effects shown in the graphs are due to covariates controlled for in these statistical models.

[c] We calculated the approximate proportion of the intervention population represented in the analysis by dividing the number of participants in the Medicare analysis (464) by the total number of participants in the Medicare (464) plus Medicaid (5,518) analytic populations.

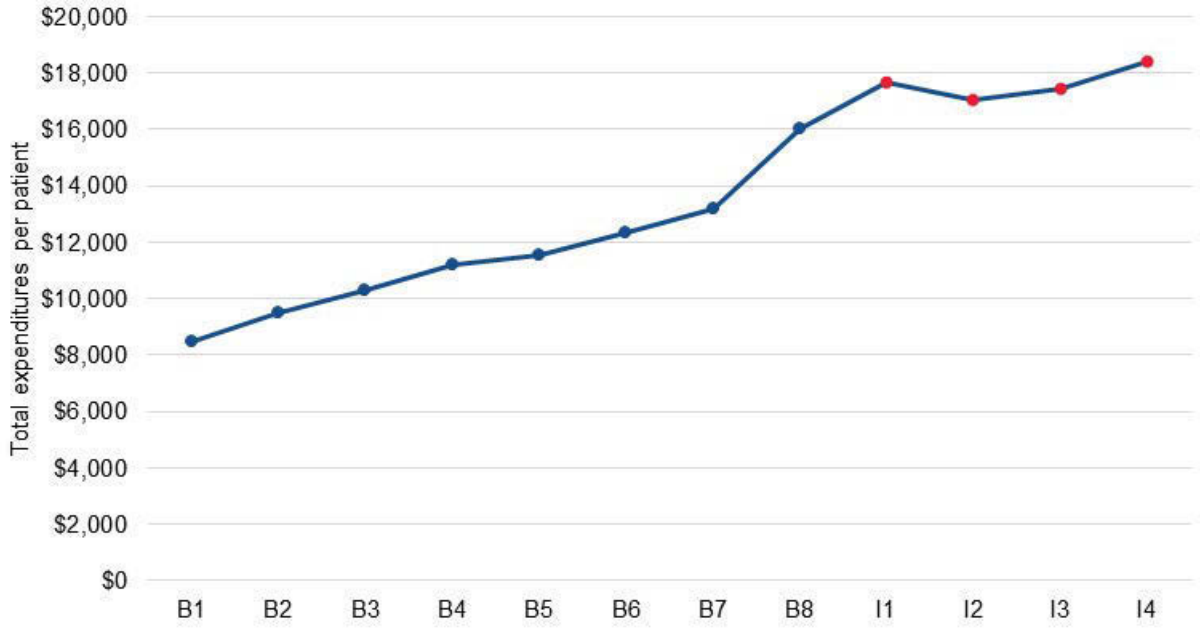
ED = emergency department.

2. Pre-post estimates for CMMI's core measures: Medicaid enrollees

Here, we present the results of the pre-post analysis with a four-year baseline period and two-year intervention period for the MMC Medicaid participants, focusing on Medicaid expenditures (Figure IX.7), hospitalizations (Figure IX.8), and ED visits (Figure IX.9). For all of these outcomes, we found significant differences at some or all of the post-intervention time periods relative to the baseline averages. Unfortunately, in the absence of a comparison group, we are unable to determine whether these differences are due to the program itself or external factors.

[64] HCC condition indicators are created as part of creating the HCC score. HCC score = Hierarchical Condition Category Score. The HCC model was developed to risk adjust Medicare payments to Medicare Advantage plans by assessing expected expenditures of enrollees. The HCC score provides a proxy of overall health status, as sicker individuals are expected to cost more than healthier individuals.

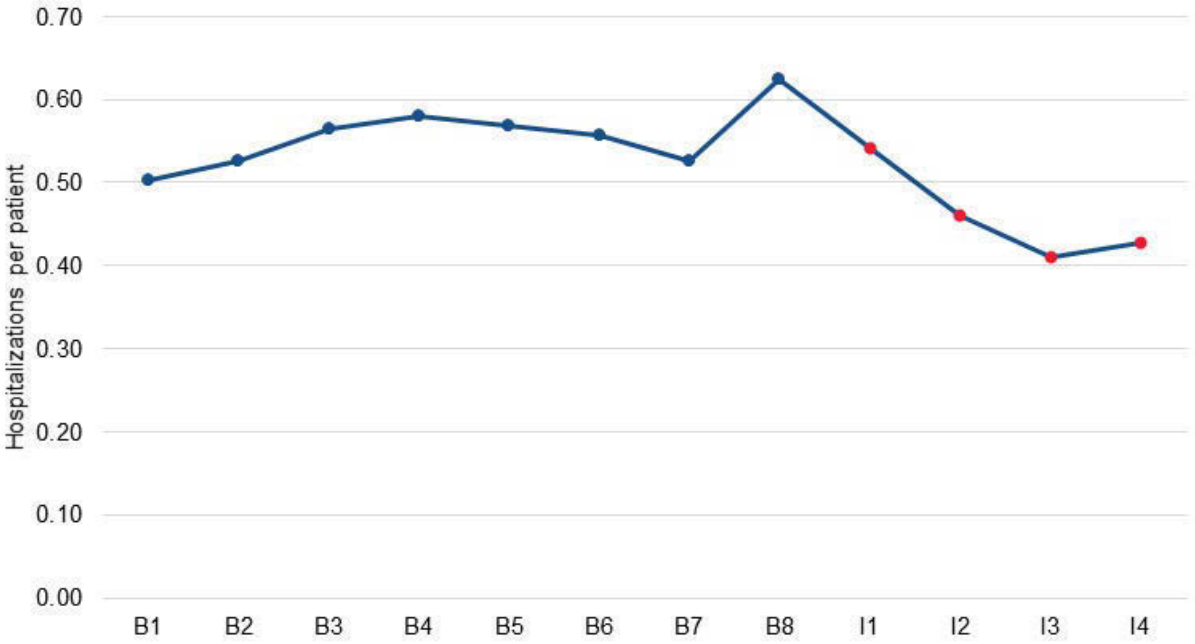
Figure IX.7. Total Medicaid expenditures for Medicaid-enrolled participants per six month period: Four years prior to and two years after start of intervention



Source: Mathematica analysis of New York Medicaid enrollment and claims data for baseline and program periods February 2009–June 2015.

Note: Regression-adjusted means for intervention population based on population characteristics in the first intervention period. Regression model controlled for age, race (white, black, Hispanic, Asian, and other), gender, disability status, SMI diagnoses (schizophrenia, bipolar disorder, depression, psychotic disorders, and disturbance of emotions specific to childhood and adolescence), Chronic Illness and Disability Payment System flags, and indicators for the calendar month and year corresponding to the first month of each six-month period.

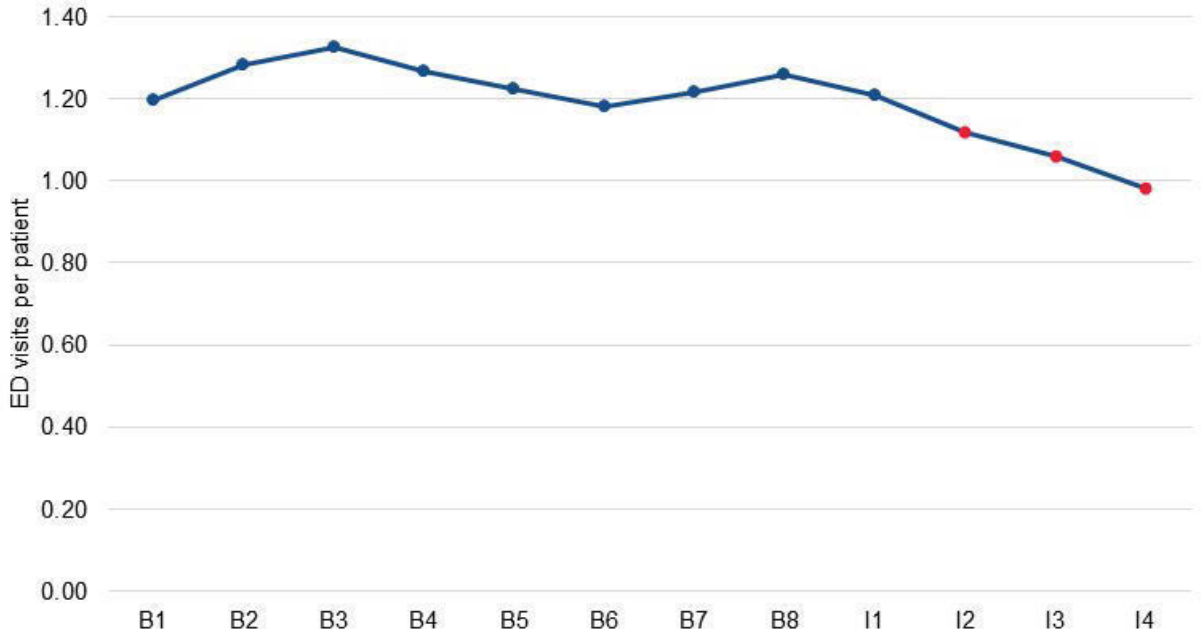
Figure IX.8. Hospitalizations per Medicaid-enrolled participant per six month period: Four years prior to and two years after start of intervention



Source: Mathematica analysis of New York Medicaid enrollment and claims data for baseline and program periods February 2009–June 2015.

Note: Regression-adjusted means for intervention population based on population characteristics in the first intervention period. Regression model controlled for age, race (white, black, Hispanic, Asian, and other), gender, disability status, SMI diagnoses (schizophrenia, bipolar disorder, depression, psychotic disorders, and disturbance of emotions specific to childhood and adolescence), Chronic Illness and Disability Payment System flags, and indicators for the calendar month and year corresponding to the first month of each six-month period.

Figure IX.9. ED visits per Medicaid-enrolled participant per six month period: Four years prior to and two years after start of intervention



Source: Mathematica analysis of New York Medicaid enrollment and claims data for baseline and program periods February 2009–June 2015.

Note: Regression-adjusted means for intervention population based on population characteristics in the first intervention period. Regression model controlled for age, race (white, black, Hispanic, Asian, and other), gender, disability status, SMI diagnoses (schizophrenia, bipolar disorder, depression, psychotic disorders, and disturbance of emotions specific to childhood and adolescence), Chronic Illness and Disability Payment System flags, and indicators for the calendar month and year corresponding to the first month of each six-month period.

ED = emergency department.

We note the following specific trends:

- Expenditures per person for the Medicaid analytic population increased steadily throughout the study period.[65] Total expenditures per patient are significantly higher relative to average baseline expenditures, but these higher values appear to be a continuation of an historical trend.
- Mean hospitalization and ED visit rates were significantly lower for all or most of the intervention time periods relative to the baseline average.

The steady increase in per capita Medicaid expenditures averaging 15 percent annually over the entire time period is striking, but its source is unclear. Changes in New York State Medicaid FFS and managed care payment rates are a contributor to the observed per enrollee expenditure increases; however, the New York state comptroller reports negative increases in per enrollee

[65] The total expenditures measure includes all types of medical claims, including pharmacy claims, which may contribute to the high expenditures per person amount shown in this figure.

spending adjusted for inflation [66] and medical inflation in the northeast region averaged only 3 percent annually in this period.[67] Thus, the rate of increase observed in the analysis population is substantially above that expected for the average NYS Medicaid enrollees in this period. The managed care spending data may not reflect actual payments to providers if the plan could not report this information. Thus, there may be some error in the expenditure reporting for managed care-covered services during the observed time period. Increased service use is another likely contributor to the Medicaid expenditure trends. Because we see an increase in expenditures but no corresponding increase in hospitalizations and ED visits, use of other services may be increasing. A substantial proportion of program participants (43 percent) are living with AIDS. Treatment for this condition may contribute to the expenditure increases. In future analyses, we will disaggregate the total expenditures measure by service type to better understand the changes in service use that may be driving increases in expenditures.

3. Analytic limitations

We note key limitations to our analyses:

- The Medicare intervention group remains relatively small (particularly in the second intervention year), making our analyses sensitive to outliers and model specifications.
- The Medicaid analysis does not allow us to draw credible conclusions because we have no comparison group, thereby preventing us from distinguishing the impacts of the program from the effect of other factors.
- Medicaid costs and service utilization for dual Medicare-Medicaid enrollees are not included in the analyses, meaning that individuals whose specialized SMI services are covered under the Medicaid options and waivers to dual enrollees are not reflected in the analyses.[68]
- We excluded participants with diagnoses of childhood emotional disturbance and psychotic disorders from our Medicare analysis because so few participants in the Medicare population had these diagnoses.
- Medicaid expenditures reported for managed care organizations may not reflect actual payments to providers if the plan was not able to report this information due to bundled or capitated payments. If the plan could not report this information, we estimated payment amounts based on the amount that would have been paid for the claim services in the state FFS system.
- We selected Philadelphia, Pittsburgh, and Chicago as comparison sites for the Medicare analysis because these cities appeared well-matched to Brooklyn, their states had not implemented a Medicaid health home program, and we could not identify any major

[66] DiNapoli, Thomas P. *Medicaid in New York: The Continuing Challenge to Improve Care and Control Costs*. Prepared by the Office of Budget and Policy Analysis, March 2015. Available at https://www.osc.state.ny.us/reports/health/medicaid_2015.pdf. Accessed August 24, 2016.

[67] Medical inflation calculated based on the medical care component of the Consumer Price Index for the northeast urban area of the United States using annual indices of the U.S. Bureau of Labor Statistics CPI Database, All Urban Consumers. See http://www.bls.gov/regions/mid-atlantic/news-release/consumerpriceindex_northeast.htm Accessed on November 27, 2016.

[68] We plan to pursue the possibility of adding the dual enrollees to the Medicaid analytic population in the future so we can calculate their Medicaid expenditures along with their Medicare-paid expenditures.

changes in behavioral health services coverage under Medicaid in the analysis period. However, although we did not identify major policy changes in these areas, we were not able to control for all possible sources of differences in trends between these cities and Brooklyn.

4. Qualitative findings on perceived effects of the program

MMC leaders and members of the workforce credited the program with numerous positive effects on participants' health, mental health, and quality of life. This perception was based on both internal monitoring data and direct interactions with program participants. Respondents told us that the program (1) reduced hospitalizations and unnecessary use of the emergency department and (2) focused attention on the social determinants of participants' health, such as housing. Staff emphasized the role of care managers in improving participants' outcomes, noting that the strong and consistent relationships care managers built with participants helped the participants maintain accountability and investment in their own health and well-being. In addition, respondents thought the care management staff played a key role in improving participants' outcomes by coordinating with other providers and helping participants find and maintain housing and other social supports.

D. Findings about the workforce

Staff in new care management roles, whether newly hired or existing employees, received the same standardized training, designed and provided by the 1199 SEIU Training and Upgrading Fund, to develop core competencies. Program staff generally found the trainings useful for their work, citing the trainings' interactive nature along with MMC's responsiveness to training needs throughout the workforce.

Most staff said they felt supported in their roles. However, our interviews and focus groups with program staff in 2014 and 2015 revealed that staff perceptions about the level of support varied from one partner organization to the next, and this disparity was at least partly attributed to the partner organizations' different leadership. For example, one staff member mentioned hearing about the resources of care management staff at other MMC partner organizations, noting, "I'm like, wow, you guys have got cell phones and tablets? We've got pencils and pads." Some staff linked the disparity in resources to the level of support they perceived from upper leadership at their organization. They thought that leaders with clinical experience were more likely than those without it to strongly support and advocate for staff.

The majority of staff said they were satisfied in their job. Staff interviewed in 2014 and 2015 almost universally reported that the most satisfying part of their job was seeing the improvement in participants' behavioral and physical health and in their level of social support. However, some staff noted high caseloads, coupled with frustration about limited resources at some partner organizations, as factors that could contribute to emotional exhaustion and dissatisfaction. This observation suggests that slight differences in the way MMC partners implemented the program could influence staff's level of satisfaction and sense of support.

The workforce respondents identified several major barriers to job performance, including characteristics of the target population, housing inaccessibility, and challenges with the program's documentation requirements in the CCP. For example, many interview and focus

group respondents indicated that the CCP was cumbersome and lacking in features that would make it useful to staff. Although many staff recognized the promise of the virtual tool, many also mentioned that the platform would be more useful if it incorporated some improvements—several of which, such as a provider landing page, were under development. In addition, many staff members thought the documentation required for the program was burdensome, particularly for care management staff. Many primary and specialty care providers did not use the CCP; as a result, the care management staff had to enter much of the documentation in the CCP.

E. Program sustainability and spread

MMC is using a variety of funding sources to sustain and expand all its program components. The awardee has expanded access to (1) the virtually co-located health home services developed under the HCIA, (2) the care coordination platform and dashboard enhanced through award funding, and (3) the training program and care standards developed under the award. Services will continue to be available to both Medicaid enrollees and the small number of participants whose care management services were covered by the award.

During our 2015 site visit, program leaders identified funding obtained through the projects they developed under the state's DSRIP initiative as the most promising source of funding for continuing and expanding the program. This plan was particularly relevant for the Medicare beneficiaries whose care management services were paid for by the HCIA. During a 2016 follow-up telephone conversation about sustainability, however, MMC leaders told us that, although DSRIP presents a good source of contingency and wraparound funding, MMC's HARP pilot best aligns with the goals of their HCIA program given its focus on those with SMI.

The HARP pilot resulted from state Medicaid reform and is designed to serve individuals with serious and persistent mental illness—MMC's HCIA target population. In contrast, DSRIP's goal is to better serve all Medicaid beneficiaries with mental illness. Through the HARP pilot, MMC and its partners have carved all of the core services coordinated by their HCIA intervention into their HARP, and have included outreach efforts by both health home care managers and managed care managers, ongoing case conferencing, and sharing the CCP.

HARP has made it possible for MMC to keep serving its Medicaid population (that is, its indirect participants under the HCIA), and the awardee has used other funding opportunities, such as health home development funds, to keep serving those who do not have Medicaid. Program staff hope that ultimately, the care managers and navigators will help many individuals enroll in Medicaid. Program leaders have used promising results from internal cost modeling and self-monitoring in their negotiations with the state about sustaining the program components for both direct and indirect participants.

As noted, external factors, such as the critical shortage of affordable housing, also influence the organization's ability to sustain the program and continue improving care for its patients. Program leaders acknowledged that Medicaid reform is not a silver bullet because it is primarily focused on beneficiaries' medical needs. Social determinants of health, such as housing and food security, are equally significant and can have profound effects on client outcomes and the cost of care. Medicaid reforms likely will provide limited opportunities to address broader social service needs that are critical to recovery. Instead, MMC hopes to use DSRIP funding to leverage capital to build housing and supportive housing.

F. Lessons learned

MMC built sustainability efforts into its operational plan and implementation from the outset. That, coupled with the organization's awareness of and ability to take advantage of state-level policy opportunities and changes, allowed the organization to use the HCIA as one piece of a financing puzzle to sustain its efforts. In addition, leaders said that positive findings from internal monitoring and financial modeling, both built into the program from its inception, have allowed the organization to make a case for its continuation to the state and other funders.

MMC staff noted that their internal analysis of Medicaid data showed a 28 percent decline in hospitalizations and a 2 percent decline in ED use, with estimates of about $51.8 million in savings over three years. We also found that hospitalizations and ED visits declined in the post-intervention period but expenditures increased for Medicaid participants. We are conducting additional analyses to identify the types of expenditures that might be driving the increases.[69]

MMC's analysis did not address Medicare expenditures or service use. Our analyses found no significant impacts of the program on expenditures or service use for participants enrolled in FFS Medicare.

MMC's HCIA program benefited from the significant amount of pre-award work the organization had already done to engage partners, thus allowing the awardee to implement its program relatively smoothly across a range of different providers. As one program leader noted, MMC's history of learning from and understanding the perspectives of partners gave it a head start at the outset of HCIA program implementation. MMC gave the partners some flexibility in how they implemented the program, allowing them to fit the program into their own partners' unique organizational structures. However, the focus groups and interviews with the workforce revealed that the different ways the program was implemented by MMC's partners may have influenced staff's levels of satisfaction and feelings of support.

As integral as it was to the program's success, the CCP by itself did not provide the mechanism for helping providers work together to coordinate services. Primary and specialty care providers did not use it in the way that MMC staff originally expected them to; instead, care management staff, rather than providers, served as the primary conduit to the platform and entered new information as available. This change illustrates MMC's flexibility and responsiveness to providers' needs over the course of the program while trying to make the CCP as useful as possible. The CCP's continuing evolution and enhancements have led other health homes in the area to adopt the platform.

[69] Because we have no comparison group, our Medicaid analysis does not distinguish the effect of the program from other factors. Thus, the utilization and expenditure trends observed in the post-intervention period may not be attributable to MMC's program.

This page has been left blank for double-sided copying.

X. VALUEOPTIONS

Findings from Mathematica's Evaluation of the ValueOptions HCIA Program

- The Massachusetts Behavioral Health Partnership tested a number of changes to its Community Support Program, which is for patients who have been discharged from detoxification facilities. The changes included additional training for staff, a shift to a case-rate payment model, and incentives for some patients. The goals of the program were to reduce addictive behaviors, enhance overall health, and improve the patient experience.
- Although early quantitative results were promising, we have not yet received all of the data necessary to evaluate the program. We expect to provide quantitative findings in an addendum to this report in early 2017.
- Our preliminary findings showed that (1) payers are uniquely positioned to create sustainable programs, but not without challenges and (2) training for paraprofessionals might be an effective way to improve workforce satisfaction.

A. Introduction

1. Program goals

The Massachusetts Behavioral Health Partnership (MBHP), a company owned by ValueOptions[70] that contracts with the Commonwealth of Massachusetts to manage behavioral health benefits for Medicaid beneficiaries, used funding from the Health Care Innovation Awards (HCIA) to test the effectiveness of three modifications to its Community Support Program (CSP). In this program, staff work with clients to help them access and coordinate medically necessary services and other community-based support services. As part of the HCIA-funded program, the title used for a subset of CSP staff was changed to recovery support navigators (RSNs). For the HCIA program, MBHP (1) trained RSNs on evidence-based treatment for substance use disorders, readiness-to-change assessments, and motivational interviewing; (2) covered RSN services through a case-rate model (a fee-for-service model is used in the CSP); and (3) offered a subset of participants in the RSN program incentives (gift cards) for achieving goals related to their recovery. MBHP hypothesized that these changes to the existing CSP model would lower costs by decreasing the repeated use of detoxification services. As part of the HCIA intervention, RSNs worked with participants who had been admitted to detoxification facilities at least twice. The Brandeis University Institute for Behavioral Health partnered with MBHP to conduct a local evaluation of the program.[71]

MBHP implemented the RSN program at four Massachusetts detoxification facilities that employ and supervise the RSNs: (1) Lahey Health Behavioral Services, (2) Stanley Street Treatment and Resources, (3) High Point Treatment Center, and (4) Spectrum Health Systems.

[70] ValueOptions and Beacon Health Strategies merged in 2015 to become Beacon Health Options. However, we refer to the company as ValueOptions throughout this report.

[71] ValueOptions received its HCIA award in July 2012 and began enrolling participants in early 2013. The awardee received a no-cost extension from CMMI through December 31, 2015, to continue evaluation activities.

When individuals were discharged from these facilities, RSNs enrolled them in the RSN program and assigned them to one of two groups:

1. **RSN+I.** Participants in this group were offered RSN support. In addition, incentive payments were given to participants who achieved specific recovery goals.
2. **RSN only.** Participants in this group were offered only RSN support.

MBHP staff assigned eligible individuals to these groups at the facility level by using a midpoint crossover design.[72] For example, during the first half of the program period, all eligible individuals discharged from the Lahey facility were assigned to the RSN+I group. During the second half of the program, individuals discharged from Lahey were assigned to the RSN-only group. At any point in time, eligible individuals from two designated facilities were assigned to the RSN-only group, while eligible individuals from the other two facilities were assigned to the RSN+I group.

There are nine non-intervention detoxification facilities in the MBHP system. All of these facilities provide the CSP.

In the short-term, MBHP expected the RSN program to improve (1) participant engagement with community-based supports and (2) participant attitudes about recovery. In the long-term, MBHP expected the program to reduce addictive behaviors, enhance overall health, and improve the patient experience.[73]

Our study findings are based on quantitative and qualitative data, as well as enrollment data. We draw on the following data sources for this chapter:

- **Program enrollment data** provided by MBHP, including information on enrollment dates and demographics for participants and members of the comparison group.
- **Data on baseline assessments of substance use and health status** among participants and comparison group members, provided by MBHP.
- **Data provided by MBHP**, including Medicaid eligibility data; medical, pharmacy, and dental claims; and information on behavioral health encounters.
- **Data from workforce surveys** that Mathematica conducted in 2014 and 2015.
- **Qualitative data**, including telephone interviews and in-person site visits in 2014 and 2015. Mathematica conducted in-depth interviews with awardee leadership, members of the workforce, and other stakeholders. In addition, we convened focus groups with members of the workforce, program participants at the RSN sites, and CSP participants at the

[72] Individuals were eligible to be enrolled as participants or as members of the comparison group if they (1) were between the ages of 18 and 64; (2) were enrolled in Medicaid; and (3) had been admitted to and discharged from a detoxification facility at least twice in the year before enrollment.

[73] A comparison group used for this study included individuals discharged from seven of the detoxification facilities that offer the CSP.

comparison sites. In spring 2016, we conducted a telephone interview with awardee leaders to discuss program sustainability.

2. Overview of administrative context

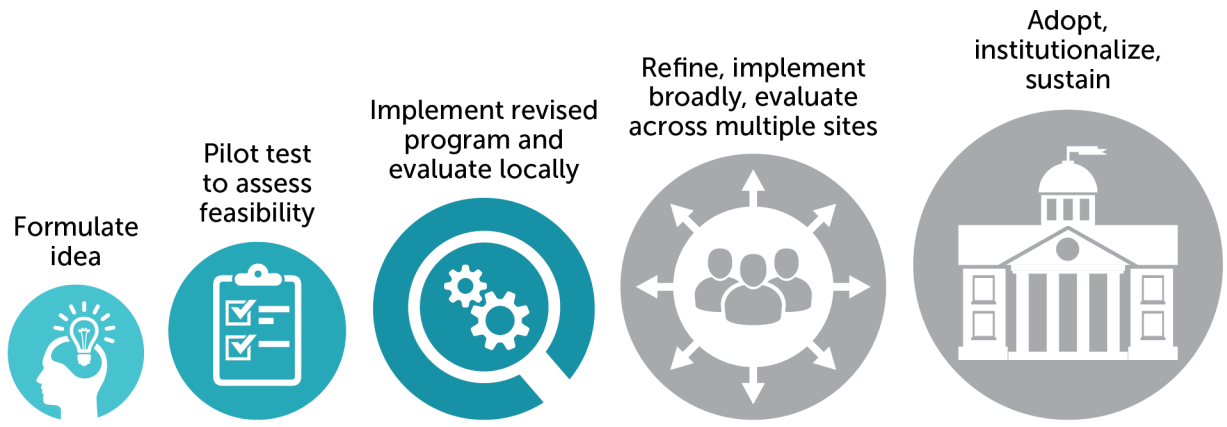
As a payer, MBHP had unique advantages and challenges in creating the RSN program. For example, MBHP was able to adjust the payment model of the CSP and also sustain this change without outside funding. However, MBHP's role as an insurance payer and its consequent lack of control over staffing at treatment sites presented challenges. MBHP leaders believed the program could have served more participants if the implementing sites had hired more RSNs. However, some sites were hesitant to hire RSNs before the need for them was demonstrated.

Working with a large number of sites also presented challenges for MBHP. Although the sites provided similar services, they had different management styles and organizational structures. This made it difficult to completely standardize the services provided to clients, which was needed for the Brandeis research study. However, working with a large number of providers also had some benefits. For example, one respondent indicated that the Brandeis study increased the visibility of both the RSN and CSP programs in the MBHP network, which renewed overall interest in both programs among detoxification facility clients.

3. Progress through the phases of innovation

Program innovation can be understood as a process that includes five distinct but fluid innovation phases (depicted in Figure X.1). During the HCIA-funded program, MBHP's RSN program progressed through the first three phases shown below: (1) MBHP created a new program, (2) MBHP implemented the program at selected sites, and (3) MBHP conducted research to compare outcomes across the sites. MBHP is analyzing data from the research study and will decide whether to progress to the next phases of innovation once those data are available.

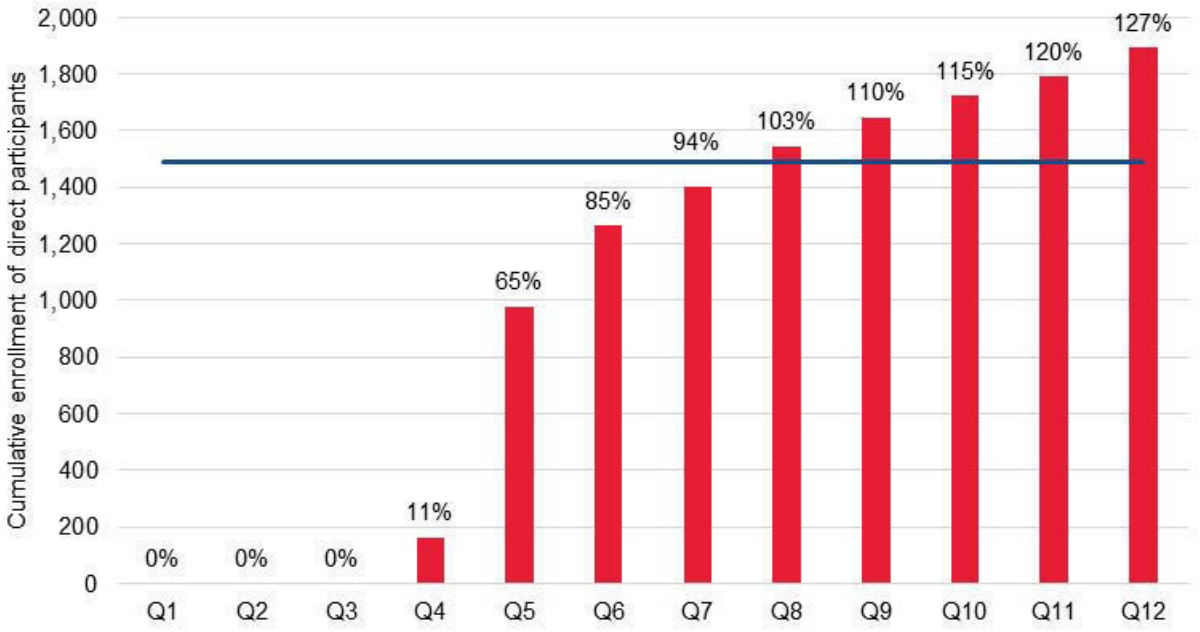
Figure X.1. Phases of program innovation: ValueOptions



4. Enrollment

By the end of the 12th quarter (June 30, 2015), MBHP had enrolled 1,893 direct participants in the program—exceeding its original goal of 1,492 participants by over 25 percent (Figure X.2). Table X.1 provides demographic information about MBHP participants who were eligible for inclusion in the core measures analysis.

Figure X.2. Percent of target enrollment achieved by quarter, Q1–Q12



Source: Awardee's enrollment data reported to the website maintained by CMMI's technical assistance contractor (the Lewin Group).

Note: The blue horizontal line represents ValueOptions' enrollment target of 1,492 unique participants.

Table X.1. Demographic characteristics of ValueOptions participants, March 2013–December 2015

	RSN		RSN+I		CSP	
	Number	Percent	Number	Percent	Number	Percent
Total	684	100	587	100	347	100
Age						
18–34	324	47.4	321	54.7	158	45.5
35–44	180	26.3	134	22.8	82	23.6
45–54	134	19.6	110	18.7	83	23.9
55–64	46	6.7	22	3.7	24	6.9
Gender						
Female	250	36.5	203	34.6	111	32
Male	434	63.5	384	65.4	236	68

Source: Mathematica analysis of electronic medical record data, March 2013–December 2015.

Note: This table includes participants who were eligible for inclusion in the core measures analysis and who had non-missing demographic data.

CSP = Community Support Program; RSN = Recovery Support Navigator program; RSN+I = Recovery Support Navigator with incentives

B. Methods

We are using a mixed-methods approach to evaluate MBHP's RSN program. As we described in the second annual report, preliminary quantitative data showed promising results, particularly in the areas of emergency department (ED) visits and expenditures. However, it is not possible to determine the impact of the program without final quantitative data. MBHP was unable to provide complete data in time to be included in this report.[74] We anticipate that we will receive complete data during summer 2016. We expect to analyze these data for an addendum to this report, which we will provide in early 2017. The addendum will include results based on the approach described below.

The quantitative analysis, to be presented in the addendum in early 2017, will rely on four data sources: (1) MBHP-provided program enrollment data, (2) baseline assessments of substance use and health status among participants and comparison group members, (3) MBHP-provided Medicaid eligibility and claims data, and (4) a workforce survey conducted by Mathematica. The analysis will include descriptive statistics related to enrollment as well as an impact analysis. This impact analysis will rely on a difference-in-differences approach to compare participants and comparison group members (CSP site enrollees).

In the second annual report, we used a difference-in-differences approach to analyze the following measures: (1) ED visits, (2) total expenditures, (3) short-term residential treatment stays, and (4) intensive day treatments. In the 2017 addendum, we will calculate these measures using the updated data. We will also expand the analysis to include an estimate of the impact of the program on initiation of and engagement with treatment for substance use disorder. In our analysis for the second annual report, we found that hospitalization rates were low among both participants and comparison group members. Therefore, for the addendum, we do not anticipate that we will have a sufficient sample size to examine changes in hospitalizations or readmissions.

Although we are unable to report on program impacts in this report, we can provide insights into program implementation from our analyses of qualitative data—including, information from site visits conducted in 2014 and 2015, as well as a telephone interview in 2016 on program sustainability. During both site visits, we conducted in-depth interviews with program and internal evaluation leaders, RSNs and their supervisors, and the trainers who worked with RSNs as part of the innovation. Interview topics included implementation effectiveness, the target population, perceived program effectiveness, spillover effects, program costs and expenditures, workforce development and experience, project leadership, partnerships, and policy implications of the awardee's work. The second site visit was similar to the first with two major differences: (1) we added interview questions about program sustainability and (2) we conducted focus groups with RSNs, participants who received RSN services, and comparison group members. Focus group discussion topics included experiences and satisfaction with the RSN program and the CSP, as well as on the perceived effectiveness of the programs.

[74] Although MBHP indicated that it would be able to provide data in April 2016, it did not provide data until early May. Furthermore, these data were incomplete. We are working with MBHP to obtain the missing information and to resolve questions related to the quality of the data before beginning the analysis.

C. Summative findings

As explained above, we will provide quantitative results for core measures in the 2017 report addendum. In the addendum, we will also provide insights into whether and how the qualitative findings below align with the final quantitative findings.

Overall, most of the program leaders and frontline staff we interviewed believed that the RSN program had positive effects on health outcomes and quality of life for participants, without causing any negative effects. RSN supervisors also indicated that the training RSNs received on motivational interviewing may have been particularly helpful in improving participant outcomes. They said the training helped staff to better understand their roles and more actively help participants achieve recovery goals. One supervisor explained, "Beforehand, we saw [the RSN role] more as mobile case management. Now, we see it more as mobile treatment. . . . I think that has been more effective and a big difference between the two."

Respondents indicated that the case rate provided an opportunity to provide participants with more services and to sustain the services over time, which may improve participant outcomes. The case rate also allowed RSNs to account for time spent on training and other professional development activities. However, it took time for RSNs to adjust to the new payment model. In addition, because MBHP was not the only payer for these types of services, RSNs at many sites continued to provide CSP services for other payers. This meant that they needed to juggle two payment models at once, which was difficult and created stress for the RSNs.

Respondents gave mixed assessments of the usefulness of the financial incentive component of the program. Many RSNs thought the incentives did not benefit participants. However, some described the structured nature of the incentive program as potentially beneficial for some participants.

D. Findings about the workforce

The RSNs received two types of training: (1) skills training in areas such as motivational interviewing and (2) topic-specific trainings in subjects that included helping clients to find housing. RSNs and their supervisors generally found the trainings helpful and believed they enhanced skills that benefitted clients. In response to the 2015 workforce survey, most staff (79 percent) indicated that the trainings they received in the past 12 months were either very or somewhat useful for their work. Only 4 percent said the trainings were not very useful. This pattern was similar to the one found in 2014 survey results, and it was consistent with feedback from staff interviews in both years.

Many of the RSNs did not have an advanced degree or formal training in working with individuals with substance use disorders. Therefore, RSN supervisors considered training to be invaluable for this workforce. In turn, RSNs appreciated the investment in their professional development. As one RSN noted, "I like the fact that [the] RSN [position] was based on . . . constant education, and I think that's something that needs to be pushed further." RSNs also valued the topic-specific trainings and indicated that trainers were responsive to requests for additional training on topics of interest.

RSNs were deeply invested in their clients; this contributed to both their satisfaction and stress levels. Although the job is difficult, it can be incredibly rewarding when clients do well. However, because it is inevitable that some clients will relapse, workforce members felt that they needed more training on how to take care of themselves while performing this difficult work. RSNs also suggested that opportunities to gather in groups at the site level to discuss difficult cases and to participate in more hands-on trainings would have been helpful. For example, one hands-on training was held at an ED. RSNs shadowed a doctor during this training. They felt that it was particularly valuable to see the skills that they were learning being implemented by an expert.

RSNs also reported receiving the support they needed to implement the program successfully. They indicated that MBHP was responsive to questions and attentive to their needs. For example, many RSNs explained that they felt comfortable calling MBHP staff when they had questions about paperwork related to the program and that they always received the support they needed. This helped to ensure that the program was implemented effectively. However, many RSNs reported that their demanding jobs were made more difficult by the paperwork required for the research study led by Brandeis. Many RSNs felt that this paperwork was repetitive with other tasks. However, they noted that this situation improved somewhat after MBHP hired a research assistant in June 2014.

E. Program sustainability and spread

MBHP has sustained the RSN case rate at the intervention sites. It will continue to sustain this aspect of the program at least until the final results from its internal evaluation are available. If the study shows that the case rate improves outcomes, it will be sustained; if the study shows otherwise, MBHP will return the intervention sites to the fee-for-service payment model. Similarly, the RSN training is on hiatus until the impact results become available. If it is reinstated, it will likely be done in a less expensive and less time-intensive way (for example, through webinars instead of on-site trainings).

MBHP ended the RSN+I incentive program because of low uptake among participants. There are no plans to reinstate it. There were several challenges to implementing this portion of the program. For example, many of the facilities where the participants lived, such as halfway houses, had policies that prevented participants from receiving the incentives.

F. Lessons learned

While preliminary quantitative results reported in the second annual report were promising, we cannot fully evaluate the MBHP program without the quantitative data that we expected to receive this summer. However, we can draw two preliminary conclusions from the quantitative, qualitative, and survey data that we have analyzed to date:

1. **Payers are in a unique position to create sustainable programs.** Payers like MBHP are uniquely situated to both create and sustain programs because they can choose to implement and sustain innovations without a need for outside funding. However, if only one payer among many changes a payment model, it can create new challenges for frontline staff, who may struggle with working under several different payment structures at once.

2. **Investing in a workforce that has minimal formal training may be an effective way to improve workforce satisfaction.** Both surveys and qualitative data indicate that the RSN workforce appreciated the training provided as part of the innovation. RSN supervisors perceived that this training increased RSN effectiveness. The final quantitative data may help us determine if this was the case.

XI. VINFEN

Findings from Mathematica's Evaluation of Vinfen's HCIA Program

- We were unable to conduct a rigorous impact evaluation of Vinfen's HCIA program because insufficient data and other limitations kept us from identifying an appropriate comparison group.
- Qualitative findings suggest that participants made positive changes in health behaviors, demonstrated increased awareness of health problems, and improved appropriate use of health and community services.
- Workforce staff identified frequent and ongoing communication, sufficient training on working with a high-needs population, and adequate supervision and support as important factors that influenced program implementation and the satisfaction and retention of its workforce.
- Vinfen is sustaining several aspects of the program, including integrating health outreach workers and nurse practitioners in behavioral health home outreach teams. This occurred partly because of Vinfen's efforts to build on existing state efforts and initiatives, develop and maintain strategic partnerships and collaborations, and draw on the support and buy in from its senior leaders.

A. Introduction

Vinfen, a community-based provider of behavioral health services, used Health Care Innovation Awards (HCIA) funding to develop a behavioral health home that integrated primary health and behavioral health care services for adults with serious mental illness who lived in the metropolitan Boston area. This program added new components to existing clinical teams and introduced self-management techniques for participants. Overall, Vinfen's goal was to reduce the participants' use of medication and acute care services, thereby reducing the cost of their care.

Vinfen's program had four key components:

1. Addition of a nurse practitioner to four existing psychiatric rehabilitation outreach teams (one Vinfen team and one team in each of three partner locations) to coordinate services and provide primary care
2. A health outreach worker (HOW), who was added to each team to provide outreach, self-management training, and other health services
3. The Health Buddy telehealth system, a small device used by a subset of participants, which allowed the nurse practitioner and the HOW to monitor individuals who needed ongoing medical attention
4. The Integrated Illness Management and Recovery (IIMR) curriculum, which the HOWs used to train the participants in self-management strategies designed to help them manage their symptoms and improve their overall health

Vinfen implemented the program in partnership with three other community-based behavioral health providers: Bay Cove Human Services, North Suffolk Mental Health, and Brookline Mental Health. Each organization had several psychiatric rehabilitation outreach teams and selected one team to implement the four program components. The Commonwealth Care Alliance (CCA), a nonprofit Medicaid managed care entity, provided and supported the

four nurse practitioners who were assigned to the program's teams. The teams served a mix of pre-existing and new clients.

Vinfen also partnered with research staff from Dartmouth College who developed the IIMR curriculum, helped train and support the HOWs in using the curriculum, and analyzed the program's effect on health outcomes. Bosch Healthcare provided the Health Buddy telehealth system as well as ongoing technical assistance and support. JEN Associates partnered with Vinfen on a cost analysis using Medicaid and Medicare claims data.

We draw on the following data sources for this report:

- Enrollment data submitted by Vinfen to the reporting website maintained by CMMI's technical assistance contractor (the Lewin Group)
- Qualitative data collected by Mathematica during site visits and telephone interviews through June 2016. We conducted in-depth interviews with key awardee staff, members of the workforce, and other stakeholders. We also convened focus groups with members of the workforce and with program participants.
- Quantitative data collected by Vinfen staff through regular assessments and submitted to Mathematica in July 2015. As described below, we used these data to describe the demographics and level of intervention use among participants. We were unable to conduct an impact analysis because insufficient individual-level data and other data-related limitations prevented us from identifying an appropriate comparison group.

This report presents a comprehensive and final summary of the findings of our evaluation of Vinfen's HCIA program.

1. Overview of administrative context

Vinfen is a community behavioral health organization that provides mental health services, intellectual and developmental disability services, and brain injury services to adolescents and adults throughout eastern Massachusetts. For the HCIA program, Vinfen sought to address the need for better coordination between medical and behavioral health care among individuals with serious mental illness because they are known to have higher rates of chronic physical health conditions and utilization of medical services. The program had three primary goals:

1. Improve the self-efficacy and self-management of health for participants with serious mental illness
2. Integrate primary care services into behavioral health outreach teams
3. Reduce unnecessary use of acute care services

In the longer term, Vinfen hoped that health plans throughout the state would adopt its HCIA program components, particularly health plans that serve the Massachusetts One Care program for dually eligible individuals.[75]

Vinfen built its HCIA program into the organization's existing Community Based Flexible Support (CBFS) program, which is a service funded by the Massachusetts Department of Mental Health (DMH). CBFS psychiatric rehabilitation outreach teams, which typically include mental health outreach workers and social workers, offer community-based rehabilitative interventions and supports to help adults with serious mental illness manage psychiatric symptoms, restore or maintain independent living, and promote physical and emotional wellness.

For the HCIA program, Vinfen partnered with three other Massachusetts-based community behavioral health organizations that had existing psychiatric rehabilitation outreach teams. Each organization incorporated a nurse practitioner or a registered nurse and a HOW into one of its existing outreach teams.[76] These organizations hired and supervised the HOWs but CCA hired and supervised the nurse practitioners. The nurse practitioner and HOW worked together on each outreach team to support HCIA participants' use of the Health Buddy system and IIMR curriculum.

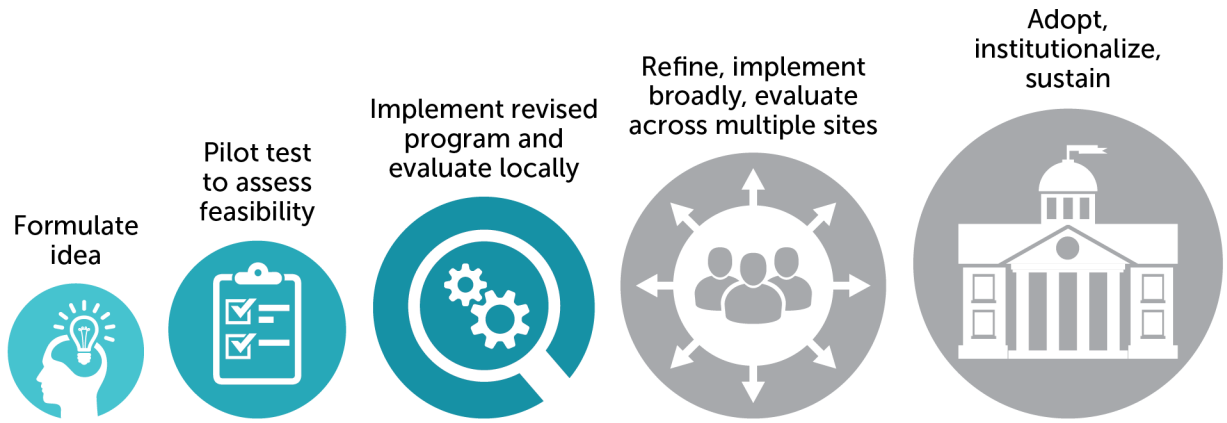
2. Progression through phases of innovation

Program innovation can be understood as a process that includes five distinct but fluid innovation phases (depicted in Figure XI.1). During the HCIA, Vinfen's program progressed through the first three phases shown below. Vinfen developed the program for the HCIA application and HCIA funds allowed Vinfen to pilot test it in the four sites and conduct research on its outcomes.

[75] The One Care program offers several health plan options to help dual Medicare-Medicaid enrollees between the ages of 21 and 64 manage their care across the two programs. One Care enrollees receive support from a dedicated care coordinator and coverage for additional community-based behavioral health services and other community support services.

[76] In the second year of the project, Vinfen modified its outreach team structure to replace nurse practitioners with registered nurses (RNs) on two of the four outreach teams. This modification was due, in part, to staff turnover among the original pool of nurse practitioners. Vinfen decided not to replace nurse practitioners who left CCA during the last six months of the project; for some outreach teams, Vinfen used RNs CCA hired. Vinfen assigned RNs to outreach teams that primarily served participants who were already well connected to a primary care home and placed the nurse practitioners on teams serving higher-needs populations—that is, those more likely to live in a group home and to lack consistent access to primary care services.

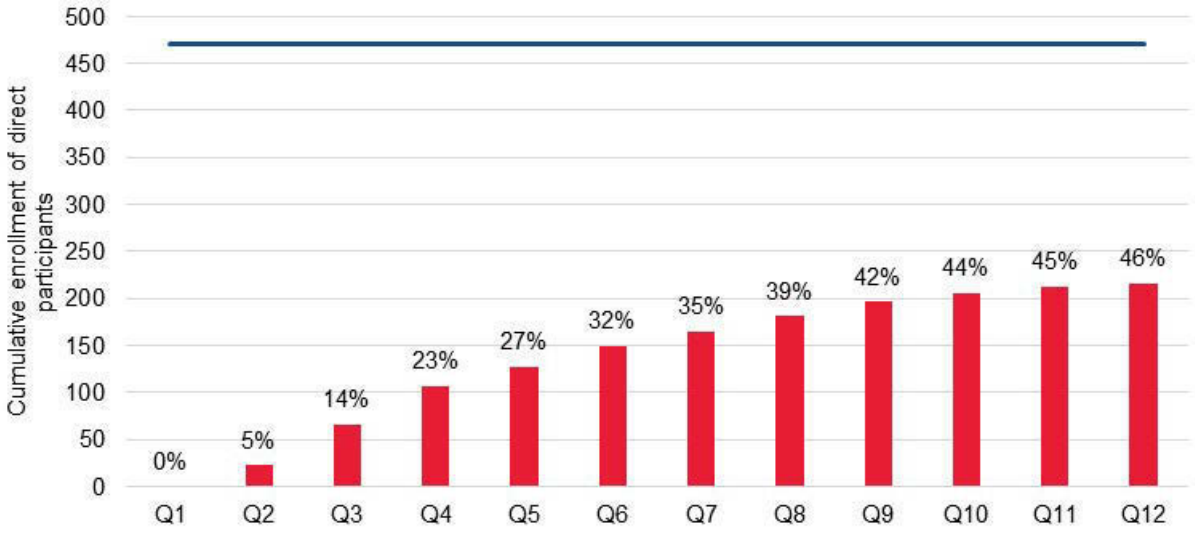
Figure XI.1. Phases of program innovation: Vinfen



3. Enrollment

Vinfen had hoped that approximately 85 percent of the 470 individuals served by all four teams, or roughly 400 persons, would engage with at least one program component. By the end of the 12th quarter, 216 individuals (about 46 percent of the 470 clients) had used at least one program component (Figure XI.2). We refer to these individuals as "direct participants." [77]

Figure XI.2. Percent of target enrollment achieved by quarter, Q1–Q12



Source: Mathematica analysis of program enrollment data that Vinfen submitted to the website maintained by CMMI's technical assistance contractor (the Lewin Group).

Note: The blue horizontal line represents Vinfen's target enrollment of 470 unique participants.

[77] The number of direct participants includes individuals who received at least one HCIA-funded service, regardless of whether they consented to participate in the evaluation.

4. Participants' demographic characteristics

During the study period, Vinfen sought to obtain its clients' consent to release information for this evaluation, and eventually obtained it from 267 individuals served by any of the four clinical sites. However, not all of the 267 individuals used an intervention component. Moreover, not all of the 216 who used at least one program component (that is, the direct participants referenced in Figure XI.2) provided consent. The information we received from Vinfen includes data on the 267 individuals who consented, even though some of them used no program services.

As shown in Table XI.1, more than half of the individuals who consented to have their information released were between ages 45 and 64 (53 percent), male (54 percent), and white (51 percent). Reliable and complete insurance data were not available. Vinfen staff believed that the large majority of participants were Medicaid eligible.

Table XI.1. Demographic characteristics of Vinfen participants

	Number of participants	Percent of participants
Total	267	100
Age		
Under 18	0	0
18–34	59	22.1
35–44	47	17.6
45–54	79	29.6
55–64	63	23.6
65 or older	19	7.1
Gender		
Female	124	46.4
Male	143	53.6
Race/ethnicity		
White	136	50.9
Black	73	27.3
Hispanic or Asian	43	16.1
Other or unknown race	15	5.6

Source: Mathematica analysis of survey and assessment data provided by Vinfen.

As shown in Table XI.2, 44 percent of the individuals who consented to have their information released did not use any of the four intervention components. The most commonly used intervention components were the HOW and the nurse practitioner, used by 48 percent and 45 percent of participants, respectively. A smaller percentage of participants used the IIMR curriculum (26 percent) and Health Buddy (21 percent).

The average age of participants who did not use any intervention component was younger than those who used the nurse practitioner, Health Buddy, the HOW, or the IIMR curriculum.

Table XI.2. Vinfen participant characteristics, by intervention component

	Nurse practitioner		Health Buddy		HOW		IIMR[c] curriculum		No intervention use	
Number of participants who used the intervention at least once during the study period[a]	119		55		127		69		118	
Percentage of all participants who used the intervention at least once during the study period	45%		21%		48%		26%		44%	
Average age among participants who used the intervention at least once during the study period	51		51		51		53		43	
Diagnosis[b]										
Psychological condition										
Anxiety	21	18%	17	31%	19	15%	13	19%	NR	NR
Bipolar	22	19%	14	25%	20	16%	15	22%	NR	NR
Major depression	14	12%	NR	NR	15	12%	13	19%	NR	NR
Post-traumatic stress disorder	11	9%	NR	NR	NR	NR	NR	NR	NR	NR
Schizophrenia or schizoaffective disorder	51	43%	16	29%	44	35%	16	23%	NR	NR
Other psychological diagnosis	22	19%	15	27%	21	17%	14	20%	NR	NR
Medical condition										
Diabetes	47	40%	25	45%	47	37%	29	42%	NR	NR
Hyperlipidemia	53	45%	33	60%	51	40%	37	54%	NR	NR
Hypertension	52	44%	31	56%	49	39%	33	48%	NR	NR
Obesity	48	40%	26	47%	48	38%	35	51%	NR	NR
Unknown or missing data	18	15%	NR	NR	35	28%	11	16%	114	97%
Dosage (median number of interactions per month during months in which the intervention was used)	2		18		3		2		-	

[a] Intervention components are not mutually exclusive; all four components were available to all participants.

[b] The diagnosis categories are not mutually exclusive; one individual may have more than one diagnosis simultaneously. The percentage of participants with each diagnosis who participated in a given intervention is calculated as the percentage of all participants who participated in that intervention. The percentage of participants with no diagnosis in a given intervention is also calculated as the percentage of all participants who participated in that intervention.

[c] IIMR = Integrated Illness Management and Recovery

NR = not reported; to mask individual identities, CMS policy prevents the reporting of cells with small sample sizes.

B. Methods

1. Quantitative methods

We were unable to conduct quantitative analysis of the effects of Vinfen's program because we did not receive sufficient data. In July 2015, Vinfen gave Mathematica assessment data for the 267 participants who consented to have their data used for evaluation purposes. These data included the demographic and service use information presented in Table XI.1, as well as various health and psychosocial assessment data, but no data on expenditures, hospitalizations,

readmissions, or ED visits. Given the small sample size associated with each intervention component, incompleteness of the data received, lack of a comparison group, and the self-reported nature of many of the assessments, we decided not to conduct any analysis.[78]

2. Qualitative methods

We collected qualitative data during site visits in 2014 and 2015, as well as a spring 2016 telephone interview during which we focused on program sustainability. Both site visits included in-depth interviews with awardee leaders, members of the workforce, and other stakeholders. Interview topics included implementation effectiveness, the target population, perceived program effectiveness, spillover effects, program costs and expenditures, workforce development and experience, project leadership, partnerships, and policy implications of the awardee's work. The second site visit was similar to the first with two major differences: (1) we added interview questions about program sustainability and (2) we conducted focus groups with outreach team members, participants who received program services, and comparison group members. Focus group discussion topics included the experiences and satisfaction of the focus group participants with the services they received, as well as perceived effectiveness of the programs.

C. Summative findings

As described above, we cannot report on the four core measures. However, site visit workforce respondents and focus group members provided the following perspectives on the program's effects on health outcomes and quality of life among participants:

- **Improvements in clinical indicators**. Anecdotally, some participants experienced improvements in such clinical indicators as blood pressure and hemoglobin A1c values. For example, one participant described the program's effects on his diabetes: "I was prediabetic. I'm no longer prediabetic. My cholesterol's gone down. My blood pressure's stable right now. My heart is stable. I've lost 18 pounds … So things are starting to … get in order." Despite these kinds of reports from some participants, workforce respondents stressed that many participants entered the program with chronic health problems, and they did not expect to see immediate or significant changes in health indicators over the relatively short course of the program period.
- **Positive changes in health behaviors.** Workforce members and individuals served by the program perceived that it empowered participants to make positive changes in their health behaviors. Anecdotally, participants had fewer missed appointments than before they entered the program, and they were better able to communicate with staff about their health. As one participant focus group member explained, "I think that self-improvement and self-management are two big benefits of the program."
- **Increased awareness of health issues**. Outreach team members described how the Health Buddy system sparked conversations about participant health and functioned as a helpful

[78] Vinfen obtained Medicaid claims data from MassHealth. However, these data were made available to Mathematica only in the aggregate and only for intervention participants who consented to have their data used for research purposes. Because we did not have access to individual-level data, we were unable to conduct trend or pre-post analyses. Through a contract with JEN Associates, Vinfen conducted various descriptive and trend analyses using these claims data; Vinfen sent these findings in a set of PowerPoint slides to us and to CMMI in early 2016.

tool for flagging health issues: "There's been many times where it's done its purpose, gave us warning signs if someone was symptomatic or needed immediate program." The system also helped train participants to use available services more appropriately, and to better recognize a true emergency, as opposed to an ongoing health issue. Workforce members also indicated that the availability of home-based care helped staff diagnose and appropriately treat long-existing health problems, particularly among participants who were afraid to access health services or struggled to make or keep appointments.

- **Improved access to and use of health care services**. Staff reported that the program helped some participants obtain care they would not have received otherwise, such as dental services, wellness visits, or such routine procedures as colonoscopies. An outreach team member provided an example of how the program had improved service use among participants: "[One of my clients] has put off going to the dentist for years because of his anxiety and we're scheduling his first appointment today."
- **Social engagement and improved quality of life.** In addition, the program helped participants access and use non-health-related services in their community: "It's great to see participants out and about and visiting and traveling and going to the clubhouse and doing activities and not always just thinking about their medical and mental issues." One participant focus group member explained, "I'm making friends. I'm a little more sociable. There's been a lot of improvement, but I still have a ways to go. So it's been a big help."

D. Findings about the workforce

Several critical workforce lessons emerged from Vinfen's HCIA experience. First, successful implementation of the program's components required adequate training on their use in real-world settings with a high-needs population. Second, the level and quality of supervision and support offered to workforce members affected job satisfaction and possibly job retention. Third, ongoing and frequent communication between the HOW and nurse practitioner dyads was essential to the success of the integrated care model. Finally, working across the multiple provider organizations involved in the program presented unique challenges for outreach team members. Below, we discuss each of the lessons in greater detail.

1. Workforce training

In general, members of the workforce who we interviewed agreed that the training they received helped them perform their jobs. However, some staff identified a need for more directive training on how the program components should be used in a real-world setting. As one HOW noted, training on use of the IIMR curriculum did not explicitly explain how the curriculum should be tailored for each client: "What's the end goal [of the IIMR curriculum], what are the exact expectations of … who do we need to do this with how many times a week and how long it should take to finish?"

Respondents also emphasized the importance of giving the staff clear guidance on the expectations associated with new workforce roles. For example, some nurses we interviewed suggested that more explicit guidance on the responsibilities of their position within the teams, which were created specifically for the program, might have been useful. Most nurse practitioners and RNs felt they received inadequate supervision from CCA, and consequently relied on ad hoc support from their outreach team leaders, the HOWs, and the Vinfen leadership

team. HOWs were generally more satisfied with the level of support and feedback they received from their provider organization and from Vinfen, noting that the regular calls with Dartmouth on the IIMR were also useful.

2. Workforce communication and collaboration

Members of the workforce cited frequent communication between nurses and HOWs and clarity on team roles and responsibilities as vital components of an effective outreach team. One team leader emphasized that successful partnership between each team's HOW and nurse practitioner depended greatly on the level of communication between team members: "I think that [the nurse practitioner's] ability to kind of remain in direct communication with the outreach worker to say 'I need [this]' or 'I'm going over today, can you meet me here and let's talk to this person?'… makes it work well. And I think at times, if somebody new was coming on [to the project], they could always take a little bit of time to get used to their style and how they … approached people and take a little time to get into the groove." When successful, collaboration between the nurse practitioners (or RNs) and the HOWs was one of the most effective components of the program. As one respondent explained, *"*Just having that extra support was fantastic … Clients loved it, staff loved it; it was something that really worked out well."

3. Working across provider organizations

Working across behavioral health providers presented additional challenges for implementation. As mentioned, nurse practitioners struggled with the level of support they received from their employer organization, which was not directly involved in program implementation. Because HOWs were employed by one organization and nurses by another, outreach team members did not share a standard EHR system, and that made it difficult to share participant information. Ensuring that the project model was implemented by all partners as originally intended was also a challenge. As one respondent explained: "It's always easier when you have one organization [implementing a project] so that you don't have to be constantly adapting pieces … Everybody was slightly adapting different things in different ways, and so it was difficult to have fidelity to a model amongst four organizations."

4. Workforce satisfaction and retention

Some members of the workforce explained that their jobs were stressful due to the high-level of needs of the population they served. In particular, most nurse practitioners, who did not feel adequately supported by their employer, described how the level of effort required to provide adequate care to participants often affected their ability to balance work and free time. The staff members we interviewed hypothesized that these high levels of stress may have affected the rates of staff retention for the program's nurse practitioners. In contrast, most HOWs who left the project were not doing so because of high stress levels but to pursue another job within their organization or to return to school. Despite the challenges, workforce respondents were generally satisfied with their jobs, and they agreed that watching participants improve their health and well-being was the most satisfying aspect of their work. As one HOW said: "I think overall, everybody that I've interacted with has either had the outcome they expected and wanted or has had an improved outcome, and that feels super good. And there have been times when people have called me and they've either gone to the emergency room when they wouldn't have and that was the right choice, or they haven't gone to the emergency room, which was absolutely the right choice. And that feels really good."

E. Program sustainability and spread

During our May 2016 follow-up call with Vinfen leaders, we learned more about their plans for sustainability and spread of each program component, as detailed below.

- **HOWs.** As of May 2016, the HOW component of Vinfen's HCIA program was being sustained in three ways. First, both Vinfen and Bay Cove incorporated HOWs into their behavioral health home outreach teams that serve the state's One Care program for dually eligible individuals. Second, Vinfen and Bay Cove were collaborating with Dartmouth College on a Research Project Grant (R01) from the National Institutes of Health to test impact outcomes linked to the IIMR and telehealth intervention. Clients enrolled in this research project have access to HOW services. Finally, a private grant through Blue Cross Blue Shield (BCBS) was funding a HOW on two Vinfen CBFS teams in Plymouth (an area that was not served by the HCIA). The BCBS grant provides three years of funding and will end in December 2018, after which Vinfen will pursue other funding opportunities. Longer-term sustainability of the HOW model will rely on statewide changes to the structure of the CBFS program. Vinfen hopes to obtain approval to integrate a HOW-like position into its CBFS teams through the re-procurement process with MassHealth.
- **Integrated care**. Vinfen and its partners could not sustain broader use of a nurse practitioner or RN as a member of the psychiatric rehabilitation teams. However, Vinfen does employ an RN and HOW as members of the Community Based Flexible Support team serving members of CCA's One Care Plan. The RN and HOW address a range of behavioral and medical health needs through the integrated model.
- **IIMR curriculum.** The IIMR curriculum was offered only to clients enrolled in the R01 project. In contrast to the HCIA program, which used a more structured model to deliver the curriculum, the R01 program uses an "a la carte" version of the curriculum, in which HOWs tailor the materials based on each client's needs. Vinfen considers the IIMR curriculum to be valuable, and plans to work with the Dartmouth team to re-design the content to be administered with less staff expense. In addition, the team hopes to translate the curriculum into an app-supported delivery system.
- **Health Buddy**. Vinfen did not expect to sustain use of the Health Buddy system, largely because Bosch went out of business and Vinfen did not have funding to maintain the system and train staff to use it. Vinfen was testing a similar system through its R01 grant with Dartmouth and exploring other ways to integrate technology into its services. Staff at Vinfen noted that it is challenging to find telehealth technology that works for this population because the technology must be easy to use and incorporate a basic reading level.

During site visit and telephone conversations, staff at Vinfen identified several factors that may contribute to long-term sustainability of the program's components.

- **Build on existing state efforts**. When possible, Vinfen has sought opportunities to align its work with existing state efforts, such as the CCA's One Care plan for dually eligible individuals, to sustain aspects of the program.
- **Develop collaborations and partnerships**. Over the course of the program, Vinfen developed relationships with the other implementing organizations that have afforded

opportunities, such as the BCBS and R01 grants, to sustain components of the program. Although the involvement of multiple provider organizations in the HCIA program presented several challenges for implementation, this organizational feature may ultimately result in broader program reach. All partners learned from the project and may draw from this experience in their future work. In the words of one respondent, the program "influenced four different groups of people rather than just one organization … and because all four organizations have made choices and have learned from it … I think it's going to influence practice."

- **Seek commitment from senior leaders**. In the process of exploring the possibility of sustaining this work, Vinfen relied on buy in from its senior leaders. For example, respondents noted that their experience with the Health Buddy system has motivated Vinfen leadership to commit to testing and integrating similar telehealth technology into its service system. As one key respondent explained: "This grant has really gotten a lot of our senior leadership a little bit more comfortable opening up and expanding to new things, which is fabulous…. We were able to test things out so that now we can say, okay that works, so that's the portion we're going to pick—let's say when we re-procure CBFS [community-based flexible support services]."

F. Lessons learned

Although we were unable to conduct an impact analysis of the program using quantitative data, our qualitative findings suggest that staff believed the program had a positive effect on participants' self-activation in regard to health behaviors, understanding their own health needs, and appropriate use of health and community resources. Respondent anecdotes suggest that the program may have positively affected the health of its participants; however, a more thorough analysis of quantitative program data is needed to understand the program's effects in this area.

Several key lessons emerged from Vinfen's experience developing and supporting its workforce. First, it was critical to ensure that workforce members understood the expectations associated with new program components and their roles and responsibilities. Second, workforce members dealing with a high-needs patient population required ongoing support and supervision from outreach team leaders and their employer organizations. Finally, workforce respondents cited frequent and ongoing communication between members of the outreach team, particularly between the nurses and HOWs, as vital to implementing the integrated care component.

Vinfen continues to sustain several components of its program, including placing a HOW and nurse practitioner pair on each behavioral health home outreach team. This outcome resulted mostly from Vinfen's tenacity in establishing and maintaining collaborative partnerships with the other organizations involved in the program, as well as efforts to build on existing state initiatives, including CCA's One Care program. Vinfen staff cited the support of its senior leaders as another facilitator in its efforts to sustain components of the program.

This page has been left blank for double-sided copying.

APPENDIX A:

TECHNICAL METHODS

This page has been left blank for double-sided copying.

I. CENTER FOR HEALTH CARE SERVICES

A. Introduction

This chapter describes the data sources and methods used in our CHCS impact analysis. We first describe the data sources (Section B). Next, we describe how the intervention group and control group members were identified (Section C). Lastly, we specify how outcome and control variables were constructed (Section D).

B. Description of data sources

To support our analysis, CHCS provided Mathematica with three data files each with records for intervention and control group members:

Enumeration file. The enumeration file, received in November 2014, included the program enrollment date, Social Security Number (SSN), date of birth, gender, employment status, highest level of education completed, living situation, and Medicaid and Medicare program identifiers, if the individual was enrolled in these programs. CHCS recorded these measures as of the program enrollment date for all intervention and control group members.

Survey assessment file. The survey assessment file, received in December 2015, included assessment data collected at three points in time: baseline, 6-month follow-up, and 12-month follow-up. Table A-I.1 displays the number and percent of intervention and control group members who completed the assessments at each point in time. The data elements in this file included an assessment round indicator (e.g., baseline, 6-month follow-up, 12-month follow-up), and assessment scores for a variety of mental and physical health indicators.

Table A-I.1. Survey assessment response rate

	Intervention group members		Control group members	
	Number	Percent	Number	Percent
Baseline	261	100	259	100
6-month follow-up	215	82	209	81
12-month follow-up	205	79	189	73

Source: Mathematica analysis of survey assessment data provided by CHCS, December 2015.

EHR file. The EHR file, received in November 2015, included service use and physical health status information. The service use information was reported in service-level records with information on service type and quantity (in minutes). The health status data included up to three records per intervention or comparison group member with health outcomes measured at three time points: baseline, at 6-month follow-up, and at 12-month follow-up. Table A-I.2 shows the health status measure collection rate at each point in time for intervention and control group members.

Table A-I.2. Health status measure collection rate

	Intervention group members		Control group members	
	Number	Percent	Number	Percent
Baseline	220	84	215	83
6-month follow-up	203	78	203	76
12-month follow-up	187	72	187	69

Source: Mathematica analysis of survey assessment data provided by CHCS, December 2015.

Records for the same individual are linkable across these three files based on a unique person identifier.

C. Identification of intervention group and control group

CHCS randomly assigned individuals recruited from Prospects Courtyard, a safe outdoor sleeping area on the Haven for Hope campus, to the intervention or control group. The intervention group members received access to program components (on-site primary care clinic and peer support specialist services) as well as to preexisting on-site behavioral health services and staff assistance with linkages and referrals to existing social services. Control group members received standard care, which only included behavioral health services and staff assistance.

Prior to conducting our analyses we assessed whether there were any statistically significant differences between the intervention and control groups at baseline. The two groups were similarly distributed in age, insurance coverage, employment status, living situation, and education level (Table A-I.3). However, the intervention group was significantly less likely to be male. BMI and blood pressure measurements were also similar for intervention and control group members at baseline.

Table A-I.3. CHCS demographics and health status measures at baseline

	Intervention group members		Control group members		Chi-square statistic
	Number	Percent[a]	Number	Percent[a]	
Total	261	100	259	100	
Demographics					
Age group					
18–34	78	30	86	33	2.10
35–44	64	25	56	22	
45–54	87	33	78	30	
55 or older	32	12	39	15	
Gender					
Male	148	57	167	64	3.29*
Female	113	43	92	36	
Insurance coverage					
Medicaid, non-dual	40	15	37	14	3.12
Medicare or dual	29	11	18	7	
Other[b]	191	73	204	79	

	Intervention group members		Control group members		Chi-square statistic
	Number	Percent[a]	Number	Percent[a]	
Employment status					
Employed	18	7	18	7	0.00
Not employed	242	93	241	93	
Living situation					
Homeless	230	90	230	92	0.48
Not homeless or other[c]	26	10	21	8	
Education level					
Less than HS	27	10	23	9	0.52
Some HS	62	24	66	26	
HS Degree or GED	107	41	104	40	
More than HS/GED	63	24	65	25	
Health status measures					
BMI					
Underweight or normal weight	91	44	80	39	2.59
Overweight	51	25	65	32	
Obese	64	31	59	29	
Blood pressure					
Hypertensive	46	22	49	24	0.10
Not hypertensive	161	78	159	76	

Source: EHR data provided by CHCS, November 2014.

[a] Percentages are reported among those with non-missing values within each category. One intervention group member was missing information on insurance coverage. One intervention group member was missing information on employment status. Five intervention group members and eight control group members were missing information on living situation. Two intervention group members and one control group member were missing information on education level. 55 intervention group members and 55 control group members were missing information on BMI at baseline. 54 intervention group members and 51 control group members were missing information on blood pressure at baseline.

[b] "Other" refers to intervention group members having neither Medicaid nor Medicare. Due to the format of the data provided, we cannot assess whether these members are uninsured or have some other form of insurance; however, anecdotally, we assume many of these are indeed uninsured.

[c] "Other" refers to correctional facilities, group quarters, and "other" living situations not explicitly classified as homeless or not homeless.

*Significantly different from zero at the .10 level.

D. Specifications of measures

In this section, we describe the specifications for the outcome and control variables used in our analyses. As noted above, the analyses were based on enrollment, service use, survey assessment, and health status measure data CHCS collected for the intervention and control group at baseline and around two follow-up time points[79]—6 and 12 months post-enrollment. Our samples for each regression were restricted to intervention and control group members with

[79] When examining the dates associated with each follow-up assessment, these measurement points did not necessarily occur exactly 6 and 12 months after a participant or control group member's enrollment date, in large part due to difficulties reaching a person for follow-up. For "6 months," CHCS measured outcomes at a point generally between 5 and 7 months after enrollment. Similarly, for "12 months," CHCS measured outcomes generally between 11 and up to 13 or more months after enrollment.

the given outcomes measured at baseline and at least one of the two follow-up points. This excluded approximately 20 percent of each of the two groups, who could not be reached during the window of availability to complete either follow-up assessment.

Below we first describe the survey assessment (Section D.1) and two health status measures (Section D.2). Then, we describe the control variables used in our analyses (Section D.3).

1. Survey assessments

The four survey assessment scores included in our analysis measure psychological distress, motivational readiness to change, quality of life (which CHCS felt would serve as a proxy for self-management capacity), and hope—all of which CHCS expected to be affected by the care team and peer support intervention components. These measures were scores derived from the following instruments:

- **Brief Symptom Index 18 (BSI-18).** The BSI-18 is a self-report tool designed to measure psychological distress. Patients rate their distress level on 18 symptom-specific questions using a Likert-type scale ranging from 0 (not at all) to 4 (extremely). Total scores, calculated by summing the question ratings, range from 0 to 72, with higher scores indicating higher global distress. This global distress score has been used and validated among drug-using populations (Wang et al. 2010) and populations with severe mental illness (Pahwa et al. 2012).
- **University of Rhode Island Change Assessment (URICA).** The URICA Committed Action (CA) score indicates motivational readiness to change. Ranging from -4 to 4, higher scores represent higher levels of motivation to change. This score has been used and validated among a sample of drug- and alcohol-dependent adults (Field et al. 2009). Further, the validity and psychometric properties of the URICA have been examined among adults with substance abuse (Henderson et al. 2004), adults with co-occurring drug abuse and severe mental illness (Nidecker et al. 2008), and adults from a general population and those with panic disorder (Dozois et al. 2004). The URICA asks respondents to rate the extent to which they agree with eight statements about each of four stages of change (pre-contemplation, contemplation, action, and maintenance) on a Likert scale ranging from 1 (strongly disagree) to 5 (strongly agree). We averaged responses within each stage to create a stage-level score.[80] Consistent with standard URICA scoring procedures, we subtracted the contemplation stage score from the action score to calculate the committed action score.
- **Short Form 36 Health Survey, Version 1 (SF-36).** The SF-36 assesses quality of life within the following domains: physical functioning, role limitations due to physical health or emotional problems, general health perceptions, bodily pain, emotional well-being, social functioning, and energy or fatigue. We combined the ratings of selected items within six dimensions of the SF-36 using a value-weighting method[81] to calculate the SF-6D

[80] Within each stage, one of the eight responses were found to not load well to the stage-level score. Therefore, in the standard scoring procedure, only seven of the eight responses within each stage are averaged to create the stage-level score. Specifically, questions 4, 8, 20, and 31 are omitted (University of Maryland, Baltimore County 2016).

[81] This method involves using regression coefficients calculated by Brazier and Roberts (2004) to weight 11 items in six dimensions from the SF-36 measuring quality of life (physical functioning, role limitations, social functioning, pain, emotional well-being, and energy/fatigue), as well as a binary indicating whether one or more of these

composite score, ranging from 0 to 1, with scores closer to 1 indicating better quality of life. The SF-6D has been used and validated among general adult populations (Petrou and Hockley 2005). The validity and psychometric properties of the SF-36 assessment have been examined among adults with chronic conditions and with depression (McHorney et al. 1993), adults with alcohol dependence (Daeppen et al. 1998), and adults with traumatic brain injury (Findler et al. 2001), but the validity and psychometric properties of the SF-6D have not been assessed in behavioral health-specific populations.

- **Adult Hope Scale (AHS).** The AHS, which assesses hope among participants, has two subscales. The first subscale, pathways, assesses the ability to plan routes to achieve desired goals. The second subscale, agency, assesses the ability to initiate and sustain the use of those pathways. Participants rate 12 statements using a Likert-type scale ranging from 1 (definitely false) to 8 (definitely true). Using standard scoring procedures, we calculated the subscale scores by summing the separate sets of four items that each comprises. We then summed the two subscale scores to calculate the global hope score, which could range from 8 to 64, with higher scores representing greater feelings of hope. This hope score has been used and validated among general adult populations (Snyder et al. 1991, Babyak et al. 1993), as well as among psychiatric patients (Brouwer et al. 2008) and traumatic injury survivors (Creamer et al. 2009).

Because intervention and control group members could have missing scores for some items for a given assessment, we calculated subscale and composite scores using the following methods:

- For the BSI-18 global distress score, we averaged the available information from 18 items. This accounted for missing items occurring for 53 intervention group records (8 percent) and 47 control group records (8 percent).
- For the URICA committed action score, we averaged the available information from the seven items within each of the four domains. Missing items occurred for 28 intervention records (4 percent) and 32 control group records (5 percent).
- For the SF-6D score, because each dimension score used in the composite only includes between one and three items, we determined there was not enough information to be able to impute responses for individuals with missing item scores.
- For the AHS global hope score, we averaged the available information from the four items within each of the two subscales. Missing items occurred for 7 intervention records (1 percent) and 7 control group records (1 percent).

No individual had 30 percent or more of item scores missing for any single subscale or composite.

dimensions is at the "most severe" level to take account of any additional effect on health state. Specifically, the 11 items used from the following dimensions include: physical functioning (items 3, 4, and 12), role limitations due to physical health or emotional problems (items 15 and 18), bodily pain (items 21 and 22), emotional well-being (items 24 and 28), social functioning (item 32), and energy or fatigue (item 27).

2. Health assessments

We also conducted analyses using two health status outcome measures: weight loss (using BMI) and reduction in blood pressure (using both systolic and diastolic blood pressure). We selected these two measures because they were applicable to the CHCS intervention and population, were feasible to construct with the EHR data provided by CHCS, and had strong evidence for their clinical importance. To examine the impact of Project HEALTH on these health status measures, we limited the analytic sample to only those intervention and control group members with a suboptimal measure at baseline. Specifically, we examined weight loss among those who were overweight or obese, and reduction in blood pressure among those who were hypertensive at baseline. We then examined the changes in these measures during the first 12 months of the intervention based on the HEDIS convention of using the last measurement within 12 months. That is, we examined impacts from baseline to either their 6 or 12 month follow-up (whichever was later to allow for more time for improvement).

Weight loss. The National Institutes of Health (2016) defines an individual as being overweight if their BMI falls between 25 to 29.9 kg/m^2, and obese as having a BMI of at least 30 kg/m^2. Two studies of obesity among homeless adults found similar mean BMI rates of 27.3 and 28.4 kg/m^2 and prevalence of overweight and obesity of 65.7 and 57.3 percent, while finding only 1.6 and 7.6 percent were underweight (Koh et al. 2012, Tsai and Rosenheck 2012).

We defined weight loss as reducing BMI by at least 5 percent. This reduction in BMI has been found to be associated with beneficial effects for other comorbid physical conditions. For example, Franz et al. (2015) found that "a weight loss of greater than 5 percent appears necessary for beneficial effects on HbA1c [a measure of diabetes], lipids [a measure of cholesterol levels], and blood pressure." We examined the proportions of intervention and control group members who lost 5 percent of BMI, as well as the difference in the change in mean BMI between the two groups.

Reduction in blood pressure. For adults 18 to 59 years old and for adults 60 to 85 years of age with diabetes, HEDIS defines hypertension as a blood pressure with either a systolic or diastolic levels greater than 140/90 mmHg. For adults 60 to 85 years old without diabetes, HEDIS identifies a slightly higher threshold of 150/90 mmHg (NCQA 2015). CHCS data did not specify whether an intervention or control group member has diabetes; thus, we used the more conservative cutoff of 140/90 mmHg for all intervention and control group members.[82] A meta-analysis of studies on hypertension among homeless adults found a pooled prevalence of hypertension of 27.0 percent (Bernstein et al. 2015).

We examined the proportions of intervention and control group members who reduced their blood pressure by enough to no longer be considered hypertensive, as well as the difference in the change in mean systolic and diastolic blood pressure levels between the two groups.

[82] We felt this extra distinction would not have made a substantial difference in our analysis, since only 5 percent of our sample with available blood pressure measures were aged 60 or above.

3. Other measures

The control variables we included in regression models for analyses of CHCS data and their specifications are listed in Table A-I.4. We derived all variables from the program enrollment, service use, and survey data provided by CHCS.

Table A-I.4. Impact analysis model control variable specifications—CHCS

Variable name	Specification
Intervention period	Categorical variable indicating time period of assessment. Categories include: baseline (reference); six months post-enrollment; twelve months post-enrollment.
Intervention group indicator	Categorical variable indicating treatment status. Categories include: control group member (reference); intervention group member.
Interaction between intervention period and intervention group indicator	Interaction between intervention period and intervention group indicator variables.
Program enrollment date	Categorical variable for month and year that individual was randomized into the study. Enrollment dates span between February 2013 (reference) and September 2014.
Quantity of medical services used	Continuous variable for member's quantity (in hours) of services used that were categorized as "medical" based on CHCS's recommended classification scheme.
Quantity of nonmedical services used	Continuous variable for member's quantity (in hours) of services used that were categorized as "nonmedical" (that is, behavioral health, case management, peer support, and "other") based on CHCS's recommended classification scheme.
Age	Categorical variable indicating member's age group at baseline. Categories include: 18–34 (reference); 35–44; 45–54; 55 and older.
Gender	Categorical variable of member's gender. Categories include: female or transgender (reference); male.
Employment status	Categorical variable of member's employment status at baseline. Categories include: employed (reference); unemployed.
Insurance status	Categorical variable of member's insurance status at baseline. Categories include: neither Medicare nor Medicaid (reference); Medicare; Medicaid; both Medicare and Medicaid.
Education status	Categorical variable of member's education status at baseline. Categories include: less than high school (reference); some high school; high school or GED; more than high school.
Living situation	Categorical variable of member's living situation at baseline. Categories include: homeless (reference); not homeless; incarcerated or other living situation.

II. THE FELTON INSTITUTE

A. Introduction

Our analysis of Felton's PREP innovation (Chapter IV of the report) reports participant demographic and diagnostic characteristics, and a pre-post analysis of two core outcome measures—hospitalizations and ED visits. Here, we first describe the data source for these analyses and then provide information on the fields used and process for constructing each measure.

B. Description of data sources

The quantitative analyses used data provided by Felton from its internal client data tracking system, the Cloud-based Integrated Reporting and Charting Environment (CIRCE), which is a HIPAA-compliant, cloud-based electronic case management and outcome tracking system that is used by all five county PREP teams to track data on PREP participants. Felton provided an extract of these data to Mathematica including all clients enrolled in PREP for at least 6 months between July 1, 2012 and December 31, 2014. We considered six months a reasonable minimum dosage period before program impacts could be detected.

The PREP county teams collected and recorded demographic information in CIRCE. They also manually extracted data on hospitalizations and ED visit from individual county mental health department systems and entered this information into the CIRCE system. We are unable to confirm the level of quality associated with data extraction and input. In addition, Felton staff were unable to confirm the completeness of the hospitalization and ED data. For example, if an ED visit or hospitalization was in a private facility or a facility in another county, it may not be reflected in the data—although Felton staff expressed confidence that most of these services were represented.

We present findings for a combined four-county sample (Alameda, Monterey, San Joaquin, and San Mateo counties), and separately for San Francisco. The samples for the two HCIA-supported programs—Monterey (N = 24) and San Joaquin (N = 28) counties—were too small to report individually. Thus, we combined these client populations with those from Alameda and San Mateo counties. The San Francisco program targets an expanded population, and was thus excluded from the combined county sample. The San Francisco program population was large enough to report separately.

C. Specifications of measures

We use multiple types of measures in these analyses. CMMI requested that we calculate four standardized outcome measures for all awardees to the extent feasible. These measures are: total Medicare and/or Medicaid expenditures, inpatient hospitalizations, hospital readmissions, and ED visits. For Felton, we were only able to calculate psychiatric-specific versions of two of the four core measures: psychiatric-related inpatient hospitalizations and psychiatric-related ED visits. Our specifications for these measures are described in the first section below (Section C.1). Section C.2 describes demographic and diagnosis-related measures used in our analyses.

1. Core measures

Receipt of PREP services was expected to help stabilize a client's condition leading to a reduction in psychiatric hospitalizations and ED visits. We analyzed the number of hospitalizations and ED visits associated with PREP participants for 12 months prior to and post program enrollment. PREP staff used their access to the county mental health data systems to identify participant psychiatric-related hospitalization and ED visits that occurred in the county. The admission and discharge date for these episodes were entered into CIRCE, and available in the data we received. Felton's standard policy was for the program staff to extract and record 12 months of pre-enrollment data for all participants upon enrollment. Once enrolled, the policy was to extract and record hospitalizations and ED visits until the client was no longer receiving services (the maximum enrollment period was two years). As a result, there was initially no standardized amount of post-enrollment data. For the purpose of this analysis, the Felton team used their access to county mental health department data systems to update CIRCE with 12 months of post-enrollment hospitalizations for all participants, regardless of when or why they exited the program. As a result of Felton's efforts, the data we received included 12 months pre- and post-enrollment hospitalization and ED visit data for all participants.

We analyzed utilization in six-month periods prior to and following enrollment. Baseline and intervention periods were defined for each intervention participant relative to their enrollment date. The first intervention period was defined as the enrollment month and the five months following that month. The second intervention period was defined as the seventh through twelfth month after enrollment. The first baseline period started in the month prior to the enrollment month and moved backward. Hospitalizations were assigned to an analysis period based on the date of discharge. ED visits were assigned to an analysis period based on admission date. The measures reported in Chapter IV of the report, Table IV.2 were calculated by summing the counts of services assigned to each analysis period.

2. Other measures

Chapter IV of the report, Table IV.1 presents demographic information on the 187 PREP program participants in the four-county sample and the 93 participants in San Francisco. There are some important caveats and notes regarding the data presented in this table:

- **Age.** This measure reflects the age at enrollment reported in the CIRCE data extract.
- **Gender.** Gender was self-reported by participants. The number of participants who reported being transgender was less than 11. Therefore we combined transgender with the male category to comply with CMS policy not to report cells smaller than 11.
- **Race/ethnicity.** Race/ethnicity was self-reported by participants. We coded 13 clients—or about 7 percent of the four-county sample—as having "other or unknown" race/ethnicity; the bulk of these include beneficiaries who reported more than one race. The total number of participants in San Francisco who identified as "African American" was less than 11. We thus report African Americans and Asian/Pacific Islanders together to comply with CMS policy.
- **Diagnosis.** Diagnosis information was determined by PREP county staff using the Structured Clinical Interview for Diagnostic and Statistical Manual, fifth addition. We reported participants diagnosed with schizophrenia, schizoaffective disorder, or schizophreniform disorder together due to group sizes less than 11.

III. FUND FOR PUBLIC HEALTH IN NEW YORK

A. Introduction

This chapter describes the methods for our analysis of Parachute NYC program impacts. We first describe the data sources (Section B) and methods for identifying the intervention group members (Section C). Then, we describe the methods for identifying the matched comparison group (Section D). Finally, we specify how we constructed outcome and control variables (Section E).

B. Description of data sources

We used two data sources for the impacts analysis: program enrollment data provided by FPHNY and Medicaid administrative data provided by the New York State Department of Health (NYSDOH).

Program enrollment data. FPHNY provided data files containing Medicaid identifiers, demographic characteristics (date of birth, gender, insurance coverage), and service begin and end dates for program participants who first used services from January 2013 through June 2015.

Medicaid administrative data. We obtained claims and enrollment data for January 2009 through June 2015 from the NYSDOH Medicaid Data Warehouse. These data included intervention group members enrolled in Medicaid for whom FPHNY provided a valid Medicaid identifier. In addition, the data included a pool of potential comparison group members who for at least one month during this period lived in New York City and had at least one claim with a behavioral health diagnosis. The claims data provided information on FFS and managed care payment amounts, service utilization, procedures, and diagnoses. The enrollment data provided monthly demographic and Medicaid enrollment information.

C. Identification of intervention group

Mathematica identified intervention group members in the NYSDOH Medicaid data based on the Medicaid identifiers provided by FPHNY; however, these identifiers were missing for about one-third of intervention group members who were identified as Medicaid enrolled. We received Parachute NYC program administrative data for 767 intervention group members who were identified as having Medicaid-only coverage, and 19 intervention group members who were identified as dually enrolled in Medicare and Medicaid. Only 507 of those intervention group members had valid Medicaid identifiers.

Starting from the 507 individuals with valid Medicaid identifiers, we narrowed the analysis population to those for whom the Medicaid administrative files would likely provide a comprehensive view of service utilization, using the following restrictions:

- **Full-benefit Medicaid coverage.** To ensure a consistent set of benefits were represented in the Medicaid administrative claims for the analysis population, we required full benefit Medicaid enrollment and no third party coverage. Individuals who were dually eligible for Medicare and Medicaid were excluded based on this restriction. This restriction excluded 193 individuals.

- **Observable in pre- and post- period.** To ensure observable data on the enrollee in the pre- and post-intervention analysis periods, we required Medicaid enrollment for at least six months prior and at least five months following the intervention enrollment month. We define the enrollment month as the first month during which the intervention group member used CRC or NA-MCT services. This restriction excluded 55 individuals.
- **Behavioral health diagnosis.** Because we believe matching to comparison individuals with the same behavioral health conditions (Table A-III.1) is important to assure comparability between the intervention and comparison group we also excluded intervention group members for whom a behavioral health diagnosis was not identified in the claims data. This restriction excluded fewer than 11 individuals.
- **Service use within three months of intervention enrollment.** Since enrollment into the FPHNY program was hypothesized to be predicated by a behavioral health-related service use, we excluded intervention group members who did not use a hospitalization, ED visit, psychiatric service, or office visit within the three months prior to their enrollment month. This restriction excluded fewer than 11 individuals.

Application of these restrictions resulted in 247 intervention group members eligible for analysis.

Table A-III.1. ICD-9 behavioral health diagnosis codes

Diagnosis codes	Label
295.00 to 295.95	Schizophrenia spectrum disorders
296.00 to 296.06, 296.40 to 296.80, 296.89, 296.10 to 296.16, 296.81	Bipolar disorders
296.20 to 296.36, 296.82, 300.4, 311, 311.0 296.90, 296.99, 293.83, 300.9	Depressive disorders Other mood disorders
305.1, 291.0-292.9, 303.0-303.9, 305.0-305.3, 292.0-292.9, 304.0-304.9, 305.2-305.9	Substance use Disorders (alcohol, tobacco, and other drug use)
300.00 to 300.11, 300.20 to 300.3, 309.81	Anxiety disorders
290.0-290.9, 294.1x	Dementia
297.0 to 298.9, V62.84, V62.85, E950, E951, E952, E953, E954, E955, E956, E957, E958, E959, 300.12 to 300.15, 300.6, 300.7 to 300.89, 301.0 to 301.9, 307.40 to 307.49, 312.0 to 312.23, 312.4 to 312.89, 313.81, 312.30 to 312.39, 302.0 to 302.9, 299.00 to 299.91, 307.1, 307.5, 307.51, 314.00 to 314.01, 307.20 to 307.3, 313.0 to 313.3, 313.82 to 316, 648.4, V65.2, V71.09, 780.09, V15.41, V15.42, V15,81, V17.0, V60.0, V62.29, V62.4, V62.81, V62.89 All other codes in the range of 290.0-299.91 and 300.00-316 (not specified above)	Other BH conditions not specified above (other psychotic disorders, suicidal or homicidal ideation, injury from suicide, dissociative disorders, somatoform disorders, personality disorders, sleep disorders, disruptive behavior disorders, impulse control behavior, sexual and gender identity disorders, ASD, eating disorders, ADHD, other disorder diagnoses in childhood, mental disorders in pregnancy, person feigning illness, observation for other suspected mental condition, other alteration of consciousness, social/contextual circumstances [violence])

D. Identification of comparison group

Propensity score matching and related matching methods are designed to create a comparison group that is similar in observable characteristics to the intervention group (Rosenbaum and Rubin 1983; Dehejia and Wahba 2002). In this section we describe how we developed a comparison pool and then applied matching methods to select the final comparison group. We also provide diagnostics to assess balance between the matched groups.

Step 1: Define comparison pool. We identified Medicaid enrollees in New York City who had at least one behavioral health-related diagnosis (Table A-III.1) between November 2012 and December 2014. We excluded individuals identified as intervention group members.[83] This resulted in a potential comparison pool of over 258,000 members.

For each potential comparison pool member, we created a pseudo-enrollment month in November 2012 through December 2014 that reflected the month when the member likely would have enrolled in the intervention if they had been an intervention group member. The pseudo-enrollment month allowed us to define the pre- and post-intervention periods for the analysis. For each person in the potential comparison pool, we identified all the months in which they had a claim, including a behavioral health diagnosis for one of four service types we hypothesized to lead to program participation: inpatient, ED visit, psychiatric visit, or office visit (Table A-III.2). Then, we randomly selected one of these months as their pseudo-enrollment month weighting the probability of selecting a given month for each potential comparison group member to assure that the distribution of enrollment or pseudo-enrollment months across the enrollment period would be similar for the intervention group and comparison population. To do this, we assigned each month in which a comparison person has an eligible service a weight equal to the proportion of intervention group members who enrolled in that month relative to the proportion of the comparison pool with an eligible service in that month. This gave greater weight to months in which there were relatively fewer comparison persons eligible relative to the proportion of participants who enrolled in that month.

Table A-III.2. Service use leading to program enrollment

Number	Variable name	Specification	Assign to month based on
1	Inpatient stay[a]	MMCOR_CD = 01, 04; SURS_SUBSYSTEM_COS_CD = 11; CLM_TYPE_CD = "I"	ADMIT_DT
2	ED visit	MMCOR_CD = 21; PROC_CD = 99281-99285; REV_CD_SUB = 0450, 0451, 0452, 0456, 0459, 0981	First SRV_DT on the claim
3	Psychiatric service	MMCOR_CD = 36, 37, 50, 55	SRV_DT
4	Office visit	MMCOR_CD = 24, 25, 48, 49	SRV_DT

[a] We considered ED visits that led to an inpatient stay part of the inpatient stay.

Once a pseudo-enrollment month was selected for each potential comparison pool member we excluded potential comparison individuals who did not have a sufficient Medicaid data to support analysis. In parallel to the exclusion for the participants, we required full benefit Medicaid enrollment with Medicaid as the primary payer for at least six months prior and at least five months following the pseudo-enrollment month, and excluded individuals who were dually enrolled in Medicaid and Medicare in this period. We also excluded potential comparison group members who did not fall within one of the strata defined by our exact match variables for the

[83] We assume that everyone remaining in this pool were not receiving FPHNY services; however, we cannot ensure that this is the case, since we only received valid Medicaid IDs for 507 out of 767 intervention group members that were identified as being enrolled in Medicaid.

intervention group (described in the next step). After these exclusions, our final comparison pool included 35,976 Medicaid enrollees.

Step 2: Match intervention group members at the individual level. Next, we conducted individual level matching in two stages. In the first stage, the matching algorithm matched the intervention group members who first used CRCs to members of the comparison pool with an inpatient stay in their pseudo-enrollment month.[84] Then, the second stage used the remaining potential comparison pool members (excluding those beneficiaries who were matched to the CRC intervention subgroup) and they were matched to intervention group members who first used the NA-MCT services.

The matching process used metrics of individual-level characteristics identified based on pre-period data to select a subset of comparison pool members who were as similar as possible to the intervention group on observable characteristics. First, the matching algorithm exact matched intervention to potential comparison members by assigning both to strata based on behavioral health diagnosis,[85] type of service used prior to enrollment (based on Table A-III.2), and whether the individual was enrolled in Medicaid for a full 12 months prior to the enrollment or pseudo-enrollment month. Then, within these strata, we used propensity score matching,[86] to match up to 20 comparison pool beneficiaries to each intervention group member. We used a nearest neighbor matching approach to select the closest comparison beneficiaries for each intervention group member.

The propensity score model used the following characteristics identified as of the enrollment (or pseudo-enrollment) month: calendar month and year of enrollment, age group (less than 18; 18–34; 35–44; 45–64), gender, race/ethnicity (African American, non-Hispanic, Hispanic, other), disability status, each of the Chronic Disability Payment System (CDPS) diagnosis flags,[87] and service use in the last 12 months (hospitalizations, ED visits, psychiatric visits, office visits, and total Medicaid expenditures).[88]

---

[84] Intervention group members using CRC services at enrollment were matched to comparison pool individuals with an inpatient stay in their enrollment month because CRC services were provided to individuals who required out-of-home care substituting for hospitalization.

[85] We created flags to represent a categorical variable indicating what we hypothesized to be a member's "most severe" diagnosis, ranked in the following order: (1) schizophrenic disorders; (2) bipolar disorders; (3) drug or alcohol-induced psychosis; (4) delusional disorder and other nonorganic psychosis; (5) suicide; (6) depressive, episodic mood disorder, or other depressive disorder; (7) anxiety, dissociative, somatoform, and adjustment reaction; and (8) other behavioral health-related diagnoses.

[86] A member's propensity score is the probability of belonging to the intervention group estimated based on a logistic regression model.

[87] The CDPS is a diagnosis-based risk adjustment model that was designed to allow Medicaid managed care organizations to adjust for health status capitation payments for TANF and disabled beneficiaries using administrative claims data. This is complemented by the MRx (Medicaid Rx) chronic disease classification, which uses NDC codes for prescription drugs. We used these complementary systems to create flags for chronic diseases and their expected level of expenditures.

[88] Instead of matching on continuous levels of use, we created categorical variables to represent the distribution of use. For hospitalizations, we created separate categories for: zero, one, two, and three or more hospitalizations. For

When this step was complete, the analysis population included 247 intervention group members and 2,917 matched comparison members.

Step 2: Assess the quality of the match. The following tests and procedures were used to verify that the intervention and matched comparison groups were similar, or balanced.

Before matching, we examined the ratio of potential comparison group members to intervention group members by exact matching stratum in order to understand how difficult it would be to match at least one potential comparison member to every intervention member. For example, if many strata had a low ratio of potential comparison members to intervention members prior to matching, we might consider reducing the number of variables used for exact matching. However, due to the very large size of the comparison pool, we observed sufficient ratios in each stratum to proceed with our strata.

We then graphically compared the propensity score distributions for the intervention and potential comparison group members prior to matching, looking for overlap in the propensity scores for the two groups (Figures A-III.1.a and A-III.1.b). Though in both figures, propensity scores appear to be generally larger for the intervention group (rightmost box plot), we determined that the amount of overlap indicated a sufficient pool of comparison group members available for matching with propensity scores similar to those observed in the intervention group.

Figure A-III.1.a. Propensity score distributions (FPHNY, CRC subgroup)

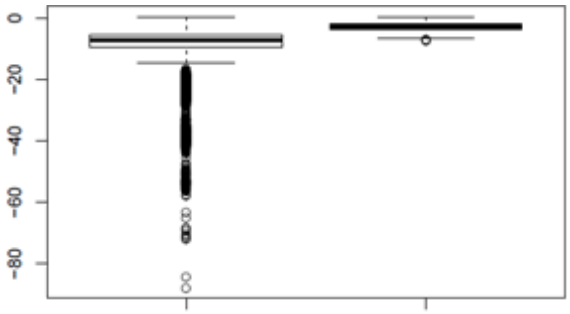
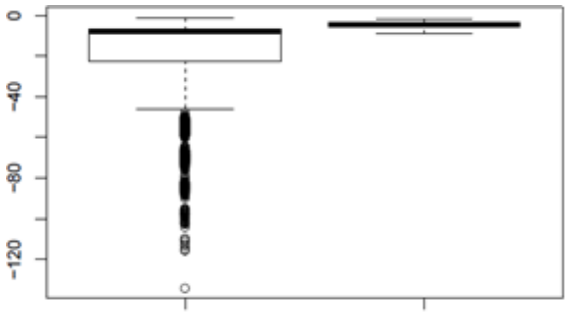


Figure A-III.1.b. Propensity score distributions (FPHNY, NA-MCT subgroup)



Note: Figures present boxplots created using the estimated propensity scores for the comparison and intervention groups (the left and right boxes, respectively). The line in the middle of each box represents the median score for the group. The lower and upper bounds of the box indicate the first and third quartile.

After we conducted matching, we examined the number of comparison members matched to each intervention group member (Table A-III.3). A large number of 1:1 matches could indicate that the matching was problematic. The match ratios in this case do not present any issue, and again demonstrate the very large size of the potential comparison pool.

ED visits, we created separate categories for: zero, one, two or three, and four or more ED visits. For each of psychiatric visits, office visits, and total Medicaid expenditures, we created categories ranking by the distribution of use (by percentile) among the intervention and potential comparison group members: zero use, those who used and were in the lowest 20 percent of use; between 20 and 40 percent of use, between 40 and 60 percent of use, between 60 and 80 percent of use, and between 80 and 100 percent of use.

Table A-III.3. Frequency table of ratio of intervention to comparison group members for each matched set (FPHNY)

Ratio of intervention to comparison group members	1:1	1:2	1:5	1:10	1:20	0:1
Number of matched sets (CRC)	0	0	102	35	26	12,510
Number of matched sets (NA-MCT)	1	3	2	4	74	33,059

Note: Each cell indicates the number of intervention group members matched to the number of comparison group members indicated for that column. The rightmost column shows the number of potential comparison group members that were not matched to an intervention group member.

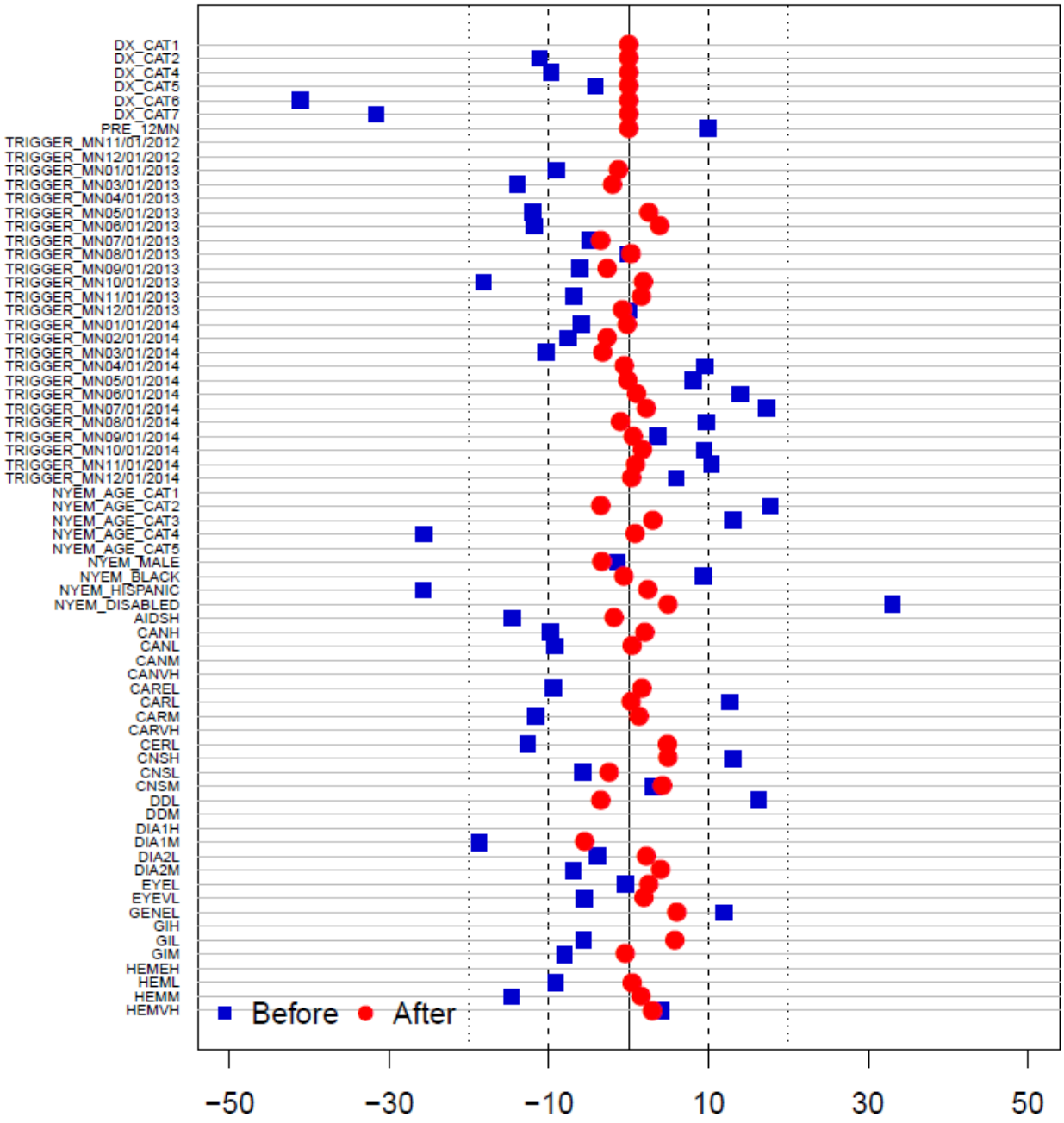
To further investigate balance between the intervention and matched comparison groups, we evaluated how matching affected the balance on all matching variables (Figures A-III.2.a and A-III.2.b; Tables A-III.4.a and A-III.4.b) by comparing the absolute and standardized difference between the intervention and comparison groups for each variable before and after matching. The standardized difference measures the difference in means relative to the pooled standard deviation of intervention and comparison group members for each variable. The standardized difference measure is advantageous in that it allows us to compare all variables on the same scale. We compared the standardized differences using plots with dashed lines at +/- 0.10 standardized differences to visually inspect whether we obtained good balance for each variable, and with a balance table that shows both absolute and standardized differences between intervention and comparison groups before and after matching. In each instance, we found that all variables are within the 0.10 standardized difference limit after matching, indicating good performance of the match.

Step 3: Create analysis weights. Weights were developed for each member of the analysis population. Weights for intervention group members were set to one. Weights for comparison group members were set to one divided by the number of comparison group members assigned to the member's associated intervention group member. For example, for an intervention group member matched to 20 comparison group members, the intervention group member would have a weight equal to one, and each comparison group member's weight would equal 0.05.

In some pre- or post-intervention analysis months,[89] intervention or comparison group members might not have had sufficient Medicaid data to be included in the analysis, because they were not enrolled in Medicaid with full-benefits or Medicaid was not their primary payer. The weight for these individuals was set to zero in analysis months where they did not meet these Medicaid coverage criteria.

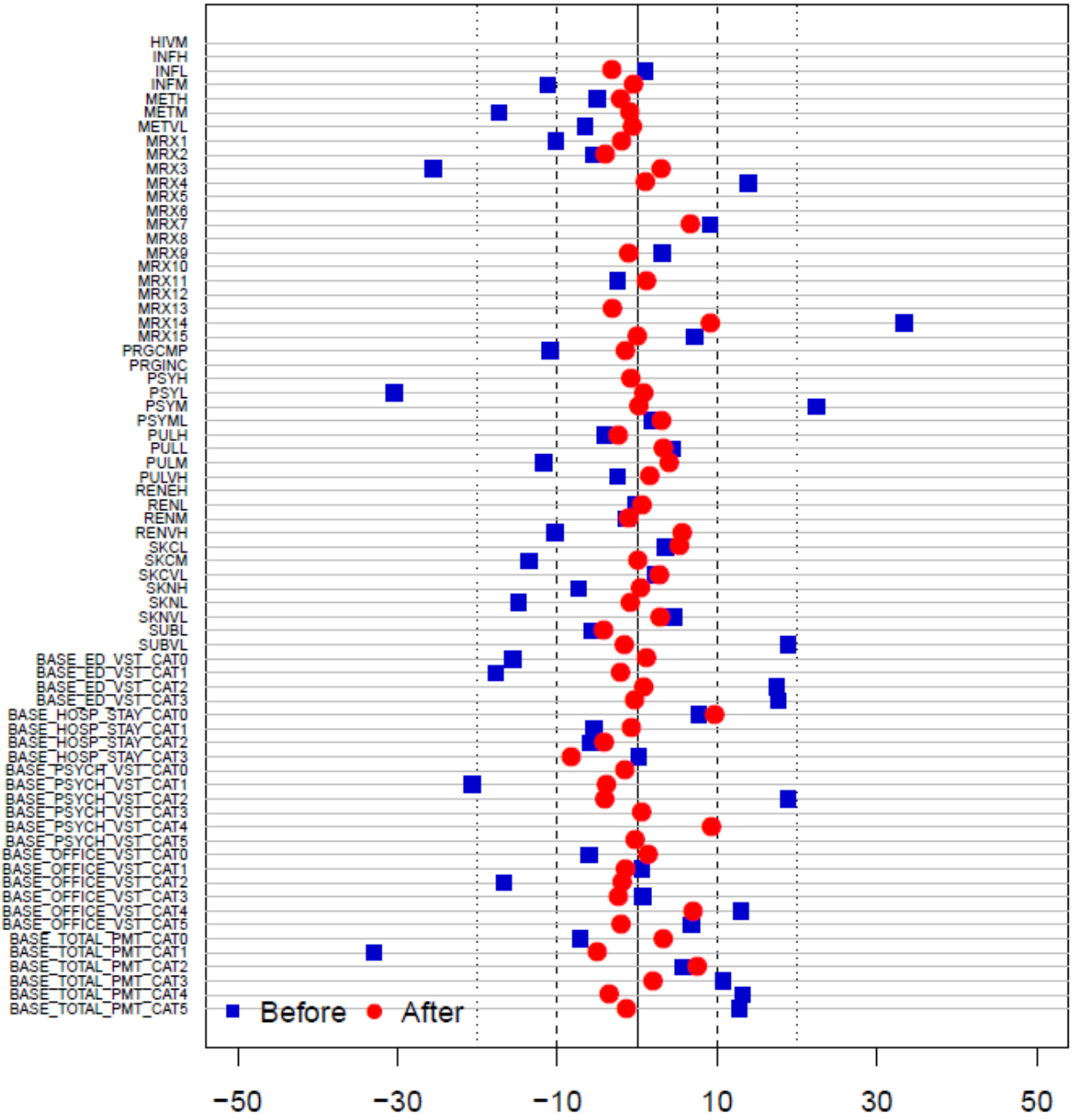
[89] We required that each intervention and comparison group member be enrolled in Medicaid with full benefits and with Medicaid as their primary payer for six months prior to the enrollment month, during the enrollment month and for five months following the program enrollment month. In contrast, the analysis period included three years prior to and one year after program enrollment.

Figure A-III.2.a. Balance plots comparing the standardized difference for each matching variable before and after matching (FPHNY, CRC subgroup)



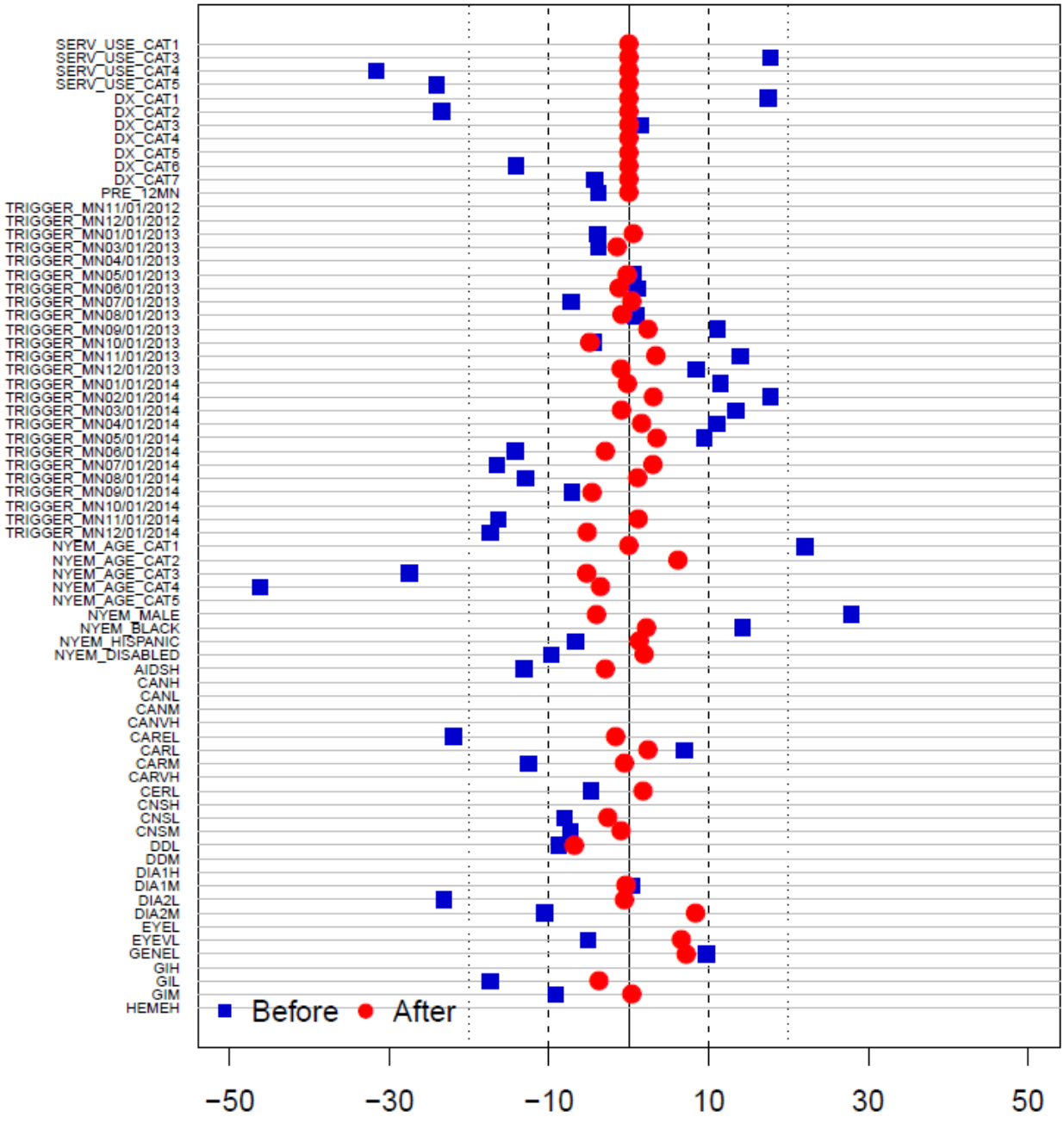
Note: Blue markers show the standardized difference before matching; red markers show the standardized difference after exact matching and propensity score modeling. See Table A-III.4.a for descriptions of the variables included in this figure.

Figure A-III.2.a *(continued)*



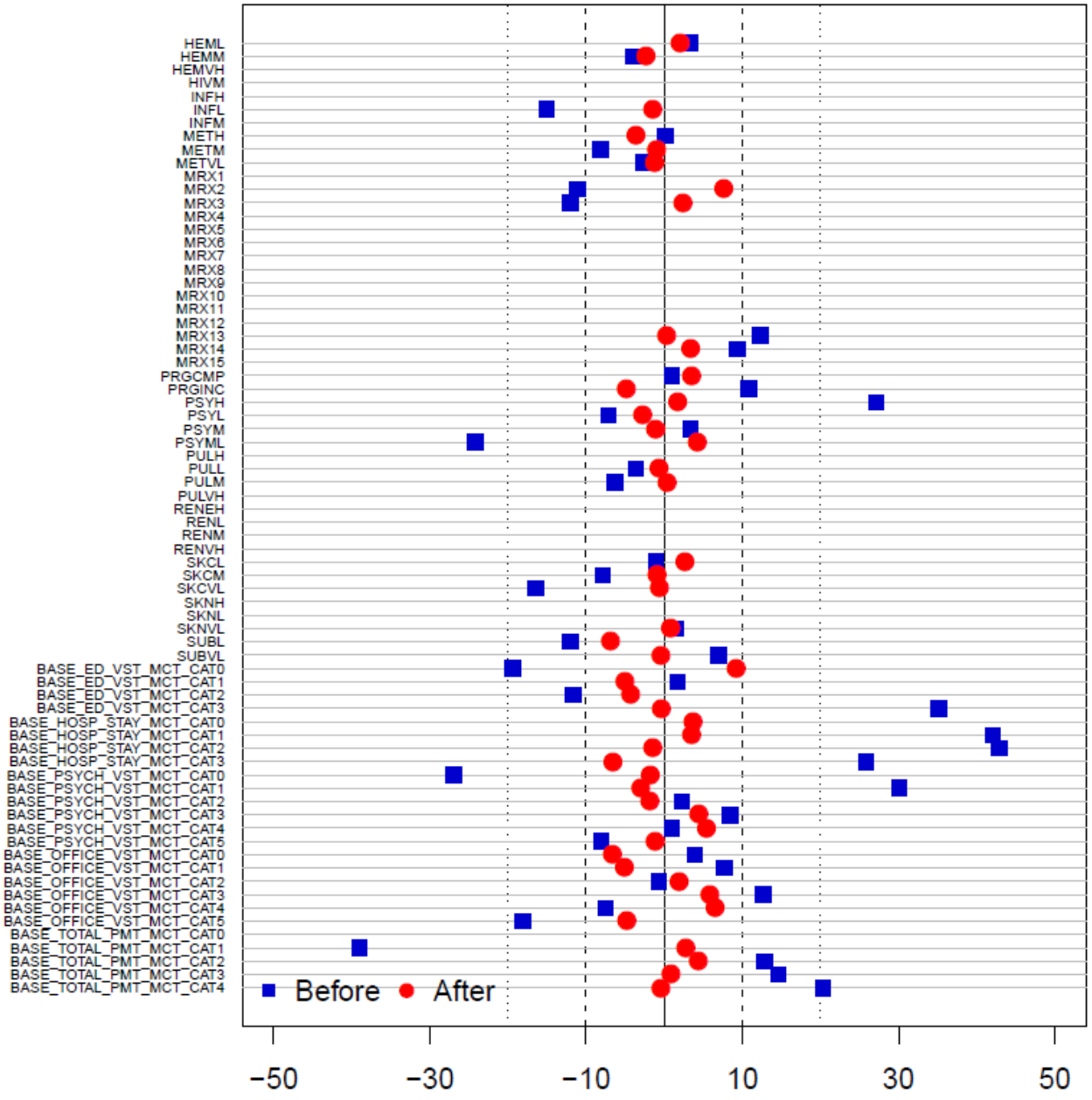
Note: Blue markers show the standardized difference before matching; red markers show the standardized difference after exact matching and propensity score modeling. See Table A-III.4.a for descriptions of the variables included in this figure.

Figure A-III.2.b. Balance plots comparing the standardized difference for each matching variable before and after matching (FPHNY, NA-MCT subgroup)



Note: Blue markers show the standardized difference before matching; red markers show the standardized difference after exact matching and propensity score modeling. See Table A-III.4.b for descriptions of the variables included in this figure.

Figure A-III.2.a *(continued)*



Note: Blue markers show the standardized difference before matching; red markers show the standardized difference after exact matching and propensity score modeling. See Table A-III.4.b for descriptions of the variables included in this figure.

Table A-III.4.a. Balance table before and after matching (FPHNY, CRC subgroup)

		Before matching					After matching				
Variable name	Variable description	Comparison	Intervention	adj.diff	std.diff	p	Comparison	Intervention	adj.diff	std.diff	p
					Exact match variables						
	Hierarchical diagnosis code prior to program enrollment										
DX_CAT1	Schizophrenia	0.2970	0.6503	0.3533	0.7726	0.0000	0.6503	0.6503	0.0000	0.0000	1.0000
DX_CAT2	Bipolar	0.2834	0.2331	-0.0503	-0.1117	0.1562	0.2331	0.2331	0.0000	0.0000	1.0000
DX_CAT4	Other nonorganic psychosis	0.0366	0.0184	-0.0182	-0.0971	0.2180	0.0184	0.0184	0.0000	0.0000	1.0000
DX_CAT5	Suicidal	0.0104	0.0061	-0.0042	-0.0419	0.5950	0.0061	0.0061	0.0000	0.0000	1.0000
DX_CAT6	Depression	0.2514	0.0736	-0.1778	-0.4113	0.0000	0.0736	0.0736	0.0000	0.0000	1.0000
DX_CAT7	Anxiety/adjustment	0.1212	0.0184	-0.1028	-0.3164	0.0001	0.0184	0.0184	0.0000	0.0000	1.0000
PRE_12MN	12 months of continuous baseline enrollment	0.8891	0.9202	0.0311	0.0992	0.2078	0.9202	0.9202	0.0000	0.0000	1.0000
					Propensity score variables						
	Program enrollment month										
TRIGGER_MN 11/01/2012	11/2012	0.0000	0.0000	0.0000	0.0000	1.0000	0.0000	0.0000	0.0000	0.0000	1.0000
TRIGGER_MN 12/01/2012	12/2012	0.0000	0.0000	0.0000	0.0000	1.0000	0.0000	0.0000	0.0000	0.0000	1.0000
TRIGGER_MN 01/01/2013	01/2013	0.0181	0.0061	-0.0120	-0.0903	0.2517	0.0074	0.0061	-0.0012	-0.0129	0.8643
TRIGGER_MN 03/01/2013	03/2013	0.0297	0.0061	-0.0236	-0.1396	0.0765	0.0080	0.0061	-0.0018	-0.0201	0.8065
TRIGGER_MN 04/01/2013	04/2013	0.0000	0.0000	0.0000	0.0000	1.0000	0.0000	0.0000	0.0000	0.0000	1.0000
TRIGGER_MN 05/01/2013	05/2013	0.0425	0.0184	-0.0241	-0.1197	0.1286	0.0150	0.0184	0.0034	0.0248	0.7518
TRIGGER_MN 06/01/2013	06/2013	0.0335	0.0123	-0.0213	-0.1186	0.1322	0.0080	0.0123	0.0043	0.0389	0.5733
TRIGGER_MN 07/01/2013	07/2013	0.0402	0.0307	-0.0096	-0.0488	0.5360	0.0371	0.0307	-0.0064	-0.0350	0.6782
TRIGGER_MN 08/01/2013	08/2013	0.0430	0.0429	0.0000	-0.0002	0.9982	0.0423	0.0429	0.0006	0.0031	0.9714
TRIGGER_MN 09/01/2013	09/2013	0.0431	0.0307	-0.0124	-0.0614	0.4359	0.0356	0.0307	-0.0049	-0.0272	0.7539
TRIGGER_MN 10/01/2013	10/2013	0.0428	0.0061	-0.0367	-0.1821	0.0208	0.0049	0.0061	0.0012	0.0182	0.8312
TRIGGER_MN 11/01/2013	11/2013	0.0376	0.0245	-0.0130	-0.0687	0.3832	0.0221	0.0245	0.0025	0.0156	0.8418
TRIGGER_MN 12/01/2013	12/2013	0.0431	0.0429	-0.0001	-0.0005	0.9946	0.0445	0.0429	-0.0015	-0.0078	0.9289
TRIGGER_MN 01/01/2014	01/2014	0.0497	0.0368	-0.0129	-0.0593	0.4516	0.0371	0.0368	-0.0003	-0.0017	0.9845
TRIGGER_MN 02/01/2014	02/2014	0.0393	0.0245	-0.0148	-0.0762	0.3337	0.0288	0.0245	-0.0043	-0.0273	0.7614

		Before matching					After matching				
Variable name	Variable description	Comparison	Intervention	adj.diff	std.diff	p	Comparison	Intervention	adj.diff	std.diff	p
TRIGGER_MN 03/01/2014	03/2014	0.0463	0.0245	-0.0218	-0.1038	0.1877	0.0298	0.0245	-0.0052	-0.0324	0.7114
TRIGGER_MN 04/01/2014	04/2014	0.0626	0.0859	0.0233	0.0958	0.2241	0.0874	0.0859	-0.0015	-0.0055	0.9479
TRIGGER_MN 05/01/2014	05/2014	0.0551	0.0736	0.0185	0.0811	0.3031	0.0739	0.0736	-0.0003	-0.0012	0.9888
TRIGGER_MN 06/01/2014	06/2014	0.0639	0.0982	0.0342	0.1395	0.0766	0.0954	0.0982	0.0028	0.0099	0.9110
TRIGGER_MN 07/01/2014	07/2014	0.0531	0.0920	0.0389	0.1727	0.0284	0.0859	0.0920	0.0061	0.0219	0.7923
TRIGGER_MN 08/01/2014	08/2014	0.0418	0.0613	0.0195	0.0973	0.2170	0.0638	0.0613	-0.0025	-0.0103	0.9038
TRIGGER_MN 09/01/2014	09/2014	0.0361	0.0429	0.0068	0.0364	0.6439	0.0417	0.0429	0.0012	0.0057	0.9413
TRIGGER_MN 10/01/2014	10/2014	0.0577	0.0798	0.0221	0.0946	0.2301	0.0752	0.0798	0.0046	0.0169	0.8384
TRIGGER_MN 11/01/2014	11/2014	0.0610	0.0859	0.0249	0.1039	0.1874	0.0834	0.0859	0.0025	0.0084	0.9148
TRIGGER_MN 12/01/2014	12/2014	0.0596	0.0736	0.0140	0.0591	0.4533	0.0727	0.0736	0.0009	0.0035	0.9666
	Age group										
NYEM_AGE_ CAT1	<18	0.0000	0.0000	0.0000	0.0000	1.0000	0.0000	0.0000	0.0000	0.0000	1.0000
NYEM_AGE_ CAT2	18-34	0.2483	0.3252	0.0768	0.1777	0.0241	0.3417	0.3252	-0.0166	-0.0348	0.6795
NYEM_AGE_ CAT3	34-45	0.1771	0.2270	0.0499	0.1305	0.0976	0.2144	0.2270	0.0126	0.0303	0.7209
NYEM_AGE_ CAT4	45-64	0.5746	0.4479	-0.1267	-0.2563	0.0011	0.4439	0.4479	0.0040	0.0080	0.9237
NYEM_AGE_ CAT5	65+	0.0000	0.0000	0.0000	0.0000	1.0000	0.0000	0.0000	0.0000	0.0000	1.0000
NYEM_MALE	Male	0.4857	0.4785	-0.0072	-0.0144	0.8546	0.4954	0.4785	-0.0169	-0.0338	0.6826
	Race/ethnicity										
NYEM_BLACK	African American, not Hispanic	0.3242	0.3681	0.0439	0.0938	0.2340	0.3712	0.3681	-0.0031	-0.0065	0.9398
NYEM_ HISPANIC	Hispanic	0.3430	0.2209	-0.1221	-0.2576	0.0011	0.2110	0.2209	0.0098	0.0238	0.7707
NYEM_ DISABLED	Disabled	0.4914	0.6564	0.1651	0.3304	0.0000	0.6319	0.6564	0.0245	0.0493	0.4897
	CDPS condition flags										
AIDSH	AIDS, High	0.1853	0.1288	-0.0565	-0.1456	0.0647	0.1350	0.1288	-0.0061	-0.0186	0.8285
CANH	Cancer, High	0.0197	0.0061	-0.0136	-0.0981	0.2130	0.0046	0.0061	0.0015	0.0201	0.7970
CANL	Cancer, Low	0.0185	0.0061	-0.0124	-0.0921	0.2423	0.0058	0.0061	0.0003	0.0040	0.9628
CANM	Cancer, Medium	0.0000	0.0000	0.0000	0.0000	1.0000	0.0000	0.0000	0.0000	0.0000	1.0000

		Before matching					After matching				
Variable name	Variable description	Comparison	Intervention	adj.diff	std.diff	p	Comparison	Intervention	adj.diff	std.diff	p
CANVH	Cancer, Very High	0.0000	0.0000	0.0000	0.0000	1.0000	0.0000	0.0000	0.0000	0.0000	1.0000
CAREL	Cardiovascular, Extra Low	0.1838	0.1472	-0.0366	-0.0945	0.2304	0.1417	0.1472	0.0055	0.0167	0.8494
CARL	Cardiovascular, Low	0.2345	0.2883	0.0539	0.1270	0.1070	0.2871	0.2883	0.0012	0.0028	0.9732
CARM	Cardiovascular, Medium	0.1243	0.0859	-0.0384	-0.1167	0.1386	0.0825	0.0859	0.0034	0.0126	0.8849
CARVH	Cardiovascular, Very High	0.0000	0.0000	0.0000	0.0000	1.0000	0.0000	0.0000	0.0000	0.0000	1.0000
CERL	Cerebrovascular, Low	0.0528	0.0245	-0.0283	-0.1269	0.1073	0.0181	0.0245	0.0064	0.0482	0.5857
CNSH	Nervous System, High	0.0040	0.0123	0.0083	0.1308	0.0970	0.0080	0.0123	0.0043	0.0489	0.5914
CNSL	Nervous System, Low	0.2009	0.1779	-0.0230	-0.0575	0.4656	0.1868	0.1779	-0.0089	-0.0243	0.7844
CNSM	Nervous System, Medium	0.0429	0.0491	0.0062	0.0304	0.6993	0.0408	0.0491	0.0083	0.0425	0.6295
DDL	Developmental Disability, Low	0.0200	0.0429	0.0229	0.1627	0.0389	0.0494	0.0429	-0.0064	-0.0350	0.7231
DDM	Developmental Disability, Medium	0.0000	0.0000	0.0000	0.0000	1.0000	0.0000	0.0000	0.0000	0.0000	1.0000
DIA1H	Diabetes, Type 1 High	0.0000	0.0000	0.0000	0.0000	1.0000	0.0000	0.0000	0.0000	0.0000	1.0000
DIA1M	Diabetes, Type 1 Low	0.0549	0.0123	-0.0426	-0.1879	0.0171	0.0196	0.0123	-0.0074	-0.0551	0.5201
DIA2L	Diabetes, Type 2 High	0.1616	0.1472	-0.0144	-0.0391	0.6197	0.1402	0.1472	0.0071	0.0221	0.8107
DIA2M	Diabetes, Type 2 Low	0.0594	0.0429	-0.0165	-0.0697	0.3763	0.0359	0.0429	0.0071	0.0398	0.6644
EYEL	Eye, Low	0.0127	0.0123	-0.0005	-0.0042	0.9573	0.0098	0.0123	0.0025	0.0250	0.7643
EYEVL	Eye, Very Low	0.0624	0.0491	-0.0133	-0.0552	0.4835	0.0451	0.0491	0.0040	0.0190	0.8201
GENEL	Genital, Extra Low	0.0993	0.1350	0.0357	0.1191	0.1305	0.1153	0.1350	0.0196	0.0603	0.4667
GIH	Gastrointestinal, High	0.0000	0.0000	0.0000	0.0000	1.0000	0.0000	0.0000	0.0000	0.0000	1.0000
GIL	Gastrointestinal, Low	0.2134	0.1902	-0.0232	-0.0567	0.4720	0.1690	0.1902	0.0212	0.0576	0.5013
GIM	Gastrointestinal, Medium	0.1044	0.0798	-0.0246	-0.0807	0.3059	0.0810	0.0798	-0.0012	-0.0046	0.9573
HEMEH	Hematological, Extra High	0.0000	0.0000	0.0000	0.0000	1.0000	0.0000	0.0000	0.0000	0.0000	1.0000
HEML	Hematological, Low	0.0431	0.0245	-0.0185	-0.0914	0.2459	0.0239	0.0245	0.0006	0.0041	0.9623
HEMM	Hematological, Medium	0.0415	0.0123	-0.0292	-0.1470	0.0620	0.0107	0.0123	0.0015	0.0151	0.8577
HEMVH	Hematological, Very High	0.0085	0.0123	0.0038	0.0410	0.6026	0.0092	0.0123	0.0031	0.0294	0.7164
HIVM	HIV, Medium	0.0000	0.0000	0.0000	0.0000	1.0000	0.0000	0.0000	0.0000	0.0000	1.0000
INFH	Infectious, High	0.0000	0.0000	0.0000	0.0000	1.0000	0.0000	0.0000	0.0000	0.0000	1.0000
INFL	Infectious, Low	0.0651	0.0675	0.0024	0.0097	0.9017	0.0755	0.0675	-0.0080	-0.0318	0.7188
INFM	Infectious, Medium	0.0228	0.0061	-0.0166	-0.1119	0.1555	0.0064	0.0061	-0.0003	-0.0043	0.9643
METH	Metabolic, High	0.0191	0.0123	-0.0068	-0.0499	0.5267	0.0144	0.0123	-0.0021	-0.0206	0.8290
METM	Metabolic, Medium	0.0767	0.0307	-0.0460	-0.1735	0.0277	0.0322	0.0307	-0.0015	-0.0094	0.9194
METVL	Metabolic, Very Low	0.0653	0.0491	-0.0162	-0.0657	0.4041	0.0503	0.0491	-0.0012	-0.0059	0.9472
PRGCMP	Pregnancy, Complete	0.0395	0.0184	-0.0210	-0.1084	0.1687	0.0206	0.0184	-0.0021	-0.0148	0.8555
PRGINC	Pregnancy, Incomplete	0.0000	0.0000	0.0000	0.0000	1.0000	0.0000	0.0000	0.0000	0.0000	1.0000
PSYH	Psychiatric, High	0.1736	0.5276	0.3540	0.9307	0.0000	0.5313	0.5276	-0.0037	-0.0079	0.8788
PSYL	Psychiatric, Low	0.1053	0.0123	-0.0930	-0.3045	0.0001	0.0110	0.0123	0.0012	0.0086	0.8790
PSYM	Psychiatric, Medium	0.1034	0.1718	0.0684	0.2239	0.0045	0.1709	0.1718	0.0009	0.0024	0.9733
PSYML	Psychiatric, Medium Low	0.1708	0.1779	0.0071	0.0188	0.8115	0.1650	0.1779	0.0129	0.0309	0.6157
PULH	Pulmonary, High	0.0102	0.0061	-0.0041	-0.0407	0.6051	0.0083	0.0061	-0.0021	-0.0235	0.7698
PULL	Pulmonary, Low	0.2626	0.2822	0.0196	0.0446	0.5710	0.2681	0.2822	0.0141	0.0331	0.7041
PULM	Pulmonary, Medium	0.1032	0.0675	-0.0357	-0.1175	0.1358	0.0583	0.0675	0.0092	0.0401	0.6424
PULVH	Pulmonary, Very High	0.0153	0.0123	-0.0031	-0.0250	0.7513	0.0107	0.0123	0.0015	0.0156	0.8646
RENEH	Renal, Extra High	0.0000	0.0000	0.0000	0.0000	1.0000	0.0000	0.0000	0.0000	0.0000	1.0000
RENL	Renal, Low	0.0680	0.0675	-0.0005	-0.0022	0.9779	0.0660	0.0675	0.0015	0.0066	0.9429

		Before matching					After matching				
Variable name	Variable description	Comparison	Intervention	adj.diff	std.diff	p	Comparison	Intervention	adj.diff	std.diff	p
RENM	Renal, Medium	0.0073	0.0061	-0.0011	-0.0134	0.8651	0.0071	0.0061	-0.0009	-0.0109	0.8924
RENVH	Renal, Very High	0.0616	0.0368	-0.0248	-0.1034	0.1893	0.0273	0.0368	0.0095	0.0565	0.5071
SKCL	Skeletal and Connective, Low	0.1116	0.1227	0.0111	0.0353	0.6545	0.1064	0.1227	0.0163	0.0530	0.5361
SKCM	Skeletal and Connective, Medium	0.1096	0.0675	-0.0422	-0.1352	0.0861	0.0672	0.0675	0.0003	0.0012	0.9883
SKCVL	Skeletal and Connective, Very Low	0.1035	0.1104	0.0069	0.0226	0.7738	0.1021	0.1104	0.0083	0.0282	0.7476
SKNH	Skin, High	0.0150	0.0061	-0.0088	-0.0730	0.3540	0.0058	0.0061	0.0003	0.0040	0.9628
SKNL	Skin, Low	0.0323	0.0061	-0.0262	-0.1488	0.0590	0.0067	0.0061	-0.0006	-0.0085	0.9301
SKNVL	Skin, Very Low	0.1143	0.1288	0.0146	0.0458	0.5610	0.1199	0.1288	0.0089	0.0288	0.7468
SUBL	Substance Abuse, Low	0.2961	0.2699	-0.0262	-0.0573	0.4667	0.2877	0.2699	-0.0178	-0.0414	0.6277
SUBVL	Substance Abuse, Very Low	0.0595	0.1043	0.0448	0.1884	0.0168	0.1089	0.1043	-0.0046	-0.0161	0.8594
	Medicaid Rx categories										
MRX1	Anti-coagulants	0.0204	0.0061	-0.0143	-0.1015	0.1975	0.0077	0.0061	-0.0015	-0.0191	0.8355
MRX2	Cardiac	0.0274	0.0184	-0.0090	-0.0550	0.4852	0.0239	0.0184	-0.0055	-0.0400	0.6652
MRX3	Psychosis/Bipolar/ Depression	0.0821	0.0123	-0.0698	-0.2555	0.0012	0.0086	0.0123	0.0037	0.0297	0.6121
MRX4	Diabetes	0.0068	0.0184	0.0116	0.1390	0.0777	0.0172	0.0184	0.0012	0.0101	0.9093
MRX5	ESRD/Renal	0.0000	0.0000	0.0000	0.0000	1.0000	0.0000	0.0000	0.0000	0.0000	1.0000
MRX6	Hemophilia/von Willebrands	0.0000	0.0000	0.0000	0.0000	1.0000	0.0000	0.0000	0.0000	0.0000	1.0000
MRX7	Hepatitis	0.0020	0.0061	0.0041	0.0908	0.2493	0.0028	0.0061	0.0034	0.0663	0.4951
MRX8	HIV	0.0000	0.0000	0.0000	0.0000	1.0000	0.0000	0.0000	0.0000	0.0000	1.0000
MRX9	Infections, high	0.0041	0.0061	0.0020	0.0317	0.6876	0.0071	0.0061	-0.0009	-0.0105	0.8931
MRX10	Inflammatory/ Autoimmune	0.0000	0.0000	0.0000	0.0000	1.0000	0.0000	0.0000	0.0000	0.0000	1.0000
MRX11	Malignancies	0.0153	0.0123	-0.0031	-0.0250	0.7513	0.0110	0.0123	0.0012	0.0117	0.8896
MRX12	Multiple Sclerosis/ Paralysis	0.0000	0.0000	0.0000	0.0000	1.0000	0.0000	0.0000	0.0000	0.0000	1.0000
MRX13	Parkinson's/tremor	0.0409	0.1534	0.1125	0.5605	0.0000	0.1629	0.1534	-0.0095	-0.0310	0.7534
MRX14	Seizure disorders	0.0324	0.0920	0.0596	0.3335	0.0000	0.0690	0.0920	0.0230	0.0918	0.2819
MRX15	Tuberculosis	0.0025	0.0061	0.0036	0.0715	0.3640	0.0061	0.0061	0.0000	0.0000	1.0000
	Number of ED visits in 12 months prior to program enrollment										
BASE_ED_ VST_CAT0	0	0.3377	0.2638	-0.0738	-0.1563	0.0473	0.2586	0.2638	0.0052	0.0117	0.8864
BASE_ED_ VST_CAT1	1	0.2138	0.1411	-0.0726	-0.1775	0.0243	0.1485	0.1411	-0.0074	-0.0206	0.8030
BASE_ED_ VST_CAT2	2-3	0.2165	0.2883	0.0719	0.1742	0.0270	0.2847	0.2883	0.0037	0.0082	0.9219
BASE_ED_ VST_CAT3	4+	0.2321	0.3067	0.0746	0.1766	0.0250	0.3083	0.3067	-0.0015	-0.0034	0.9684

		Before matching					After matching				
Variable name	Variable description	Comparison	Intervention	adj.diff	std.diff	p	Comparison	Intervention	adj.diff	std.diff	p
	Number of hospitalizations in 12 months prior to program enrollment										
BASE_HOSP_STAY_CAT0	0	0.4709	0.5092	0.0383	0.0767	0.3303	0.4607	0.5092	0.0485	0.0970	0.2219
BASE_HOSP_STAY_CAT1	1	0.1995	0.1779	-0.0216	-0.0540	0.4928	0.1807	0.1779	-0.0028	-0.0073	0.9315
BASE_HOSP_STAY_CAT2	2	0.1035	0.0859	-0.0176	-0.0577	0.4637	0.0972	0.0859	-0.0113	-0.0413	0.6457
BASE_HOSP_STAY_CAT3	3+	0.2261	0.2270	0.0009	0.0021	0.9792	0.2613	0.2270	-0.0344	-0.0822	0.3350
	Relative number of psychiatric visits in 12 months prior to program enrollment[a]										
BASE_PSYCH_VST_CAT0	0	0.5452	0.1043	-0.4409	-0.8886	0.0000	0.1098	0.1043	-0.0055	-0.0151	0.7941
BASE_PSYCH_VST_CAT1	Up to 20%	0.1843	0.1043	-0.0800	-0.2068	0.0087	0.1178	0.1043	-0.0135	-0.0387	0.5928
BASE_PSYCH_VST_CAT2	20-40%	0.1271	0.1902	0.0631	0.1891	0.0164	0.2064	0.1902	-0.0163	-0.0407	0.6254
BASE_PSYCH_VST_CAT3	40-60%	0.0472	0.1902	0.1430	0.6654	0.0000	0.1880	0.1902	0.0021	0.0058	0.9475
BASE_PSYCH_VST_CAT4	60-80%	0.0433	0.2025	0.1591	0.7686	0.0000	0.1684	0.2025	0.0340	0.0933	0.2712
BASE_PSYCH_VST_CAT5	80-100%	0.0529	0.2086	0.1557	0.6863	0.0000	0.2095	0.2086	-0.0009	-0.0024	0.9781
	Relative number of office visits in 12 months prior to program enrollment[a]										
BASE_OFFICE_VST_CAT0	0	0.0842	0.0675	-0.0167	-0.0601	0.4453	0.0641	0.0675	0.0034	0.0140	0.8683
BASE_OFFICE_VST_CAT1	Up to 20%	0.1696	0.1718	0.0022	0.0058	0.9418	0.1773	0.1718	-0.0055	-0.0147	0.8633
BASE_OFFICE_VST_CAT2	20-40%	0.1806	0.1166	-0.0640	-0.1667	0.0344	0.1227	0.1166	-0.0061	-0.0186	0.8225
BASE_OFFICE_VST_CAT3	40-60%	0.1752	0.1779	0.0028	0.0072	0.9268	0.1874	0.1779	-0.0095	-0.0237	0.7658
BASE_OFFICE_VST_CAT4	60-80%	0.1554	0.2025	0.0470	0.1296	0.1000	0.1758	0.2025	0.0267	0.0702	0.4129
BASE_OFFICE_VST_CAT5	80-100%	0.2351	0.2638	0.0287	0.0677	0.3898	0.2727	0.2638	-0.0089	-0.0201	0.8104

		Before matching					After matching				
Variable name	Variable description	Comparison	Intervention	adj.diff	std.diff	p	Comparison	Intervention	adj.diff	std.diff	p
	Relative total Medicaid expenditures in 12 months prior to program enrollment[a]										
BASE_TOTAL_PMT_CAT0	0	0.0148	0.0061	-0.0086	-0.0718	0.3624	0.0034	0.0061	0.0028	0.0328	0.5997
BASE_TOTAL_PMT_CAT1	Up to 20%	0.2978	0.1472	-0.1505	-0.3299	0.0000	0.1678	0.1472	-0.0206	-0.0499	0.4728
BASE_TOTAL_PMT_CAT2	20-40%	0.1863	0.2086	0.0223	0.0572	0.4681	0.1785	0.2086	0.0301	0.0754	0.3487
BASE_TOTAL_PMT_CAT3	40-60%	0.1351	0.1718	0.0367	0.1073	0.1733	0.1644	0.1718	0.0074	0.0201	0.8155
BASE_TOTAL_PMT_CAT4	60-80%	0.1658	0.2147	0.0489	0.1314	0.0955	0.2288	0.2147	-0.0141	-0.0351	0.6834
BASE_TOTAL_PMT_CAT5	80-100%	0.2003	0.2515	0.0512	0.1279	0.1045	0.2571	0.2515	-0.0055	-0.0134	0.8765

[a] For each of office visits, psychiatric visits, and total Medicaid expenditures, we created categories ranking by the distribution of use (by percentile) among the intervention and potential comparison group members.

Table A-III.4.b. Balance table before and after matching (FPHNY, NA-MCT subgroup)

		Before matching					After matching				
Variable name	Variable description	Comparison	Intervention	adj.diff	std.diff	P	Comparison	Intervention	adj.diff	std.diff	p
					Exact match variables						
	Service use category prior to program enrollment										
SERV_USE_CAT1	Hospitalization or ED visit leading to hospitalization	0.0919	0.2976	0.2057	0.7107	0.0000	0.2976	0.2976	0.0000	0.0000	1.0000
SERV_USE_CAT3	ED visit not leading to hospitalization	0.1207	0.1786	0.0579	0.1776	0.1039	0.1786	0.1786	0.0000	0.0000	1.0000
SERV_USE_CAT4	Psychiatric visit	0.5271	0.3690	-0.1581	-0.3166	0.0038	0.3690	0.3690	0.0000	0.0000	1.0000
SERV_USE_CAT5	Office visit	0.2603	0.1548	-0.1055	-0.2406	0.0277	0.1548	0.1548	0.0000	0.0000	1.0000
	Hierarchical diagnosis code prior to program enrollment										
DX_CAT1	Schizophrenia	0.6050	0.6905	0.0854	0.1748	0.1096	0.6905	0.6905	0.0000	0.0000	1.0000
DX_CAT2	Bipolar	0.2978	0.1905	-0.1073	-0.2347	0.0317	0.1905	0.1905	0.0000	0.0000	1.0000
DX_CAT3	Drug/alcohol-induced psychosis	0.0104	0.0119	0.0015	0.0145	0.8947	0.0119	0.0119	0.0000	0.0000	1.0000
DX_CAT4	Other nonorganic psychosis	0.0026	0.0476	0.0450	0.8718	0.0000	0.0476	0.0476	0.0000	0.0000	1.0000
DX_CAT5	Suicidal	0.0003	0.0119	0.0116	0.6862	0.0000	0.0119	0.0119	0.0000	0.0000	1.0000
DX_CAT6	Depression	0.0394	0.0119	-0.0275	-0.1413	0.1957	0.0119	0.0119	0.0000	0.0000	1.0000
DX_CAT7	Anxiety/adjustment	0.0445	0.0357	-0.0088	-0.0428	0.6952	0.0357	0.0357	0.0000	0.0000	1.0000
PRE_12MN	12 months of continuous baseline enrollment	0.9267	0.9167	-0.0101	-0.0386	0.7239	0.9167	0.9167	0.0000	0.0000	1.0000
					Propensity score variables						
	Program enrollment month										
TRIGGER_MN 11/01/2012	11/2012	0.0000	0.0000	0.0000	0.0000	1.0000	0.0000	0.0000	0.0000	0.0000	1.0000
TRIGGER_MN 12/01/2012	12/2012	0.0000	0.0000	0.0000	0.0000	1.0000	0.0000	0.0000	0.0000	0.0000	1.0000
TRIGGER_MN 01/01/2013	01/2013	0.0170	0.0119	-0.0051	-0.0394	0.7183	0.0113	0.0119	0.0006	0.0054	0.9600
TRIGGER_MN 03/01/2013	03/2013	0.0303	0.0238	-0.0065	-0.0378	0.7291	0.0262	0.0238	-0.0024	-0.0150	0.9033
TRIGGER_MN 04/01/2013	04/2013	0.0000	0.0000	0.0000	0.0000	1.0000	0.0000	0.0000	0.0000	0.0000	1.0000
TRIGGER_MN 05/01/2013	05/2013	0.0465	0.0476	0.0011	0.0054	0.9605	0.0482	0.0476	-0.0006	-0.0027	0.9800
TRIGGER_MN 06/01/2013	06/2013	0.0337	0.0357	0.0020	0.0110	0.9200	0.0381	0.0357	-0.0024	-0.0122	0.9115
TRIGGER_MN 07/01/2013	07/2013	0.0374	0.0238	-0.0136	-0.0718	0.5110	0.0232	0.0238	0.0006	0.0038	0.9729

		Before matching					After matching				
Variable name	Variable description	Comparison	Intervention	adj.diff	std.diff	P	Comparison	Intervention	adj.diff	std.diff	p
TRIGGER_MN 08/01/2013	08/2013	0.0457	0.0476	0.0019	0.0091	0.9340	0.0494	0.0476	-0.0018	-0.0086	0.9416
TRIGGER_MN 09/01/2013	09/2013	0.0478	0.0714	0.0236	0.1108	0.3104	0.0655	0.0714	0.0060	0.0237	0.8292
TRIGGER_MN 10/01/2013	10/2013	0.0449	0.0357	-0.0092	-0.0446	0.6833	0.0458	0.0357	-0.0101	-0.0494	0.6616
TRIGGER_MN 11/01/2013	11/2013	0.0431	0.0714	0.0283	0.1394	0.2019	0.0631	0.0714	0.0083	0.0334	0.7693
TRIGGER_MN 12/01/2013	12/2013	0.0425	0.0595	0.0170	0.0842	0.4406	0.0619	0.0595	-0.0024	-0.0099	0.9355
TRIGGER_MN 01/01/2014	01/2014	0.0568	0.0833	0.0266	0.1147	0.2936	0.0839	0.0833	-0.0006	-0.0021	0.9851
TRIGGER_MN 02/01/2014	02/2014	0.0461	0.0833	0.0372	0.1772	0.1048	0.0750	0.0833	0.0083	0.0309	0.7771
TRIGGER_MN 03/01/2014	03/2014	0.0532	0.0833	0.0301	0.1341	0.2196	0.0857	0.0833	-0.0024	-0.0090	0.9401
TRIGGER_MN 04/01/2014	04/2014	0.0776	0.1071	0.0295	0.1103	0.3125	0.1024	0.1071	0.0048	0.0161	0.8901
TRIGGER_MN 05/01/2014	05/2014	0.0710	0.0952	0.0242	0.0941	0.3889	0.0857	0.0952	0.0095	0.0349	0.7631
TRIGGER_MN 06/01/2014	06/2014	0.0726	0.0357	-0.0369	-0.1423	0.1928	0.0417	0.0357	-0.0060	-0.0297	0.7860
TRIGGER_MN 07/01/2014	07/2014	0.0469	0.0119	-0.0350	-0.1656	0.1295	0.0089	0.0119	0.0030	0.0301	0.7728
TRIGGER_MN 08/01/2014	08/2014	0.0361	0.0119	-0.0242	-0.1298	0.2347	0.0107	0.0119	0.0012	0.0111	0.9193
TRIGGER_MN 09/01/2014	09/2014	0.0373	0.0238	-0.0135	-0.0712	0.5148	0.0321	0.0238	-0.0083	-0.0464	0.6687
TRIGGER_MN 10/01/2014	10/2014	0.0000	0.0000	0.0000	0.0000	1.0000	0.0000	0.0000	0.0000	0.0000	1.0000
TRIGGER_MN 11/01/2014	11/2014	0.0638	0.0238	-0.0400	-0.1637	0.1339	0.0220	0.0238	0.0018	0.0116	0.9119
TRIGGER_MN 12/01/2014	12/2014	0.0495	0.0119	-0.0376	-0.1736	0.1120	0.0190	0.0119	-0.0071	-0.0521	0.6358
	Age group										
NYEM_AGE_ CAT1	<18	0.0822	0.1429	0.0607	0.2208	0.0433	0.1429	0.1429	0.0000	0.0000	1.0000
NYEM_AGE_ CAT2	18-34	0.2856	0.5595	0.2740	0.6064	0.0000	0.5292	0.5595	0.0304	0.0613	0.4747
NYEM_AGE_ CAT3	34-45	0.1759	0.0714	-0.1045	-0.2747	0.0119	0.0863	0.0714	-0.0149	-0.0528	0.6219
NYEM_AGE_ CAT4	45-64	0.4563	0.2262	-0.2301	-0.4622	0.0000	0.2417	0.2262	-0.0155	-0.0358	0.7196
NYEM_AGE_ CAT5	65+	0.0000	0.0000	0.0000	0.0000	1.0000	0.0000	0.0000	0.0000	0.0000	1.0000
NYEM_MALE	Male	0.5036	0.6429	0.1393	0.2786	0.0108	0.6619	0.6429	-0.0190	-0.0406	0.7146

Variable name	Variable description	Before matching					After matching				
		Comparison	Intervention	adj.diff	std.diff	P	Comparison	Intervention	adj.diff	std.diff	p
	Race/ethnicity										
NYEM_BLACK	African American, not Hispanic	0.3149	0.3810	0.0660	0.1421	0.1932	0.3702	0.3810	0.0107	0.0221	0.8444
NYEM_ HISPANIC	Hispanic	0.3654	0.3333	-0.0320	-0.0665	0.5426	0.3274	0.3333	0.0060	0.0126	0.9076
NYEM_ DISABLED	Disabled	0.5484	0.5000	-0.0484	-0.0972	0.3735	0.4905	0.5000	0.0095	0.0190	0.8569
	CDPS condition flags										
AIDSH	AIDS, High	0.1131	0.0714	-0.0417	-0.1317	0.2279	0.0798	0.0714	-0.0083	-0.0297	0.7795
CANH	Cancer, High	0.0000	0.0000	0.0000	0.0000	1.0000	0.0000	0.0000	0.0000	0.0000	1.0000
CANL	Cancer, Low	0.0000	0.0000	0.0000	0.0000	1.0000	0.0000	0.0000	0.0000	0.0000	1.0000
CANM	Cancer, Medium	0.0000	0.0000	0.0000	0.0000	1.0000	0.0000	0.0000	0.0000	0.0000	1.0000
CANVH	Cancer, Very High	0.0000	0.0000	0.0000	0.0000	1.0000	0.0000	0.0000	0.0000	0.0000	1.0000
CAREL	Cardiovascular, Extra Low	0.1942	0.1071	-0.0870	-0.2201	0.0439	0.1125	0.1071	-0.0054	-0.0164	0.8756
CARL	Cardiovascular, Low	0.1647	0.1905	0.0258	0.0695	0.5248	0.1810	0.1905	0.0095	0.0241	0.8208
CARM	Cardiovascular, Medium	0.0351	0.0119	-0.0232	-0.1261	0.2484	0.0125	0.0119	-0.0006	-0.0054	0.9613
CARVH	Cardiovascular, Very High	0.0000	0.0000	0.0000	0.0000	1.0000	0.0000	0.0000	0.0000	0.0000	1.0000
CERL	Cerebrovascular, Low	0.0182	0.0119	-0.0063	-0.0474	0.6647	0.0101	0.0119	0.0018	0.0175	0.9117
CNSH	Nervous System, High	0.0000	0.0000	0.0000	0.0000	1.0000	0.0000	0.0000	0.0000	0.0000	1.0000
CNSL	Nervous System, Low	0.1216	0.0952	-0.0263	-0.0806	0.4606	0.1036	0.0952	-0.0083	-0.0267	0.8031
CNSM	Nervous System, Medium	0.0228	0.0119	-0.0109	-0.0732	0.5027	0.0131	0.0119	-0.0012	-0.0101	0.9249
DDL	Developmental Disability, Low	0.0258	0.0119	-0.0139	-0.0879	0.4211	0.0214	0.0119	-0.0095	-0.0684	0.5453
DDM	Developmental Disability, Medium	0.0000	0.0000	0.0000	0.0000	1.0000	0.0000	0.0000	0.0000	0.0000	1.0000
DIA1H	Diabetes, Type 1 High	0.0000	0.0000	0.0000	0.0000	1.0000	0.0000	0.0000	0.0000	0.0000	1.0000
DIA1M	Diabetes, Type 1 Low	0.0234	0.0238	0.0005	0.0030	0.9780	0.0244	0.0238	-0.0006	-0.0037	0.9723
DIA2L	Diabetes, Type 2 High	0.1239	0.0476	-0.0763	-0.2317	0.0340	0.0488	0.0476	-0.0012	-0.0055	0.9592
DIA2M	Diabetes, Type 2 Low	0.0298	0.0119	-0.0179	-0.1053	0.3350	0.0054	0.0119	0.0065	0.0836	0.4297
EYEL	Eye, Low	0.0000	0.0000	0.0000	0.0000	1.0000	0.0000	0.0000	0.0000	0.0000	1.0000
EYEVL	Eye, Very Low	0.0465	0.0357	-0.0107	-0.0510	0.6404	0.0250	0.0357	0.0107	0.0659	0.6124
GENEL	Genital, Extra Low	0.0704	0.0952	0.0249	0.0971	0.3740	0.0756	0.0952	0.0196	0.0715	0.4995
GIH	Gastrointestinal, High	0.0000	0.0000	0.0000	0.0000	1.0000	0.0000	0.0000	0.0000	0.0000	1.0000
GIL	Gastrointestinal, Low	0.1444	0.0833	-0.0610	-0.1738	0.1117	0.0946	0.0833	-0.0113	-0.0378	0.7244
GIM	Gastrointestinal, Medium	0.0423	0.0238	-0.0185	-0.0919	0.4004	0.0232	0.0238	0.0006	0.0038	0.9720
HEMEH	Hematological, Extra High	0.0000	0.0000	0.0000	0.0000	1.0000	0.0000	0.0000	0.0000	0.0000	1.0000
HEML	Hematological, Low	0.0192	0.0238	0.0046	0.0334	0.7598	0.0208	0.0238	0.0030	0.0199	0.8544
HEMM	Hematological, Medium	0.0169	0.0119	-0.0050	-0.0390	0.7209	0.0149	0.0119	-0.0030	-0.0237	0.8241
HEMVH	Hematological, Very High	0.0000	0.0000	0.0000	0.0000	1.0000	0.0000	0.0000	0.0000	0.0000	1.0000
HIVM	HIV, Medium	0.0000	0.0000	0.0000	0.0000	1.0000	0.0000	0.0000	0.0000	0.0000	1.0000
INFH	Infectious, High	0.0000	0.0000	0.0000	0.0000	1.0000	0.0000	0.0000	0.0000	0.0000	1.0000
INFL	Infectious, Low	0.0422	0.0119	-0.0303	-0.1508	0.1674	0.0137	0.0119	-0.0018	-0.0148	0.8885
INFM	Infectious, Medium	0.0000	0.0000	0.0000	0.0000	1.0000	0.0000	0.0000	0.0000	0.0000	1.0000
METH	Metabolic, High	0.0117	0.0119	0.0002	0.0018	0.9865	0.0167	0.0119	-0.0048	-0.0365	0.7380
METM	Metabolic, Medium	0.0245	0.0119	-0.0125	-0.0813	0.4568	0.0131	0.0119	-0.0012	-0.0101	0.9246

		Before matching					After matching				
Variable name	Variable description	Comparison	Intervention	adj.diff	std.diff	P	Comparison	Intervention	adj.diff	std.diff	p
METVL	Metabolic, Very Low	0.0282	0.0238	-0.0044	-0.0266	0.8077	0.0256	0.0238	-0.0018	-0.0123	0.9204
PRGCMP	Pregnancy, Complete	0.0340	0.0357	0.0018	0.0097	0.9295	0.0298	0.0357	0.0060	0.0347	0.7547
PRGINC	Pregnancy, Incomplete	0.0120	0.0238	0.0118	0.1084	0.3211	0.0327	0.0238	-0.0089	-0.0489	0.6418
PSYH	Psychiatric, High	0.4252	0.5595	0.1343	0.2716	0.0129	0.5512	0.5595	0.0083	0.0169	0.7799
PSYL	Psychiatric, Low	0.0374	0.0238	-0.0136	-0.0717	0.5118	0.0280	0.0238	-0.0042	-0.0275	0.8177
PSYM	Psychiatric, Medium	0.1315	0.1429	0.0113	0.0335	0.7588	0.1464	0.1429	-0.0036	-0.0109	0.9122
PSYML	Psychiatric, Medium Low	0.1900	0.0952	-0.0948	-0.2417	0.0269	0.0833	0.0952	0.0119	0.0424	0.6780
PULH	Pulmonary, High	0.0000	0.0000	0.0000	0.0000	1.0000	0.0000	0.0000	0.0000	0.0000	1.0000
PULL	Pulmonary, Low	0.2416	0.2262	-0.0155	-0.0361	0.7410	0.2292	0.2262	-0.0030	-0.0071	0.9500
PULM	Pulmonary, Medium	0.0354	0.0238	-0.0116	-0.0628	0.5655	0.0232	0.0238	0.0006	0.0038	0.9720
PULVH	Pulmonary, Very High	0.0000	0.0000	0.0000	0.0000	1.0000	0.0000	0.0000	0.0000	0.0000	1.0000
RENEH	Renal, Extra High	0.0000	0.0000	0.0000	0.0000	1.0000	0.0000	0.0000	0.0000	0.0000	1.0000
RENL	Renal, Low	0.0000	0.0000	0.0000	0.0000	1.0000	0.0000	0.0000	0.0000	0.0000	1.0000
RENM	Renal, Medium	0.0000	0.0000	0.0000	0.0000	1.0000	0.0000	0.0000	0.0000	0.0000	1.0000
RENVH	Renal, Very High	0.0000	0.0000	0.0000	0.0000	1.0000	0.0000	0.0000	0.0000	0.0000	1.0000
SKCL	Skeletal and Connective, Low	0.0862	0.0833	-0.0029	-0.0102	0.9256	0.0762	0.0833	0.0071	0.0262	0.8090
SKCM	Skeletal and Connective, Medium	0.0535	0.0357	-0.0178	-0.0790	0.4697	0.0375	0.0357	-0.0018	-0.0090	0.9319
SKCVL	Skeletal and Connective, Very Low	0.0804	0.0357	-0.0447	-0.1644	0.1324	0.0369	0.0357	-0.0012	-0.0062	0.9543
SKNH	Skin, High	0.0000	0.0000	0.0000	0.0000	1.0000	0.0000	0.0000	0.0000	0.0000	1.0000
SKNL	Skin, Low	0.0000	0.0000	0.0000	0.0000	1.0000	0.0000	0.0000	0.0000	0.0000	1.0000
SKNVL	Skin, Very Low	0.0912	0.0952	0.0041	0.0141	0.8970	0.0929	0.0952	0.0024	0.0080	0.9410
SUBL	Substance Abuse, Low	0.2421	0.1905	-0.0517	-0.1206	0.2695	0.2196	0.1905	-0.0292	-0.0688	0.4979
SUBVL	Substance Abuse, Very Low	0.0556	0.0714	0.0159	0.0693	0.5260	0.0726	0.0714	-0.0012	-0.0045	0.9671
	Medicaid Rx categories										
MRX1	Anti-coagulants	0.0000	0.0000	0.0000	0.0000	1.0000	0.0000	0.0000	0.0000	0.0000	1.0000
MRX2	Cardiac	0.0312	0.0119	-0.0193	-0.1110	0.3095	0.0060	0.0119	0.0060	0.0760	0.4982
MRX3	Psychosis/bipolar/ depression	0.0336	0.0119	-0.0217	-0.1207	0.2694	0.0095	0.0119	0.0024	0.0241	0.8278
MRX4	Diabetes	0.0000	0.0000	0.0000	0.0000	1.0000	0.0000	0.0000	0.0000	0.0000	1.0000
MRX5	ESRD/Renal	0.0000	0.0000	0.0000	0.0000	1.0000	0.0000	0.0000	0.0000	0.0000	1.0000
MRX6	Hemophilia/von Willebrands	0.0000	0.0000	0.0000	0.0000	1.0000	0.0000	0.0000	0.0000	0.0000	1.0000
MRX7	Hepatitis	0.0000	0.0000	0.0000	0.0000	1.0000	0.0000	0.0000	0.0000	0.0000	1.0000
MRX8	HIV	0.0000	0.0000	0.0000	0.0000	1.0000	0.0000	0.0000	0.0000	0.0000	1.0000
MRX9	Infections, high	0.0000	0.0000	0.0000	0.0000	1.0000	0.0000	0.0000	0.0000	0.0000	1.0000
MRX10	Inflammatory/ Autoimmune	0.0000	0.0000	0.0000	0.0000	1.0000	0.0000	0.0000	0.0000	0.0000	1.0000
MRX11	Malignancies	0.0000	0.0000	0.0000	0.0000	1.0000	0.0000	0.0000	0.0000	0.0000	1.0000
MRX12	Multiple Sclerosis/ Paralysis	0.0000	0.0000	0.0000	0.0000	1.0000	0.0000	0.0000	0.0000	0.0000	1.0000
MRX13	Parkinson's/Tremor	0.1155	0.1548	0.0393	0.1228	0.2611	0.1536	0.1548	0.0012	0.0032	0.9752
MRX14	Seizure disorders	0.0410	0.0595	0.0185	0.0933	0.3932	0.0518	0.0595	0.0077	0.0336	0.7537

		Before matching					After matching				
Variable name	Variable description	Comparison	Intervention	adj.diff	std.diff	P	Comparison	Intervention	adj.diff	std.diff	p
MRX15	Tuberculosis	0.0000	0.0000	0.0000	0.0000	1.0000	0.0000	0.0000	0.0000	0.0000	1.0000
	Number of ED visits in 12 months prior to program enrollment										
BASE_ED_VST_CAT0	0	0.3921	0.2976	-0.0944	-0.1935	0.0765	0.2577	0.2976	0.0399	0.0919	0.2825
BASE_ED_VST_CAT1	1	0.2308	0.2381	0.0073	0.0172	0.8747	0.2601	0.2381	-0.0220	-0.0503	0.6449
BASE_ED_VST_CAT2	2-3	0.2016	0.1548	-0.0468	-0.1167	0.2855	0.1708	0.1548	-0.0161	-0.0432	0.6877
BASE_ED_VST_CAT3	4+	0.1755	0.3095	0.1340	0.3520	0.0013	0.3113	0.3095	-0.0018	-0.0038	0.9711
	Number of hospitalizations in 12 months prior to program enrollment										
BASE_HOSP_STAY_CAT0	0	0.6490	0.2976	-0.3514	-0.7364	0.0000	0.2810	0.2976	0.0167	0.0367	0.5079
BASE_HOSP_STAY_CAT1	1	0.1739	0.3333	0.1594	0.4203	0.0001	0.3173	0.3333	0.0161	0.0350	0.7355
BASE_HOSP_STAY_CAT2	2	0.0764	0.1905	0.1141	0.4288	0.0001	0.1964	0.1905	-0.0060	-0.0148	0.8912
BASE_HOSP_STAY_CAT3	3+	0.1006	0.1786	0.0779	0.2588	0.0178	0.2054	0.1786	-0.0268	-0.0660	0.5257
	Relative number of psychiatric visits in 12 months prior to program enrollment[a]										
BASE_PSYCH_VST_CAT0	0	0.2889	0.1667	-0.1223	-0.2698	0.0135	0.1732	0.1667	-0.0065	-0.0181	0.7615
BASE_PSYCH_VST_CAT1	Up to 20%	0.1534	0.2619	0.1085	0.3009	0.0059	0.2756	0.2619	-0.0137	-0.0304	0.7660
BASE_PSYCH_VST_CAT2	20-40%	0.2053	0.2143	0.0090	0.0222	0.8389	0.2220	0.2143	-0.0077	-0.0186	0.8648
BASE_PSYCH_VST_CAT3	40-60%	0.1267	0.1548	0.0280	0.0843	0.4404	0.1393	0.1548	0.0155	0.0440	0.6892
BASE_PSYCH_VST_CAT4	60-80%	0.1041	0.1071	0.0030	0.0098	0.9284	0.0911	0.1071	0.0161	0.0540	0.6122
BASE_PSYCH_VST_CAT5	80-100%	0.1215	0.0952	-0.0262	-0.0804	0.4619	0.0988	0.0952	-0.0036	-0.0119	0.9124
	Relative number of office visits in 12 months prior to program enrollment[a]										
BASE_OFFICE_VST_CAT0	0	0.0620	0.0714	0.0095	0.0392	0.7196	0.0899	0.0714	-0.0185	-0.0663	0.5599

		Before matching					After matching				
Variable name	Variable description	Comparison	Intervention	adj.diff	std.diff	P	Comparison	Intervention	adj.diff	std.diff	p
BASE_OFFICE_VST_CAT1	Up to 20%	0.1844	0.2143	0.0299	0.0772	0.4799	0.2357	0.2143	-0.0214	-0.0510	0.6569
BASE_OFFICE_VST_CAT2	20-40%	0.2172	0.2143	-0.0029	-0.0071	0.9480	0.2065	0.2143	0.0077	0.0188	0.8656
BASE_OFFICE_VST_CAT3	40-60%	0.2434	0.2976	0.0542	0.1262	0.2479	0.2714	0.2976	0.0262	0.0583	0.5971
BASE_OFFICE_VST_CAT4	60-80%	0.1454	0.1190	-0.0264	-0.0748	0.4934	0.0994	0.1190	0.0196	0.0651	0.5577
BASE_OFFICE_VST_CAT5	80-100%	0.1476	0.0833	-0.0643	-0.1812	0.0971	0.0970	0.0833	-0.0137	-0.0479	0.6494
	Relative total Medicaid expenditures in 12 months prior to program enrollment[a]										
BASE_TOTAL_PMT_CAT0	0	0.0000	0.0000	0.0000	0.0000	1.0000	0.0000	0.0000	0.0000	0.0000	1.0000
BASE_TOTAL_PMT_CAT1	Up to 20%	0.4439	0.2500	-0.1939	-0.3904	0.0004	0.2381	0.2500	0.0119	0.0277	0.7576
BASE_TOTAL_PMT_CAT2	20-40%	0.1878	0.2381	0.0503	0.1288	0.2382	0.2202	0.2381	0.0179	0.0433	0.7037
BASE_TOTAL_PMT_CAT3	40-60%	0.1606	0.2143	0.0537	0.1462	0.1809	0.2107	0.2143	0.0036	0.0087	0.9381
BASE_TOTAL_PMT_CAT4	60-80%	0.1140	0.1786	0.0646	0.2032	0.0629	0.1804	0.1786	-0.0018	-0.0046	0.9676
BASE_TOTAL_PMT_CAT5	80-100%	0.0937	0.1190	0.0253	0.0868	0.4268	0.1506	0.1190	-0.0315	-0.0897	0.4221

[a] For each of office visits, psychiatric visits, and total Medicaid expenditures, we created categories ranking by the distribution of use (by percentile) among the intervention and potential comparison group members.

E. Specifications of measures

CMMI requested that we calculate four standardized outcome measures for all awardees to the extent feasible. These measures are: total Medicare and/or Medicaid expenditures, inpatient hospitalizations, hospital readmissions, and ED visits. If it was possible to calculate these core measures identified by CMS, and if the measures were appropriate to the intervention, we used them. Our specifications for these measures in Medicare and Medicaid administrative data are described in Section E.1 below. Our analyses used several other types of measures as control variables, described in Section E.2.

1. Core measures

For FPHNY, we were able to develop monthly measures for three of the standardized outcomes: total Medicaid expenditures, inpatient hospitalizations, and ED visits. We determined that our sample size was too small to detect effects in readmissions, and therefore we did not include the measure in our analysis.

We first describe how we identified observation periods, then describe how each of these three outcome measures were constructed for each period.

a. Identifying observation periods

Baseline and intervention periods were defined for each intervention or comparison group member relative to their program enrollment (or pseudo-enrollment) month. The first intervention period was defined as the enrollment month and the five months[90] following that month. Where applicable, the second intervention period was the six months following the first intervention period. The first baseline period was the six months prior to the program enrollment or start month and additional baseline periods were identified by moving backward six months from the first baseline period. For each individual included in the analysis, the proportion of each baseline and intervention period for which the individual was eligible for the analysis was calculated. This proportion was used to pro-rate the expenditure and utilization measures for individuals enrolled for less than the full analysis period. It was also used to weight observations in the regression analysis.

b. Calculating total Medicaid expenditures

All claims in the NY Medicaid administrative data were considered for inclusion in the analysis; however, duplicate and denied claims were excluded. The total cost of care was based on the total amount paid to the provider for the approved claim. For claims with services spanning more than one day, all expenditures were counted on the first date of service. These expenditures include both fee-for-service and managed care payments, but do not include capitated payments. When service level payment information was not available for managed care covered services, these payment amounts were estimated based on fee-for-service payment guidelines.

[90] For brevity, we only discuss the six-month baseline and intervention periods created for the analysis of trends in regression-adjusted means. We also defined a second set of baseline and intervention periods and weights corresponding to years from program enrollment, used in the impact and total savings table in the report (Table V.2).

c. Calculating hospitalizations

The specifications for hospitalization counts were developed to align with the CMMI priority all-cause admissions per patient measure. We describe the steps to develop these counts here.

Step 1: Identify hospitalization claims. Inpatient hospital claims were identified by using the Medicaid Managed Care Operating Report code (MMCOR_CD) values of 01 ("Inpatient Psych, Acute Detox Subabuse") or 04 ("Medical/surgical"), Surveillance and Utilization Review System Category of Service code (SURS_SUBSYSTEM_COS_CD) value of 11 ("Inpatient"), and the eMedNY claim type code (CLAIM_TYPE_CD) value of "I" ("Inpatient").

Step 2: Eliminate duplicate or denied claims. We identified claims with the same information in all fields and only kept one of these claims. We also excluded denied claims from our analysis.

Step 3: Combine claims that represent the same stay and combine transfer stays with initial stays. We identified and combined initial and interim claims into one discharge. Interim claims had (1) the same admission date as the initial claim; (2) an admission date that was equal to the discharge date from the initial or another interim claim and the status on the other (previous) claim was "still a patient"; or (3) a claim with an admission date that was equal to one day after the discharge date of the initial or another interim claim and the status on the other previous claim was "still a patient." Such claims were combined to count as a single stay.

Next, we identified and combined claims associated with a transfer into a single stay. We identified claims indicating that the patient was transferred to either another short-term hospital, a Critical Access Hospital (CAH), another type institution for inpatient care, a federal hospital, or a psychiatric hospital or unit. Then we combined these claims with claims for the same beneficiary at a different facility where the admission date fell within one day of the discharge date of the first claim.

Step 4: Sum the number of discharges in each month. Once claims representing a single stay were combined, we summed the number of unique discharges for each enrollee for each month. Inpatient stays were counted in the month of the discharge date.

d. ED visits

Outpatient ED visit utilization is reflected in CMMI priority measure 62. This measure includes ED visits and observation stays that do not lead to an inpatient stay.

We reviewed claims not identified as inpatient and considered them as ED visits if the procedure code, cost center revenue code, or managed care operating report code indicated ED visit.

ED visits that led to inpatient stays (i.e., ones that overlapped with or were adjacent to an inpatient stay) were excluded. If two or more ED visits or observation stays had the same patient identifier and beginning date of service, we counted them as one visit.

2. Other measures

The control variables included in the FPHNY regression models are listed in Table A-III.5 along with the specifications for the variables.

Table A-III.5. Impact analysis model control variable specifications—FPHNY

Variable name	Specification
Intervention period	Categorical variable indicating time period of assessment. Categories include: at enrollment (reference); six months post-enrollment
Treatment indicator	Categorical variable indicating treatment status. Categories include: control group member (reference); participant
Entry indicator	Categorical variable indicating first type of service used at entry. Categories include use of CRC and NA-MCT
Interaction between intervention period and treatment	Interaction between intervention period and treatment indicator variables.
Interaction between intervention period and entry	Interaction between intervention period and entry indicator variables
Interaction between treatment and entry	Interaction between treatment indicator and entry indicator variables
Interaction between intervention period, treatment, and entry	Interaction between intervention period, treatment indicator, and entry indicator variables
Intervention period begin date	Categorical variable for member's first observed month and year in a given intervention period. Dates span between January 2010 (reference – three years prior to earliest enrollment date) and June 2015
Age	Continuous variable indicating age on the begin date of the intervention period
Age squared	Continuous variable measuring age as defined above squared
Sex	Categorical variable of member's sex. Categories include: female (reference); male
Disabled	Categorical variable indicating whether member was eligible for Medicaid based on disability
Race	Categorical variable of member's race. Categories include: White or other (reference), African American, or Hispanic
CDPS score	Continuous variable measuring member's CDPS score
AIDS, high	Categorical variable indicating whether member had AIDS, pneumocystis pneumonia, cryptococcosis, or Kaposi's sarcoma
Cardiovascular, extra low	Categorical variable indicating whether member had hypertension
Cardiovascular, low	Categorical variable indicating whether member had endocardial disease, myocardial infarction, angina, coronary atherosclerosis, or dysrhythmias
Nervous system, low	Categorical variable indicating whether member had epilepsy, Parkinson's disease, cerebral palsy, migraine, or cerebral degeneration
Genital, extra low	Categorical variable indicating whether member had uterine and pelvic inflammatory disease, endometriosis, or hyperplasia of prostate
Gastrointestinal, low	Categorical variable indicating whether member had ulcer, hernia, GI hemorrhage, intestinal infectious disease, or intestinal obstruction
Folate deficiency	Categorical variable indicating whether member had folate deficiency
Psychiatric, high	Categorical variable indicating whether member had schizophrenia
Psychiatric, medium	Categorical variable indicating whether member had bipolar affective disorder
Psychiatric, low	Categorical variable indicating whether member had other depression, panic disorder, or phobic disorder
Pulmonary, low	Categorical variable indicating whether member had viral pneumonias, chronic bronchitis, asthma, COPD, or emphysema
Skeletal, low	Categorical variable indicating whether member had rheumatoid arthritis, osteomyelitis, systemic lupus, or traumatic amputation of foot or leg
Substance abuse, low	Categorical variable indicating whether member had opioid, barbiturate, cocaine, or amphetamine abuse or dependence, or drug psychoses

Variable name	Specification
ICD-9 diagnosis category	Hierarchical categorical variable indicating most "important" diagnosis category, based on ICD-9 diagnosis codes, in the following order: (1) schizophrenic disorders; (2) bipolar disorders; (3) drug or alcohol-induced psychosis; (4) delusional disorder and other nonorganic psychosis; (5) suicide; (6) depressive, episodic mood disorder, or other depressive disorder; (7) anxiety, dissociative, somatoform, and adjustment reaction; (8) other behavioral health-related diagnoses
Adjustment reaction	Categorical variable indicating whether member had ICD-9 diagnosis code of adjustment reaction
Anxiety	Categorical variable indicating whether member had ICD-9 diagnosis code of anxiety, dissociative, and somatoform
Bipolar	Categorical variable indicating whether member had ICD-9 diagnosis code of bipolar disorders
Delusional disorder	Categorical variable indicating whether member had ICD-9 diagnosis code of delusional disorder and other nonorganic psychosis
Dementia	Categorical variable indicating whether member had ICD-9 diagnosis code of dementia
Depressive	Categorical variable indicating whether member had ICD-9 diagnosis code of episodic mood disorder, depressive
Drug or alcohol psychosis	Categorical variable indicating whether member had ICD-9 diagnosis code of drug or alcohol psychosis
Drug and alcohol	Categorical variable indicating whether member had ICD-9 diagnosis code of drug or alcohol-related disorders
Other depressive	Categorical variable indicating whether member had ICD-9 diagnosis code of other depressive disorder
Other psychosis	Categorical variable indicating whether member had ICD-9 diagnosis code of other psychoses not listed in other categories
Other nonpsychotic mental disorders	Categorical variable indicating whether member had ICD-9 diagnosis code of other nonpsychotic mental disorders listed in other categories
Persistent mental disorders	Categorical variable indicating whether member had ICD-9 diagnosis code of persistent mental disorders due to conditions classified elsewhere
Schizophrenia	Categorical variable indicating whether member had ICD-9 diagnosis code of schizophrenic disorders
Suicide	Categorical variable indicating whether member had ICD-9 diagnosis code related to suicide

IV. HEALTHLINKNOW

A. Introduction

The goal of this quantitative evaluation was to determine whether the HLN telepsychiatry program had a measurable impact on health care utilization or expenditures. We applied a difference-in-differences regression model to estimate program impacts on Medicare beneficiaries. Below, we describe the methods for conducting this analysis. We first describe the data sources (Section B) and approach to selecting the analytic intervention population (Section C). Then, we describe the steps taken to select the matched comparison group (Section D). Finally, we specify how outcome and control variables were constructed (Section E).

B. Description of data sources

We used two sources of data for our impact analyses: program enrollment data obtained from HLN and Medicare claims and enrollment data extracted from the CMS Virtual Research Data Center (VRDC).

Program enrollment data. To support this analysis, HLN provided Mathematica with data files for program participants containing Medicare identifiers, demographic characteristics (date of birth, gender, disability status, insurance coverage), and start and end dates of services for participants who used services between March 1, 2013 and June 30, 2015.

Medicare administrative data. For participants and a comparison pool we extracted final action claims with dates of service from January 2010 through June 2015.[91] Mathematica extracted data for program participants by linking the Medicare identifier provided in HLN's enrollment data to the BENE_ID in the VRDC cross-reference files. Comparison group members were selected based on service use, as described in Section D. We extracted standard analytic base and revenue-center/line-item claims datasets for the following claim types: carrier, durable medical equipment, home health, hospice, inpatient, outpatient, and skilled nursing facility. To obtain information on beneficiary Medicare enrollment spans we used the Master Beneficiary Summary File (MBSF). The MBSF includes information on date of birth, gender, most recent county of residence, Medicare Advantage enrollment, and third party insurance coverage.

C. Identification of intervention group

Our analysis was limited to program participants enrolled in Medicare-only or dually enrolled in Medicaid and Medicare. For each participant we analyzed use of the four service types identified in Table A-IV.1, any of which might have led to program enrollment. We identified use of these services in the program enrollment month and the 11 months prior. The last column in Table A-IV.1 indicates the claim date used to assign each service to a month. In addition, we identified whether each service was associated with a behavioral health and selected somatic diagnoses (Table A-IV.2). To broaden the base of conditions that could have triggered referral to telepsychiatry services, we tracked select somatic conditions. These conditions were added to a set of diagnosis codes potentially leading to program enrollment because we noted a

[91] Mathematica extracted data for program participants on December 21, 2015. Data for the comparison population were extracted on April 17, 2016.

number of intervention group members with these somatic diagnoses and no behavioral health diagnoses in the pre-period data. We limited inclusion of somatic conditions to those potentially co-occurring with a behavioral health condition (diagnosed or undiagnosed) or those not directly attributable to a behavioral health condition but that may be associated with an underlying physical or mental health problem.

Table A-IV.1. Service use leading to program enrollment

Variable name	Specification	Assign to month based on
Office visit	CLM_TYPE = 40 (Outpatient) and [CPT Code = 99201-99205, 99211-99215, 99304-99310, 99315-99316, 99318, 99324-99328, 99334-99350, 99381-99387, 99391-99397, 99401-99404, 99406-99409, 99411-99412, 99420, 99429, 99432, 99461 or ICD-9 = V20.2, V20.3, V70.0, V70.3, V70.5, V70.6, V70.8, V70.9 or HPCPS = G0438, G0439]	FROM_DATE
ED visit	CLM_TYPE = 40 (Outpatient and revenue center codes = (450, 451, 452, 456, 459, 760, 762, 981 or CPT code = G0378) where at least one line item has a CPT and BETOS code that is not a laboratory or imaging (CPT=70000-79999 or 80000-89999 or BETOS T1-T2, I1-I4)	FROM_DATE
Inpatient stay	CLM_TYPE = 60 (Inpatient claim)	ADMSN_DT
SNF service	CLM_TYPE = 20 or 30 (SNF claim)	Any day from FROM_DT through THRU_DT

Table A-IV.2. ICD-9 mental health and somatic diagnosis codes associated with services included in the analytic file

Diagnosis codes	Label
295.00 to 295.95	Schizophrenia spectrum disorders
296.00 to 296.06, 296.40 to 296.80, 296.89, 296.10 to 296.16, 296.81	Bipolar disorders
296.20 to 296.36, 296.82, 300.4, 311, 311.0 296.90, 296.99, 293.83, 300.9	Depressive disorders Other mood disorders
305.1, 291.0-292.9, 303.0-303.9, 305.0-305.3, 292.0-292.9, 304.0-304.9, 305.2-305.9	Substance use disorders (alcohol, tobacco, and other drug use)
300.00 to 300.11, 300.20 to 300.3, 309.81	Anxiety disorders
290.0-290.9, 294.1x	Dementia
297.0 to 298.9, V62.84, V62.85, E950, E951, E952, E953, E954, E955, E956, E957, E958, E959, 300.12 to 300.15, 300.6, 300.7 to 300.89, 301.0 to 301.9, 307.40 to 307.49, 312.0 to 312.23, 312.4 to 312.89, 313.81, 312.30 to 312.39, 302.0 to 302.9, 299.00 to 299.91, 307.1, 307.5, 307.51, 314.00 to 314.01, 307.20 to 307.3, 313.0 to 313.3, 313.82 to 316, 648.4, V65.2, V71.09, 780.09, V15.41, V15.42, V15.81, V17.0, V60.0, V62.29, V62.4, V62.81, V62.89 All other codes in the range of 290.0-299.91 and 300.00-316 (not specified above)	Other BH conditions not specified above (other psychotic disorders, suicidal or homicidal ideation, injury from suicide, dissociative disorders, somatoform disorders, personality disorders, sleep disorders, disruptive behavior disorders, impulse control behavior, sexual and gender identity disorders, autism spectrum disorder, eating disorders, ADHD, other disorders diagnosed in childhood, mental disorders in pregnancy, person feigning illness, observation for other suspected mental condition, other alteration of consciousness, social/contextual circumstances [violence])
338.0-338.4, 729.1, 719.4x	Pain, chronic pain, myalgia
780.7	Fatigue
784.0, 346.xx	Headache, migraine
780.52, 780.54	Insomnia
780.1, 780.93	Hallucinations, memory loss

HLN's program enrollment data included 266 participants with Medicare identifiers. We excluded 34 individuals from the analytic sample for the following reasons:

- **Lack of Medicare data.** Three participants did not have a MBSF record in their program enrollment year. These data were necessary to obtain demographic characteristics.[92]
- **Insufficient Medicare enrollment**. Twenty-six participants were not continuously enrolled with FFS Medicare as their primary payer for six months around their program enrollment month (three months before their enrollment month, the month of enrollment, two months after enrollment). This level of FFS Medicare enrollment is a minimum required to measure the individual's pre- and post-enrollment Medicare utilization and expenditures.
- **Place of residence**. Three participants did not reside in Washington, Montana, or Wyoming or in three border counties at the time of program enrollment. Individuals outside these areas were not residing in the program service area.
- **Service use with behavioral health diagnosis.** Two participants did not have a claim for one of the four service types with a behavioral health or selected somatic diagnosis. Without information on relevant diagnoses, finding well-matched comparison beneficiaries would be difficult.

Our final pre-matching intervention group contained 232 participants.

D. Identification of comparison group

We used matching techniques to identify a comparison group. Propensity score matching and related matching methods are designed to create a comparison group that is similar in observable characteristics to the treatment group (Rosenbaum and Rubin 1983; Stuart 2010). Limiting the comparison group to a matched subsample may also reduce differences between intervention and comparison group members in terms of unobserved characteristics if those characteristics are correlated with matching variables. In this section we describe the steps taken to select the comparison group.

Step 1: Identified pool of potential comparison group beneficiaries. The comparison group for this analysis was selected from Medicare enrollees who, between March 2013 and June 2015, had a claim for least one of the four service types hypothesized to lead to enrollment (Table A-IV.1) associated with a targeted behavioral health or somatic diagnosis (Table A-IV.2) and resided in the following 10 neighbor states: North Dakota, South Dakota, Idaho, Oregon, Nevada, Utah, Colorado, Nebraska, Arizona, and New Mexico. We chose a comparison group outside of the HLN intervention states because HLN tried to enroll all eligible providers within the intervention states. As such, we were concerned that any providers within the intervention states who chose not to enroll would be systematically different from those who enrolled. For example, providers who chose to enroll might have less access to psychiatric services and might have less personal training in psychiatric illness relative to those who chose not to enroll.

[92] Lack of MBSF data suggests that these individuals were not enrolled in Medicare during the year associated with their enrollment month.

For each potential comparison group member, we needed to create a pseudo-enrollment month in March 2013 through June 2015 that reflects the month when the member likely would have enrolled in the intervention if they had been a participant. The pseudo-enrollment month allows us to define the pre-period and the post-period timeframe, similar to pre- and post-periods for intervention participants. For each person in the comparison pool, we identified all the months in which they had a claim for one of the four service types hypothesized to lead to enrollment with an associated behavioral health or selected somatic diagnoses and then we randomly selected one of these months as their pseudo-enrollment month. To ensure that the distribution of pseudo-enrollment months for the comparison population would be similar to the distribution of enrollment months among participants, we weighted the probability of selecting a given calendar month based on the proportion of participants who enrolled in that month relative to the proportion of the comparison population with an eligible service in that month.

After a pseudo-enrollment month was selected for each comparison group member, we excluded individuals who were not continuously enrolled with FFS Medicare as their primary payer for six months around their pseudo-enrollment month (three months before their enrollment month, the month of enrollment, two months after enrollment). Our comparison pool included 1,475,916 beneficiaries.

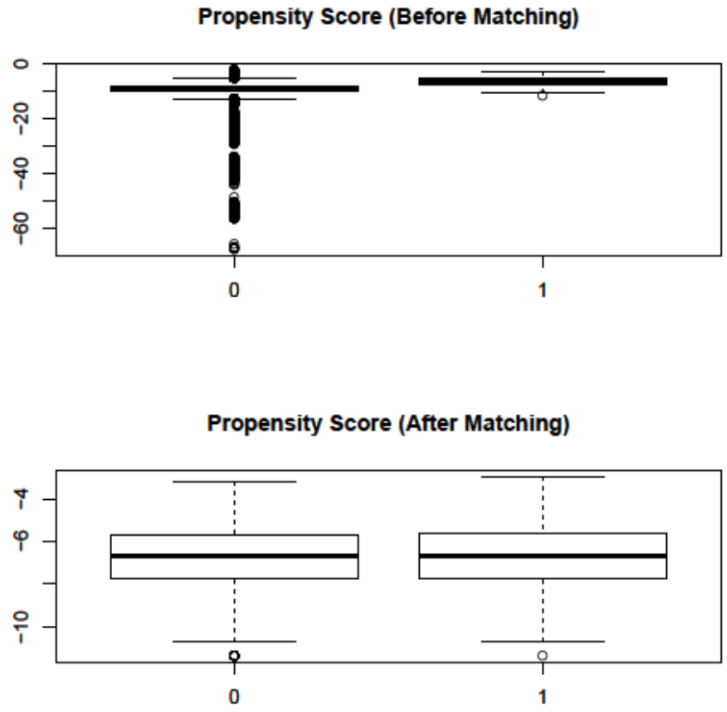
Step 2: Matched intervention participants at the individual level. Next, we used metrics of individual-level characteristics identified based on pre-period Medicare administrative data to select a subset of comparison pool members who were as similar as possible to the intervention group beneficiaries. Using propensity score matching, we matched up to 20 comparison group beneficiaries to each treatment group beneficiary. We used a nearest neighbor matching approach to select the closest comparison beneficiaries for each treatment group member, minimizing differences between treatment and comparison beneficiaries on observed characteristics during the pre-period. The matching algorithm first exact matched on county urbanicity, Medicare disability status, behavioral health diagnosis, and service use associated with either a behavioral health or somatic condition diagnosis at program entry.

We then fit a propensity score model including the following variables: enrollment date, age, gender, race, ethnicity, pre-period FFS service claims availability, 12-month service utilization (psychiatric visits, hospitalizations, ED visits, primary care visits, nursing home utilization), total Medicare spending, county-level unemployment, county-level poverty, Medicaid and Medicare dual enrollment status at study enrollment month, CMS Hierarchical Condition Category (HCC) codes, CMS Chronic Conditions Data Warehouse conditions not represented in HCC, and selected somatic diagnoses (pain, fatigue, headache, insomnia, hallucinations). The target matching ratio of treatment to comparison beneficiaries varied based on the number of comparison group members available within the exact matching strata. Up to 20 comparison beneficiaries could be selected for each participant.

Step 3: Assessed quality of the match. We employed several diagnostic tests to assess the quality of the matches. First, we graphically compared the propensity score distributions for all treatment and comparison beneficiaries prior to matching, checking for overlap in the propensity scores for the treatment and comparison groups. The propensity scores for both participants and comparison pool members prior to matching were concentrated near -10 with some overlap for the scores of all participants (Figure A-IV.1) suggesting we would be able to find matches for all

participants.[93] The post-matching distribution of propensity scores is similar for the two groups which signals that the groups are similar on observed characteristics.

Figure A-IV.1. Propensity score distributions for intervention and potential comparison groups (HLN)



Note: Figure presents boxplots created using the estimated propensity scores for the comparison (0) and intervention (1) groups, the left and right panels respectively. The width of the boxplots corresponds to the amount of data that contributed to the plots. Propensity scores are presented on the logit scale to better visualize the range of scores.

After matching, we examined the number of comparison beneficiaries matched to each intervention beneficiary (Table A-IV.3). A large number of 1:1 matches could indicate that the matching was problematic. However, we observed that the majority of the matched sets are of a higher ratio (e.g., 1:20), suggesting that the observable characteristics of many comparison pool members were similar to those of the intervention participants. In the regression analysis, each comparison group member was weighted by the inverse of the number of comparison group members matched to the same intervention group member. We were able to find matches for all 232 intervention beneficiaries.

[93] Note propensity scores are presented on the logit scale to better visualize the range of scores.

Table A-IV.3. Frequency table of ratio of intervention beneficiaries to comparison beneficiaries for each matched set (HLN)

Ratio of intervention to comparison beneficiaries	1:1	1:2	1:5	1:10	1:20
Number of matched sets	2	5	3	11	213

Note: In the second row, each cell indicates the number of intervention beneficiaries matched to the number of comparison beneficiaries indicated for that column. In this example, most of the intervention beneficiaries (213 out of 232) were matched to 20 comparison beneficiaries.

We also examined the overall covariate balance between intervention and comparison beneficiaries with a balance plot (Figure A-IV.2). A balance plot is a way to compare visually how close intervention and comparison beneficiaries are for each variable of interest. We compared the absolute and standardized differences between treatment and comparison groups for each matching variable before and after matching (Table A-IV.4). The standardized difference measures the difference in means in units of the pooled standard deviation. The standardized difference measure is advantageous in that it allows us to compare all variables on the same scale. We compared the standardized differences using plots to visually inspect whether we obtained good balance for each of the matching variables. Most matching variables met the condition of having less than .10 standardized difference between intervention and comparison beneficiaries. However, we were not able to obtain ideal balance on two variables: county-level unemployment and primary care service utilization during the pre-period. Both of these conditions were statistically significantly higher among the intervention group relative to the comparison group.

Figure A-IV.2. Balance plot comparing the standardized difference between intervention and comparison groups for each matching variable before and after matching (HLN)

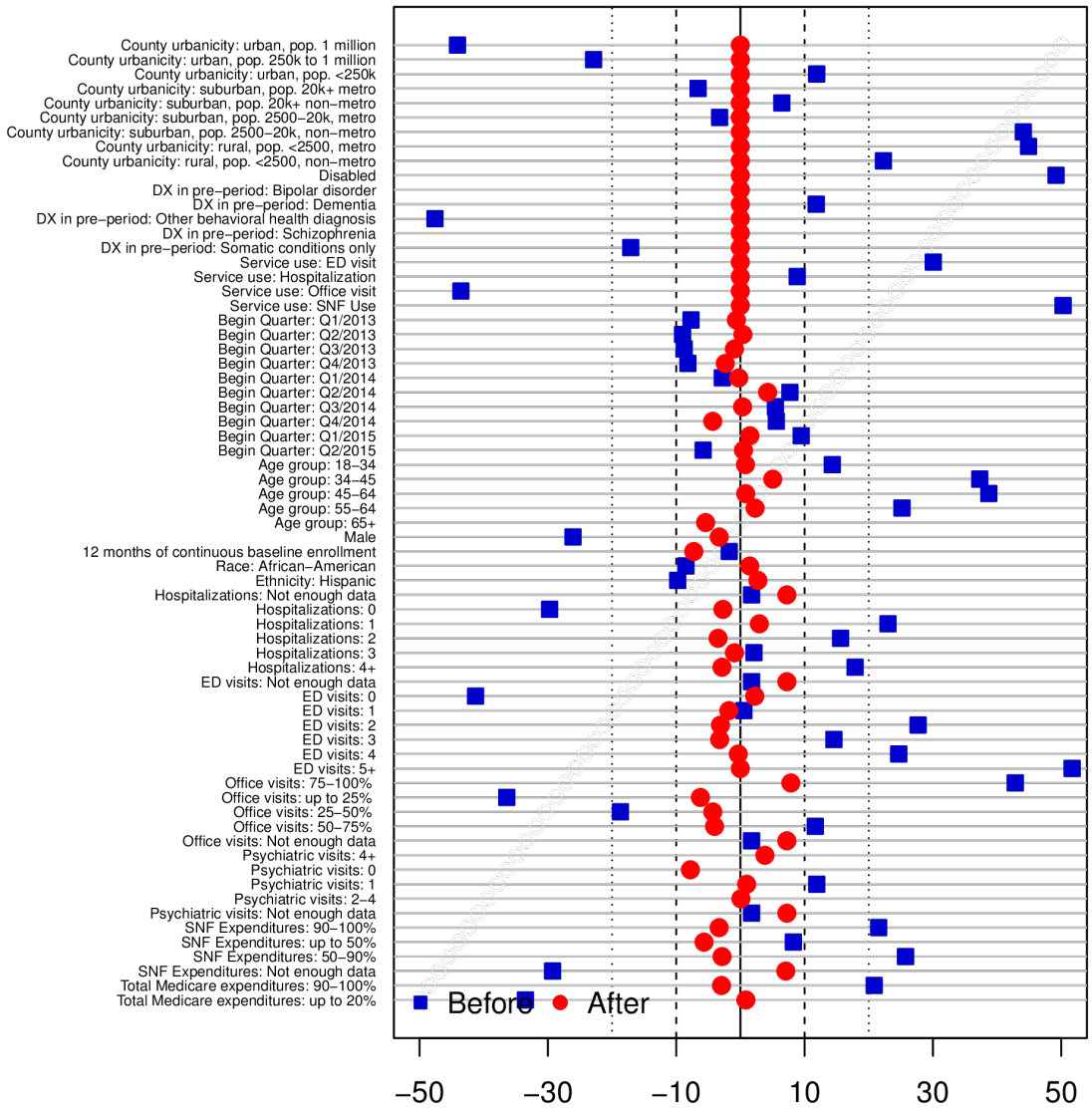


Figure A-IV.2 *(continued)*

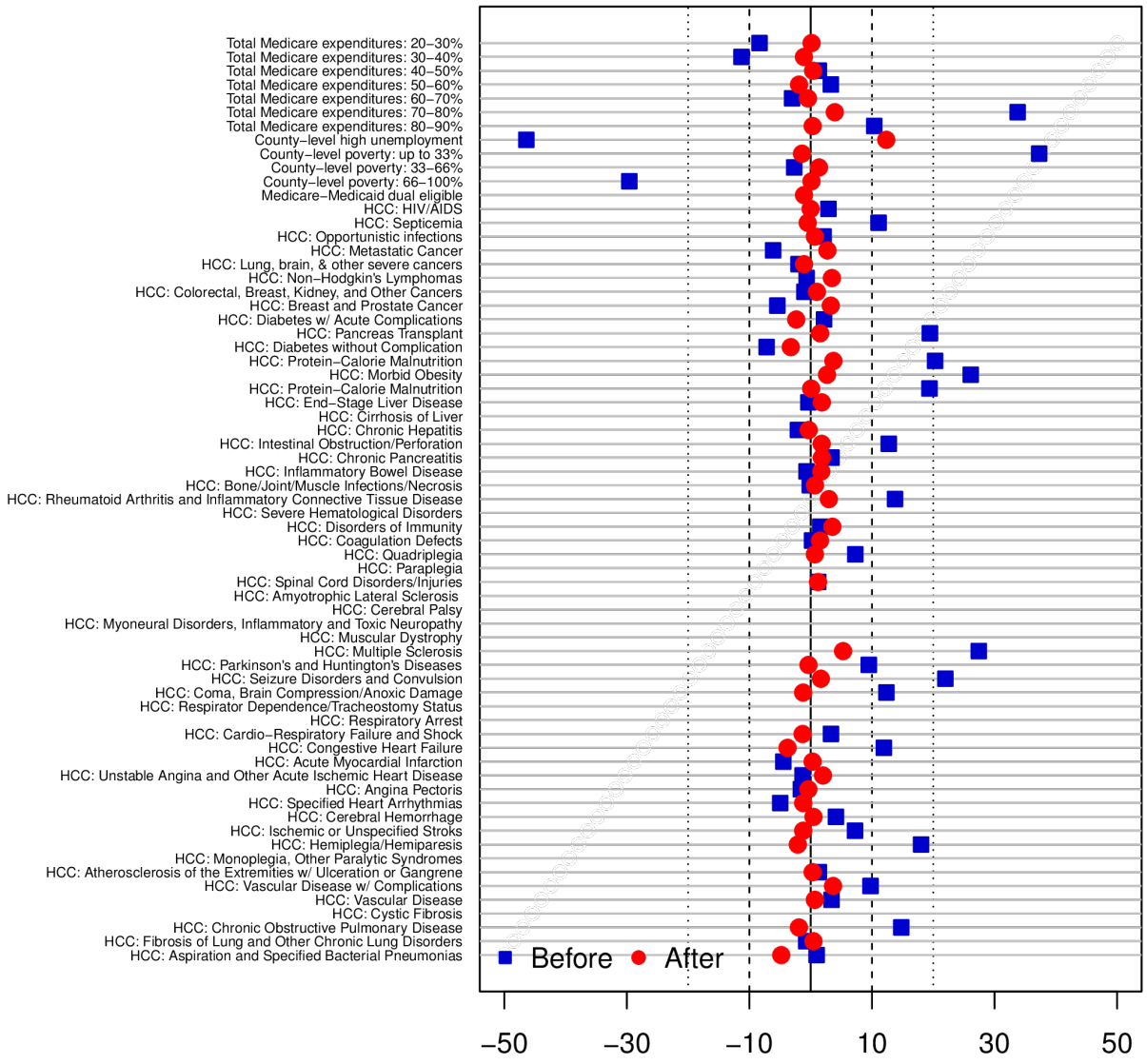
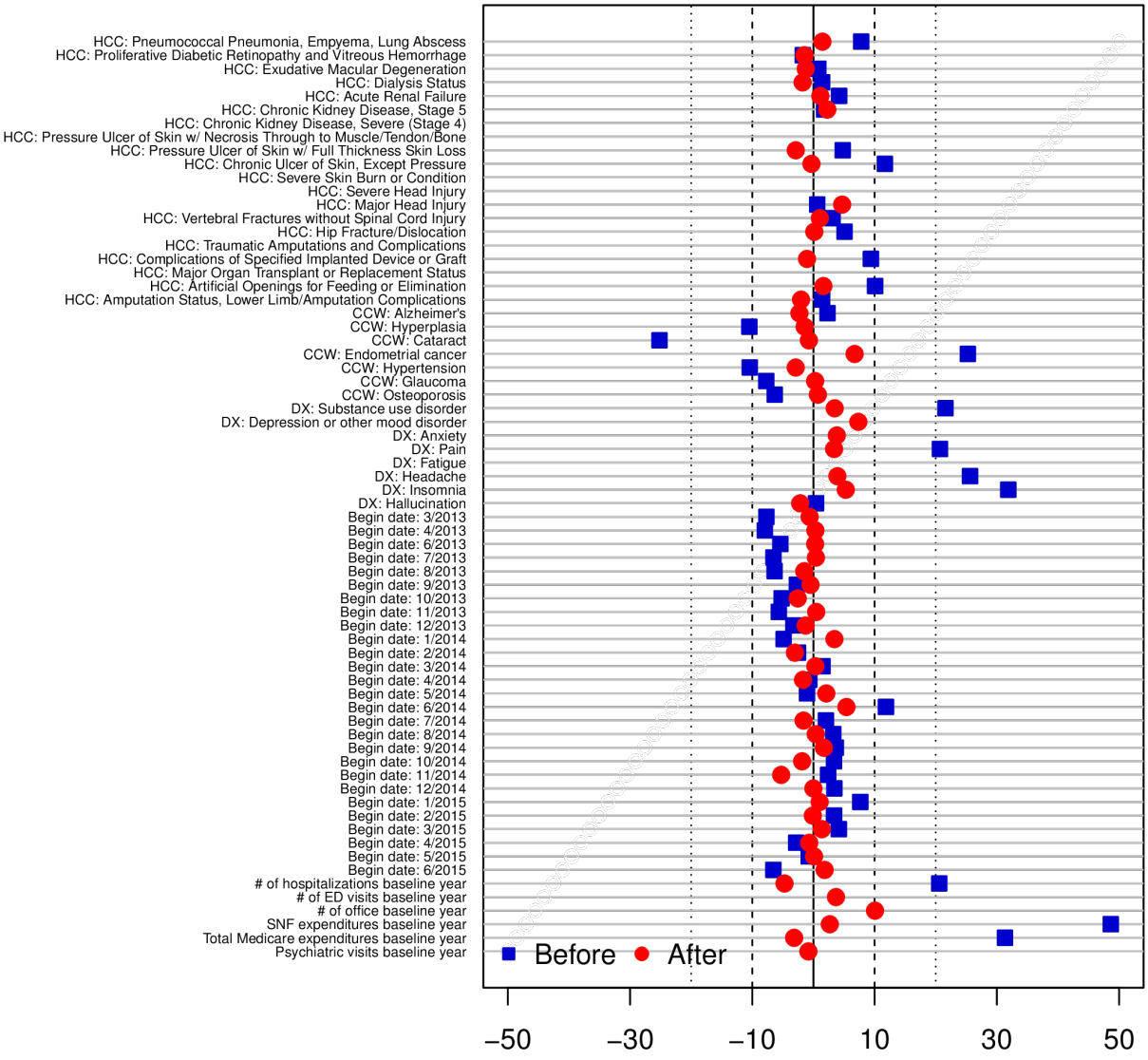


Figure A-IV.2 *(continued)*



Overall, we obtained an analytic sample well matched on medical conditions and core measures before the intervention. We controlled for baseline differences in unemployment and primary care visits in the regression models.

Table A-IV.4. Balance table before and after matching (HLN)

	Before matching					After matching				
	Comparison	Intervention	adj.diff	std.diff	*p*-value	Comparison	Intervention	adj.diff	std.diff	*p*-value
County urbanicity: urban, pop. 1 million	0.29	0.08	-0.21	-0.46	0.00	0.08	0.08	0.00	0.00	1.00
County urbanicity: urban, pop. 250k to 1 million	0.11	0.02	-0.09	-0.29	0.00	0.02	0.02	0.00	0.00	1.00
County urbanicity: urban, pop. <250k	0.26	0.31	0.05	0.11	0.10	0.31	0.31	0.00	0.00	1.00
County urbanicity: suburban, pop. 20k+ metro	0.07	0.08	0.01	0.03	0.68	0.08	0.08	0.00	0.00	1.00
County urbanicity: suburban, pop. 20k+ non-metro	0.10	0.13	0.03	0.12	0.08	0.13	0.13	0.00	0.00	1.00
County urbanicity: suburban, pop. 2500-20k, metro	0.02	0.02	0.00	-0.01	0.82	0.02	0.02	0.00	0.00	1.00
County urbanicity: suburban, pop. 2500-20k, non-metro	0.09	0.23	0.13	0.46	0.00	0.23	0.23	0.00	0.00	1.00
County urbanicity: rural, pop. <2500, metro	0.01	0.06	0.05	0.43	0.00	0.06	0.06	0.00	0.00	1.00
County urbanicity: rural, pop. <2500, non-metro	0.03	0.07	0.03	0.19	0.00	0.07	0.07	0.00	0.00	1.00
Disabled	0.25	0.51	0.26	0.59	0.00	0.51	0.51	0.00	0.00	1.00
DX in pre-period: Bipolar disorder	0.03	0.17	0.15	0.89	0.00	0.17	0.17	0.00	0.00	1.00
DX in pre-period: Dementia	0.02	0.05	0.02	0.15	0.02	0.05	0.05	0.00	0.00	1.00
DX in pre-period: Other behavioral health diagnosis	0.90	0.68	-0.22	-0.75	0.00	0.68	0.68	0.00	0.00	1.00
DX in pre-period: Schizophrenia	0.01	0.07	0.06	0.69	0.00	0.07	0.07	0.00	0.00	1.00
DX in pre-period: Somatic conditions only	0.04	0.03	-0.01	-0.06	0.40	0.03	0.03	0.00	0.00	1.00
Service use: ED visit	0.06	0.14	0.09	0.38	0.00	0.14	0.14	0.00	0.00	1.00
Service use: Hospitalization	0.03	0.08	0.05	0.29	0.00	0.08	0.08	0.00	0.00	1.00
Service use: Office visit	0.91	0.72	-0.19	-0.67	0.00	0.72	0.72	0.00	0.00	1.00
Service use: SNF Use	0.01	0.06	0.06	0.69	0.00	0.06	0.06	0.00	0.00	1.00
Begin Quarter: Q1/2013	0.01	0.00	0.00	-0.03	0.63	0.00	0.00	0.00	-0.01	0.92
Begin Quarter: Q2/2013	0.02	0.03	0.01	0.04	0.59	0.03	0.03	0.00	0.00	0.95
Begin Quarter: Q3/2013	0.04	0.06	0.02	0.12	0.07	0.06	0.06	0.00	-0.01	0.89
Begin Quarter: Q4/2013	0.03	0.05	0.02	0.11	0.08	0.06	0.05	-0.01	-0.02	0.74
Begin Quarter: Q1/2014	0.11	0.08	-0.03	-0.10	0.15	0.08	0.08	0.00	0.00	0.97
Begin Quarter: Q2/2014	0.17	0.16	-0.01	-0.02	0.78	0.14	0.16	0.02	0.04	0.53
Begin Quarter: Q3/2014	0.17	0.16	-0.02	-0.04	0.54	0.15	0.16	0.00	0.00	0.96
Begin Quarter: Q4/2014	0.17	0.16	-0.01	-0.04	0.57	0.17	0.16	-0.02	-0.04	0.53
Begin Quarter: Q1/2015	0.17	0.21	0.04	0.10	0.11	0.20	0.21	0.01	0.01	0.82
Begin Quarter: Q2/2015	0.12	0.10	-0.02	-0.06	0.39	0.10	0.10	0.00	0.01	0.94
Age group: 18-34	0.02	0.04	0.02	0.18	0.01	0.04	0.04	0.00	0.01	0.90
Age group: 34-45	0.03	0.09	0.07	0.43	0.00	0.08	0.09	0.01	0.05	0.43
Age group: 45-64	0.05	0.15	0.10	0.45	0.00	0.15	0.15	0.00	0.01	0.89
Age group: 55-64	0.08	0.17	0.08	0.30	0.00	0.16	0.17	0.01	0.02	0.71
Age group: 65+	0.82	0.55	-0.27	-0.72	0.00	0.57	0.55	-0.03	-0.05	0.12
Male	0.37	0.25	-0.12	-0.25	0.00	0.26	0.25	-0.01	-0.03	0.62
12 months of continuous baseline enrollment	0.93	0.94	0.01	0.04	0.56	0.95	0.94	-0.02	-0.07	0.28

	Before matching					After matching				
	Comparison	Intervention	adj.diff	std.diff	*p*-value	Comparison	Intervention	adj.diff	std.diff	*p*-value
Race: African American	0.02	0.00	-0.01	-0.09	0.17	0.00	0.00	0.00	0.01	0.83
Ethnicity: Hispanic	0.02	0.00	-0.01	-0.10	0.12	0.00	0.00	0.00	0.03	0.67
Hospitalizations: Not enough data	0.07	0.06	-0.01	-0.04	0.56	0.05	0.06	0.02	0.07	0.28
Hospitalizations: 0	0.73	0.63	-0.10	-0.23	0.00	0.64	0.63	-0.01	-0.03	0.56
Hospitalizations: 1	0.14	0.21	0.07	0.19	0.00	0.20	0.21	0.01	0.03	0.63
Hospitalizations: 2	0.04	0.06	0.03	0.15	0.02	0.07	0.06	-0.01	-0.03	0.60
Hospitalizations: 3	0.01	0.01	0.00	0.02	0.72	0.01	0.01	0.00	-0.01	0.89
Hospitalizations: 4+	0.01	0.02	0.02	0.20	0.00	0.03	0.02	0.00	-0.03	0.66
ED visits: Not enough data	0.07	0.06	-0.01	-0.04	0.56	0.05	0.06	0.02	0.07	0.28
ED visits: 0	0.59	0.39	-0.20	-0.41	0.00	0.38	0.39	0.01	0.02	0.70
ED visits: 1	0.20	0.21	0.01	0.03	0.68	0.22	0.21	-0.01	-0.02	0.79
ED visits: 2	0.07	0.15	0.08	0.30	0.00	0.16	0.15	-0.01	-0.03	0.65
ED visits: 3	0.03	0.06	0.03	0.16	0.02	0.06	0.06	-0.01	-0.03	0.65
ED visits: 4	0.01	0.04	0.03	0.26	0.00	0.04	0.04	0.00	0.00	0.96
ED visits: 5+	0.02	0.09	0.07	0.54	0.00	0.09	0.09	0.00	0.00	1.00
Office visits: 75-100%	0.23	0.40	0.17	0.41	0.00	0.36	0.40	0.04	0.08	0.23
Office visits: up to 25%	0.26	0.10	-0.15	-0.35	0.00	0.12	0.10	-0.02	-0.06	0.34
Office visits: 25-50%	0.19	0.12	-0.07	-0.18	0.01	0.13	0.12	-0.01	-0.04	0.52
Office visits: 50-75%	0.25	0.31	0.06	0.14	0.04	0.33	0.31	-0.02	-0.04	0.55
Office visits: Not enough data	0.07	0.06	-0.01	-0.04	0.56	0.05	0.06	0.02	0.07	0.28
Psychiatric visits: 4+	0.06	0.24	0.18	0.77	0.00	0.23	0.24	0.02	0.04	0.54
Psychiatric visits: 0	0.65	0.27	-0.39	-0.81	0.00	0.30	0.27	-0.04	-0.08	0.16
Psychiatric visits: 1	0.13	0.16	0.03	0.10	0.13	0.16	0.16	0.00	0.01	0.88
Psychiatric visits: 2-4	0.08	0.26	0.18	0.65	0.00	0.26	0.26	0.00	0.00	0.99
Psychiatric visits: Not enough data	0.07	0.06	-0.01	-0.04	0.56	0.05	0.06	0.02	0.07	0.28
SNF Expenditures: 90-100%	0.01	0.02	0.02	0.20	0.00	0.03	0.02	-0.01	-0.03	0.63
SNF Expenditures: up to 50%	0.03	0.04	0.02	0.09	0.16	0.06	0.04	-0.01	-0.06	0.37
SNF Expenditures: 50-90%	0.02	0.06	0.04	0.23	0.00	0.07	0.06	-0.01	-0.03	0.64
SNF Expenditures: Not enough data	0.94	0.88	-0.07	-0.29	0.00	0.85	0.88	0.03	0.07	0.14
Total Medicare expenditures: 90-100%	0.08	0.13	0.05	0.19	0.00	0.14	0.13	-0.01	-0.03	0.62
Total Medicare expenditures: up to 20%	0.21	0.06	-0.14	-0.35	0.00	0.06	0.06	0.00	0.01	0.89
Total Medicare expenditures: 20-30%	0.10	0.08	-0.02	-0.08	0.24	0.08	0.08	0.00	0.00	0.98
Total Medicare expenditures: 30-40%	0.10	0.07	-0.03	-0.11	0.10	0.07	0.07	0.00	-0.01	0.87
Total Medicare expenditures: 40-50%	0.10	0.11	0.01	0.03	0.68	0.11	0.11	0.00	0.00	0.95
Total Medicare expenditures: 50-60%	0.11	0.12	0.01	0.05	0.47	0.13	0.12	-0.01	-0.02	0.78
Total Medicare expenditures: 60-70%	0.11	0.10	-0.01	-0.02	0.74	0.10	0.10	0.00	0.00	0.95
Total Medicare expenditures: 70-80%	0.10	0.21	0.11	0.35	0.00	0.19	0.21	0.02	0.04	0.55
Total Medicare expenditures: 80-90%	0.09	0.12	0.03	0.09	0.18	0.12	0.12	0.00	0.00	0.96

	Before matching					After matching				
	Comparison	Intervention	adj.diff	std.diff	*p*-value	Comparison	Intervention	adj.diff	std.diff	*p*-value
County-level high unemployment	0.66	0.43	-0.23	-0.48	0.00	0.37	0.43	0.06	0.12	0.04
County-level poverty: up to 33%	0.23	0.40	0.17	0.40	0.00	0.41	0.40	-0.01	-0.01	0.83
County-level poverty: 33-66%	0.32	0.30	-0.02	-0.05	0.48	0.30	0.30	0.01	0.01	0.84
County-level poverty: 66-100%	0.44	0.30	-0.15	-0.30	0.00	0.30	0.30	0.00	0.00	0.98
Medicare-Medicaid dual eligible	0.20	0.48	0.28	0.70	0.00	0.49	0.48	-0.01	-0.01	0.85
HCC: HIV/AIDS	0.00	0.00	0.00	0.04	0.55	0.00	0.00	0.00	0.00	1.00
HCC: Septicemia	0.02	0.03	0.01	0.08	0.20	0.04	0.03	0.00	0.00	0.94
HCC: Opportunistic infections	0.00	0.00	0.00	0.02	0.80	0.00	0.00	0.00	0.01	0.92
HCC: Metastatic Cancer	0.01	0.00	-0.01	-0.06	0.34	0.00	0.00	0.00	0.03	0.67
HCC: Lung, brain, & other severe cancers	0.01	0.01	0.00	-0.02	0.75	0.01	0.01	0.00	-0.01	0.87
HCC: Non-Hodgkin's Lymphomas	0.01	0.01	0.00	-0.01	0.90	0.01	0.01	0.00	0.03	0.60
HCC: Colorectal, Breast, Kidney, and Other Cancers	0.02	0.02	0.00	-0.01	0.85	0.02	0.02	0.00	0.01	0.88
HCC: Breast and Prostate Cancer	0.06	0.04	-0.01	-0.06	0.35	0.04	0.04	0.01	0.03	0.63
HCC: Diabetes with Acute Complications	0.00	0.00	0.00	0.02	0.72	0.01	0.00	0.00	-0.02	0.74
HCC: Pancreas Transplant	0.09	0.15	0.06	0.20	0.00	0.14	0.15	0.01	0.02	0.81
HCC: Diabetes without Complication	0.15	0.12	-0.03	-0.07	0.28	0.13	0.12	-0.01	-0.03	0.63
HCC: Protein-Calorie Malnutrition	0.02	0.04	0.03	0.20	0.00	0.04	0.04	0.01	0.04	0.58
HCC: Morbid Obesity	0.04	0.09	0.05	0.26	0.00	0.08	0.09	0.01	0.03	0.69
HCC: Protein-Calorie Malnutrition	0.03	0.07	0.03	0.19	0.00	0.07	0.07	0.00	0.00	0.99
HCC: End-Stage Liver Disease	0.00	0.00	0.00	0.00	0.96	0.00	0.00	0.00	0.02	0.78
HCC: Cirrhosis of Liver	0.00	0.00	0.00	N/A	1.00	0.00	0.00	0.00	N/A	1.00
HCC: Chronic Hepatitis	0.01	0.00	0.00	-0.02	0.78	0.00	0.00	0.00	0.00	0.96
HCC: Intestinal Obstruction/Perforation	0.02	0.03	0.01	0.11	0.08	0.03	0.03	0.00	0.02	0.78
HCC: Chronic Pancreatitis	0.00	0.00	0.00	0.04	0.59	0.00	0.00	0.00	0.02	0.78
HCC: Inflammatory Bowel Disease	0.01	0.01	0.00	-0.01	0.86	0.01	0.01	0.00	0.02	0.79
HCC: Bone/Joint/Muscle Infections/Necrosis	0.01	0.01	0.00	-0.01	0.92	0.01	0.01	0.00	0.01	0.91
HCC: Rheumatoid Arthritis and Inflammatory Connective Tissue Disease	0.07	0.10	0.03	0.14	0.04	0.09	0.10	0.01	0.03	0.65
HCC: Severe Hematological Disorders	0.00	0.00	0.00	N/A	1.00	0.00	0.00	0.00	N/A	1.00
HCC: Disorders of Immunity	0.01	0.01	0.00	0.01	0.83	0.01	0.01	0.00	0.04	0.60
HCC: Coagulation Defects	0.05	0.04	0.00	-0.01	0.87	0.04	0.04	0.00	0.02	0.82
HCC: Quadriplegia	0.00	0.00	0.00	0.08	0.24	0.00	0.00	0.00	0.01	0.92
HCC: Paraplegia	0.00	0.00	0.00	N/A	1.00	0.00	0.00	0.00	N/A	1.00
HCC: Spinal Cord Disorders/Injuries	0.01	0.01	0.00	0.01	0.88	0.01	0.01	0.00	0.01	0.85
HCC: Amyotrophic Lateral Sclerosis	0.00	0.00	0.00	N/A	1.00	0.00	0.00	0.00	N/A	1.00
HCC: Cerebral Palsy	0.00	0.00	0.00	N/A	1.00	0.00	0.00	0.00	N/A	1.00
HCC: Myoneural Disorders, Inflammatory and Toxic Neuropathy	0.00	0.00	0.00	N/A	1.00	0.00	0.00	0.00	N/A	1.00

	Before matching					After matching				
	Comparison	Intervention	adj.diff	std.diff	*p*-value	Comparison	Intervention	adj.diff	std.diff	*p*-value
HCC: Muscular Dystrophy	0.00	0.00	0.00	N/A	1.00	0.00	0.00	0.00	N/A	1.00
HCC: Multiple Sclerosis	0.01	0.03	0.03	0.29	0.00	0.03	0.03	0.01	0.05	0.42
HCC: Parkinson's and Huntington's Diseases	0.02	0.03	0.01	0.09	0.17	0.04	0.03	0.00	0.00	0.96
HCC: Seizure Disorders and Convulsion	0.04	0.08	0.04	0.22	0.00	0.07	0.08	0.00	0.02	0.81
HCC: Coma, Brain Compression/Anoxic Damage	0.00	0.01	0.01	0.12	0.08	0.01	0.01	0.00	-0.01	0.87
HCC: Respirator Dependence/Tracheostomy Status	0.00	0.00	0.00	N/A	1.00	0.00	0.00	0.00	N/A	1.00
HCC: Respiratory Arrest	0.00	0.00	0.00	N/A	1.00	0.00	0.00	0.00	N/A	1.00
HCC: Cardio-Respiratory Failure and Shock	0.03	0.03	0.00	0.02	0.80	0.04	0.03	0.00	-0.01	0.85
HCC: Congestive Heart Failure	0.12	0.16	0.04	0.11	0.08	0.17	0.16	-0.01	-0.04	0.57
HCC: Acute Myocardial Infarction	0.01	0.00	-0.01	-0.06	0.39	0.00	0.00	0.00	0.00	0.97
HCC: Unstable Angina and Other Acute Ischemic Heart Disease	0.02	0.01	0.00	-0.02	0.73	0.01	0.01	0.00	0.02	0.76
HCC: Angina Pectoris	0.02	0.01	0.00	-0.02	0.78	0.01	0.01	0.00	0.00	0.96
HCC: Specified Heart Arrhythmias	0.13	0.11	-0.02	-0.06	0.33	0.12	0.11	0.00	-0.01	0.86
HCC: Cerebral Hemorrhage	0.01	0.01	0.00	0.03	0.66	0.01	0.01	0.00	0.00	0.94
HCC: Ischemic or Unspecified Strokes	0.04	0.05	0.01	0.06	0.36	0.05	0.05	0.00	-0.01	0.86
HCC: Hemiplegia/Hemiparesis	0.01	0.03	0.02	0.17	0.01	0.04	0.03	0.00	-0.02	0.76
HCC: Monoplegia, Other Paralytic Syndromes	0.00	0.00	0.00	N/A	1.00	0.00	0.00	0.00	N/A	1.00
HCC: Atherosclerosis of the Extremities with Ulceration or Gangrene	0.00	0.00	0.00	0.01	0.84	0.00	0.00	0.00	0.00	0.96
HCC: Vascular Disease with Complications	0.02	0.04	0.01	0.09	0.17	0.03	0.04	0.01	0.04	0.60
HCC: Vascular Disease	0.13	0.14	0.01	0.03	0.70	0.14	0.14	0.00	0.01	0.92
HCC: Cystic Fibrosis	0.00	0.00	0.00	N/A	1.00	0.00	0.00	0.00	N/A	1.00
HCC: Chronic Obstructive Pulmonary Disease	0.17	0.22	0.05	0.15	0.03	0.23	0.22	-0.01	-0.02	0.78
HCC: Fibrosis of Lung and Other Chronic Lung Disorders	0.01	0.01	0.00	-0.01	0.88	0.01	0.01	0.00	0.00	0.94
HCC: Aspiration and Specified Bacterial Pneumonias	0.01	0.01	0.00	0.00	0.98	0.01	0.01	-0.01	-0.05	0.52
HCC: Pneumococcal Pneumonia, Empyema, Lung Abscess	0.00	0.01	0.00	0.07	0.27	0.01	0.01	0.00	0.01	0.82
HCC: Proliferative Diabetic Retinopathy and Vitreous Hemorrhage	0.01	0.00	0.00	-0.02	0.79	0.01	0.00	0.00	-0.01	0.86
HCC: Exudative Macular Degeneration	0.02	0.02	0.00	0.01	0.91	0.02	0.02	0.00	-0.01	0.85
HCC: Dialysis Status	0.00	0.00	0.00	0.01	0.86	0.01	0.00	0.00	-0.02	0.80
HCC: Acute Renal Failure	0.04	0.04	0.00	0.02	0.80	0.04	0.04	0.00	0.01	0.87
HCC: Chronic Kidney Disease, Stage 5	0.00	0.00	0.00	0.02	0.75	0.00	0.00	0.00	0.02	0.73
HCC: Chronic Kidney Disease, Severe (Stage 4)	0.00	0.00	0.00	N/A	1.00	0.00	0.00	0.00	N/A	1.00
HCC: Pressure Ulcer of Skin with Necrosis Through to Muscle, Tendon, or Bone	0.00	0.00	0.00	N/A	1.00	0.00	0.00	0.00	N/A	1.00

	Before matching					After matching				
	Comparison	Intervention	adj.diff	std.diff	*p*-value	Comparison	Intervention	adj.diff	std.diff	*p*-value
HCC: Pressure Ulcer of Skin with Full Thickness Skin Loss	0.00	0.00	0.00	0.04	0.51	0.01	0.00	0.00	-0.03	0.71
HCC: Chronic Ulcer of Skin, Except Pressure	0.02	0.04	0.02	0.12	0.07	0.04	0.04	0.00	0.00	0.96
HCC: Severe Skin Burn or Condition	0.00	0.00	0.00	N/A	1.00	0.00	0.00	0.00	N/A	1.00
HCC: Severe Head Injury	0.00	0.00	0.00	N/A	1.00	0.00	0.00	0.00	N/A	1.00
HCC: Major Head Injury	0.01	0.01	0.00	0.00	0.98	0.01	0.01	0.00	0.05	0.51
HCC: Vertebral Fractures without Spinal Cord Injury	0.02	0.02	0.00	0.02	0.71	0.02	0.02	0.00	0.01	0.88
HCC: Hip Fracture/Dislocation	0.02	0.02	0.00	0.03	0.60	0.02	0.02	0.00	0.00	0.98
HCC: Traumatic Amputations and Complications	0.00	0.00	0.00	N/A	1.00	0.00	0.00	0.00	N/A	1.00
HCC: Complications of Specified Implanted Device or Graft	0.02	0.03	0.01	0.09	0.20	0.04	0.03	0.00	-0.01	0.88
HCC: Major Organ Transplant or Replacement Status	0.00	0.00	0.00	N/A	1.00	0.00	0.00	0.00	N/A	1.00
HCC: Artificial Openings for Feeding or Elimination	0.01	0.02	0.01	0.10	0.13	0.02	0.02	0.00	0.02	0.82
HCC: Amputation Status, Lower Limb/Amputation Complications	0.00	0.00	0.00	0.02	0.81	0.01	0.00	0.00	-0.02	0.77
CCW: Alzheimer's	0.05	0.06	0.00	0.02	0.74	0.06	0.06	-0.01	-0.02	0.71
CCW: Hyperplasia	0.07	0.04	-0.03	-0.11	0.08	0.04	0.04	0.00	-0.01	0.83
CCW: Cataract	0.18	0.07	-0.10	-0.27	0.00	0.08	0.07	0.00	-0.01	0.91
CCW: Endometrial cancer	0.00	0.02	0.02	0.25	0.00	0.01	0.02	0.01	0.07	0.33
CCW: Hypertension	0.58	0.51	-0.06	-0.13	0.05	0.53	0.51	-0.01	-0.03	0.66
CCW: Glaucoma	0.08	0.06	-0.02	-0.08	0.20	0.06	0.06	0.00	0.00	0.97
CCW: Osteoporosis	0.08	0.06	-0.02	-0.08	0.24	0.06	0.06	0.00	0.01	0.91
DX: Substance use disorder	0.25	0.35	0.10	0.22	0.00	0.33	0.35	0.02	0.03	0.57
DX: Depression or other mood disorder	0.45	0.80	0.35	0.71	0.00	0.77	0.80	0.03	0.07	0.18
DX: Anxiety	0.35	0.62	0.27	0.57	0.00	0.60	0.62	0.02	0.04	0.53
DX: Pain	0.50	0.62	0.11	0.23	0.00	0.60	0.62	0.02	0.03	0.61
DX: Fatigue	0.00	0.00	0.00	N/A	1.00	0.00	0.00	0.00	N/A	1.00
DX: Headache	0.14	0.24	0.09	0.26	0.00	0.22	0.24	0.02	0.04	0.55
DX: Insomnia	0.15	0.26	0.11	0.31	0.00	0.24	0.26	0.02	0.05	0.42
DX: Hallucination	0.07	0.07	0.00	-0.01	0.87	0.07	0.07	-0.01	-0.02	0.75
Begin date: 3/2013	0.01	0.00	0.00	-0.03	0.63	0.00	0.00	0.00	-0.01	0.92
Begin date: 4/2013	0.01	0.00	0.00	-0.03	0.65	0.00	0.00	0.00	0.00	0.96
Begin date: 6/2013	0.01	0.02	0.01	0.06	0.33	0.02	0.02	0.00	0.00	0.96
Begin date: 7/2013	0.01	0.01	0.00	0.02	0.72	0.01	0.01	0.00	0.00	0.95
Begin date: 8/2013	0.01	0.01	0.00	0.03	0.66	0.01	0.01	0.00	-0.02	0.83
Begin date: 9/2013	0.02	0.03	0.02	0.13	0.05	0.04	0.03	0.00	0.00	0.94
Begin date: 10/2013	0.01	0.02	0.01	0.07	0.30	0.03	0.02	0.00	-0.03	0.71
Begin date: 11/2013	0.01	0.01	0.00	0.02	0.72	0.01	0.01	0.00	0.00	0.94
Begin date: 12/2013	0.01	0.02	0.01	0.10	0.15	0.02	0.02	0.00	-0.01	0.86

	Before matching					After matching				
	Comparison	Intervention	adj.diff	std.diff	*p*-value	Comparison	Intervention	adj.diff	std.diff	*p*-value
Begin date: 1/2014	0.02	0.01	-0.01	-0.08	0.25	0.01	0.01	0.00	0.03	0.60
Begin date: 2/2014	0.03	0.02	-0.01	-0.06	0.36	0.03	0.02	0.00	-0.03	0.66
Begin date: 3/2014	0.05	0.04	-0.01	-0.03	0.61	0.04	0.04	0.00	0.00	0.96
Begin date: 4/2014	0.04	0.03	-0.01	-0.05	0.46	0.04	0.03	0.00	-0.02	0.80
Begin date: 5/2014	0.04	0.03	-0.01	-0.05	0.43	0.03	0.03	0.00	0.02	0.75
Begin date: 6/2014	0.08	0.09	0.01	0.05	0.46	0.08	0.09	0.01	0.05	0.42
Begin date: 7/2014	0.05	0.05	-0.01	-0.03	0.66	0.05	0.05	0.00	-0.02	0.82
Begin date: 8/2014	0.06	0.05	0.00	-0.02	0.76	0.05	0.05	0.00	0.00	0.95
Begin date: 9/2014	0.06	0.06	0.00	-0.02	0.80	0.05	0.06	0.00	0.02	0.79
Begin date: 10/2014	0.07	0.06	-0.01	-0.02	0.73	0.06	0.06	0.00	-0.02	0.78
Begin date: 11/2014	0.05	0.04	0.00	-0.02	0.74	0.05	0.04	-0.01	-0.05	0.46
Begin date: 12/2014	0.06	0.05	0.00	-0.02	0.80	0.05	0.05	0.00	0.00	1.00
Begin date: 1/2015	0.06	0.08	0.02	0.08	0.21	0.08	0.08	0.00	0.01	0.88
Begin date: 2/2015	0.05	0.06	0.01	0.04	0.57	0.06	0.06	0.00	0.00	0.99
Begin date: 3/2015	0.06	0.07	0.01	0.05	0.48	0.07	0.07	0.00	0.01	0.84
Begin date: 4/2015	0.04	0.03	-0.01	-0.03	0.67	0.04	0.03	0.00	-0.01	0.92
Begin date: 5/2015	0.05	0.05	0.00	-0.01	0.94	0.05	0.05	0.00	0.00	0.99
Begin date: 6/2015	0.03	0.02	-0.01	-0.07	0.32	0.02	0.02	0.00	0.02	0.79
# of hospitalizations baseline year	1.08	1.17	0.09	0.22	0.00	1.20	1.17	-0.03	-0.05	0.43
# of ED visits baseline year	1.28	2.08	0.80	0.81	0.00	2.42	2.50	0.07	0.04	0.57
# of office baseline year	10.46	14.86	4.40	0.56	0.00	13.96	14.86	0.90	0.10	0.12
SNF expenditures baseline year	833.56	2936.48	2102.92	0.47	0.00	2695.12	2936.48	241.36	0.03	0.65
Total Medicare expenditures baseline year	10061.49	15215.98	5154.49	0.30	0.00	15934.20	15215.98	-718.22	-0.03	0.59
Psychiatric visits baseline year	1.16	4.16	3.00	0.78	0.00	4.22	4.16	-0.06	-0.01	0.90
# of hospitalizations baseline year N/A	0.79	0.66	-0.13	-0.32	0.00	0.67	0.66	-0.01	-0.02	0.71
# of ED visits baseline year N/A	0.65	0.41	-0.23	-0.49	0.00	0.40	0.41	0.01	0.02	0.74

E. Specifications of measures

We used multiple types of measures in these analyses. CMMI requested that we calculate four core outcome measures for all awardees to the extent feasible: total Medicare expenditures, inpatient hospitalizations, hospital readmissions, and ED visits. The intervention group had only 35 readmissions over four years, far below the threshold of analysis (11 per quarter) used in previous reports. Due to this insufficient sample size, we did not pursue the readmission analysis. We describe our specifications for the other three standardized outcome measures in the first section below (Section E.1). In Section E.2, we then discuss specifications for the regression model covariates.

1. Core measures

We first describe how we identified the observation periods for the core measures, and then we describe how we constructed each of these outcome measures.

a. Defining baseline and intervention periods.

We defined baseline and intervention years for each participant or comparison group beneficiary relative to their enrollment[94] or pseudo-enrollment month (see Section D Step 1 above). The first intervention year was defined as the enrollment start month and the 11 months following that month. The first baseline year was the 12 months prior to the enrollment month, and additional baseline years were identified by moving backward from the first baseline year. For each individual included in the analysis, we calculated the proportion of each baseline and intervention year for which the individual was eligible for inclusion. This proportion was used to pro-rate the expenditure and utilization measures for individuals enrolled for less than the full analysis period. We also included this as a weighting variable in the regression analysis.

b. Calculating total Medicare expenditures

We included the following claim types in the analysis of Medicare expenditures: carrier, DME, home health, hospice, inpatient, outpatient, and skilled nursing facility. Only FFS data were included in this analysis. Part D services were excluded. Duplicate and denied claims were excluded. The total payment amount on each Medicare claim was summed across all file types to calculate total expenditures. For services that extended beyond a single day (for example, an inpatient or long-term care stay) we counted all Medicare payments recorded on a single date. Inpatient stay expenditures were counted in the month of the discharge date. For other types of claims, all expenditures were assigned based on the claim "from" date. We excluded expenditures from this analysis if they were assigned to a month during which the Medicare beneficiary was deemed ineligible for the analysis. We summed expenditures within the months assigned to a given analysis period to estimate total expenditures for the period.

c. Calculating inpatient hospitalizations

Estimating the number of hospitalizations in each period required several steps.

[94] For participants, we defined the enrollment month based on the date the individual first used a program service, according to the program enrollment data HLN provided (see Section B above).

Step 1: Identify hospitalization claims. Inpatient hospital claims were identified by claim type. For this measure only acute stays or psychiatric stays were included in the analysis. We identified and excluded rehabilitation and long-term care stays based on provider identifier codes.

Step 2: Eliminate duplicate or denied claims. We identified claims with the same information in all fields and only kept one of these claims. We also excluded denied claims from our analysis.

Step 3: Combine claims that represent the same stay and combine transfer stays with initial stays. We identified and combined initial and interim claims into one discharge. Interim claims had (1) the same admission date as the initial claim; (2) an admission date that was equal to the discharge date from the initial or another interim claim and the status on the other (previous) claim was "still a patient"; or (3) a claim with an admission date that was equal to one day after the discharge date of the initial or another interim claim and the status on the other previous claim was "still a patient." Such claims were combined to count as a single stay.

Next, we identified and combined claims associated with a transfer into a single stay. We identified claims indicating that the patient was transferred to either another short-term hospital, a Critical Access Hospital (CAH), another type institution for inpatient care, a federal hospital, or a psychiatric hospital or unit. Then we combined these claims with claims for the same beneficiary at a different facility where the admission date fell within one day of the discharge date of the first claim.

Step 4: Sum the number of discharges in each month. After we combined claims representing a single stay, we summed the number of unique discharges for each enrollee for each month. We counted inpatient stays in the month of the discharge date. We summed hospitalizations within the months assigned to a given analysis period to estimate total expenditures for the period.

d. Calculating emergency department visits

Outpatient ED visit utilization is reflected in CMMI priority measure 62. This measure includes ED visits and observation stays that do not lead to an inpatient stay.

In the Medicare outpatient file, we identified outpatient ED claims as those with a revenue center value indicating an ED visit, excluding any claims that involved only lab or imaging services in the ED. We identified observation claims based on the combination of revenue center code, CPT-code, and a unit count of greater than or equal to eight hours. We assigned ED visits and observation stays to a given calendar month based on the beginning service date. ED visits that led to inpatient stays (i.e., ones that share the same start date with an inpatient stay) were excluded. If two or more ED visits or observation stays had the same patient identifier and date of service, we counted them as one visit.

2. Other measures

The control variables included in the HLN regression models are listed in Table A-IV.5 along with the specifications for the variables. Because we expected similar inter-relationships

among these factors, we used the same set of control variables for each of the three outcome models.

Table A-IV.5. Impact analysis model control variable specifications (HLN)

Variable name	Specification
Intervention period	Categorical variable indicating time period. Categories are one year prior to enrollment, two years prior to enrollment, three years prior to enrollment, and intervention year
Treatment indicator	Categorical variable indicating treatment status. Categories include: comparison group member (reference); intervention participant
Interaction between intervention period and treatment	Interaction between intervention period and treatment indicator variables
Age	Continuous measure (age on first day of observation month —calculated based on date of birth)
Age (squared)	Age as specified above, squared
Gender	Categorical variable of member's gender. Categories include: female (reference) or male
Disability status	Binary variable indicating whether or not disability is original reason for Medicare entitlement
Service use	Categorical variable for type of service use at program entry. Categories include inpatient stays, primary care visits, SNF use
HCC score	Continuous measure of HCC score in the 12 months prior to intervention start date
Select CCW indicators	Binary indicators for the presence of the following conditions: Alzheimer's disease; hypoplasia; cataracts; endocarditis; hypertension; glycemia; osteoporosis
Select behavioral health conditions	Binary indicators for the presence of the following conditions diagnosed in the 12 months prior to program enrollment: schizophrenia; bipolar disorder; depressive disorder (any); substance use disorder (any); dementia; anxiety disorder (any)
Select somatic condition indicators	Binary indicators for the presence of the following conditions diagnosed in the 12 months prior to program enrollment: pain; headache; insomnia; hallucinations
Dual status	Categorical variable of dual status. Categories include: Medicare only (reference) and both Medicare and Medicaid
Pre-period history	Binary indicator for whether we had 12-month FFS claims history for person prior to begin date
Urbanicity	Categorical variable of county-level urbanicity based on patient residence at enrollment. Categories include: urban (reference); suburban; rural
Poverty	Categorical variable of county-level percent poverty based on patient residence at enrollment. Categories include tertiles indicating low, medium, and high poverty
Unemployment	Binary indicator of county-level unemployment based on patient residence. Categories include high unemployment or low unemployment (reference)

V. ICSI

A. Introduction

Our quantitative analyses for ICSI included pre-post analyses of the four core outcomes and health status measures. We did not construct a comparison group for ICSI because measures of program enrollment criteria were not observable for a potential comparison pool. Specifically, the COMPASS program uses scores from the Patient Health Questionnaire-9 (PHQ-9) and numerous diagnostic criteria (i.e., blood pressure, A1C levels, and lipid panels) to determine eligibility for the intervention. These measures were not available in Medicare or Medicaid claims data, and we were not able to identify any alternative data sources with comparable information that we could use to identify a comparison group of individuals who met the program enrollment criteria but who were not enrolled.

This chapter describes the methods for our pre-post analyses. We first describe the data sources (Section B) and approach to identifying the intervention group (Section C). Then, we specify how we constructed the outcome and control variables (Section D). Finally, we present the results from supplemental analyses (Section E).

B. Description of data sources

We used data from two sources:

COMPASS Registry data. Four of ICSI's eight partners [ICSI, Mayo Health System, Pittsburg Regional Healthy Initiative (PRHI), and Michigan Center for Clinical Systems Improvement (MiCCSI)] provided Mathematica with extracts from the COMPASS Registry (EHR). These data included enrollment date, demographic information, health status measures, care plans data, diagnoses, care management notes, medication history, and hospitalizations for all participants at each site. The data included Medicare identifiers for participants enrolled in Medicare.

Medicare administrative data. We extracted Medicare claims and enrollment data for Medicare-covered participants based on identifiers provided by the four sites linked to beneficiary identification numbers in the VRDC cross-reference files.[95] We extracted standard analytic base and revenue-center/line-item claims datasets for the following claim types: carrier, DME, home health, hospice, inpatient, outpatient, and skilled nursing facility. Then we extracted information on Medicare enrollment spans from the MBSF. The MBSF includes information on date of birth, gender, most recent county of residence, enrollment in Medicare Parts A and B, enrollment in Medicare Advantage, and third party insurance coverage.

C. Identification of intervention group

In this section, we provide background on the subset of intervention group members included in our analyses. As noted above, our analyses for ICSI were limited to a subset of four of eight partner sites. Below we discuss why we focused on this subset (Section C.1). Next, we

[95] Mathematica extracted data for program participants on April 29nd and 30th, 2016.

discuss the characteristics of the participants included in the Medicare FFS analysis (Section C.2) and the EHR data analysis (Section C.3).

1. Selection of sites to be included in the analysis

We selected sites for these analyses based on the size of the Medicare fee-for-service (FFS) population included in the intervention. (We did not base decisions on which sites to include based on Medicaid data because of delays in Medicaid data availability.) CMMI requested that we calculate four standardized outcome measures for all awardees to the extent feasible. These measures are: total Medicare and/or Medicaid expenditures, all-cause inpatient hospitalizations, 30-day unplanned hospital readmissions, and outpatient ED visits. The primary data sources available to calculate these measures were the Medicare and Medicaid administrative and claims data for FFS beneficiaries available through the VRDC. We excluded Medicare managed care enrollees from our analyses because claims for Medicare Advantage enrolled participants are excluded from CMS's Medicare administrative data. Similarly, Medicaid managed care encounter data available through the VRDC may not be reliable or comprehensive in many states.

In September 2014, ICSI provided Mathematica with data on each of their partners' enrolled Medicare and Medicaid populations. These data showed only 29 percent of program participants were enrolled in Medicare FFS. There were too few Medicaid FFS enrollees spread across multiple states to allow for robust analysis. Based on this, we selected partner sites for inclusion in this analysis based on the total number of Medicare FFS participants in the COMPASS intervention at the site. For example, we excluded Kaiser Permanente sites due to low numbers of Medicare FFS enrollees. We included four ICSI partners (50 percent of COMPASS partners) who had the largest number of Medicare FFS beneficiaries enrolled in the intervention. Across these four partners, the intervention was delivered at a total of 18 different clinic sites. The list of partners and their clinics is presented below in Table A-V.1.

Table A-V.1. ICSI partners and clinics included in analysis

ICSI partner	State	Clinics
Mayo Health System	Minnesota	Austin, Albert Lea, Owatonna/Faribault, Mankato, Red Wing, Rochester, and Jacksonville
Michigan Center for Clinical Symptoms Improvement (MiCCSI)	Michigan	Advantage Health, Lakeshore, Spectrum
Pittsburg Regional Health Initiative (PRHI)	Pennsylvania	Excela, Premier, St. Vincent's Medical Group
Institute for Clinical Symptoms Improvement (ICSI)	Minnesota	Entira, Essentia, Lakeview/Stillwater, North Memorial

Note: The partners not included in this analysis were Community Health Plan of Washington, Mount Auburn Cambridge Independent Practice Association, Kaiser Permanente of Southern California, and Kaiser Permanente of Colorado

2. Medicare FFS analysis population

As noted above, we extracted Medicare enrollment and claims data for Medicare FFS clients at the four sites included in these analyses (N = 1,148). Then, we identified the intervention enrollment month and the 24 months just prior to enrollment for each participant (or, the "pre-period"). We create analysis measures using each participant's pre-period and post-intervention

period data. These measures included: the states and counties of residence at enrollment, Medicare enrollment history throughout the pre- and post-periods, demographics at enrollment, health status at enrollment (based on diagnoses reported on claims), and the core outcomes measures.

We applied several exclusion criteria to the sample.

- **Insufficient FFS Medicare enrollment.** We excluded participants (n = 592) from the sample if they had less than six months of FFS Medicare Part A and B enrollment with Medicare as their primary payer during the pre- and post-periods. This included excluding participants who changed insurance (e.g., switched to a private insurance or Medicare Advantage resulting in less than 6-months of utilization data) or died during our intervention periods. This level of FFS Medicare enrollment was a minimum required to measure the individual's pre- and post-enrollment Medicare utilization and expenditures.
- **Residence in site service area.** We also excluded individuals (n = 52) from the ICSI treatment group if they did not reside in one of the three states listed in the Table A.V1 above (Michigan, Minnesota, or Pennsylvania) or Wisconsin (since the Mayo Health System has sites located on the border of Minnesota and Wisconsin) at the start of the intervention period.
- **Lack of demographic data or missing an enrollment date.** We excluded individuals (n = 75) who lacked demographic data or were missing enrollment information in the MBSF.

After applying these exclusion criteria (and some participants met more than one), our final analytic sample included N = 481 participants across the four partners. Table A-V.2 shows the demographics for the final sample.

Table A-V.2. Demographics of participants in Medicare FFS analysis

Characteristic	Number of participants	Percentage of participants
Total	**481**	**100**
Age group		
18-44	37*	8
45–54	73	15
55–64	98	20
65 or older	273	57
Gender		
Male	152	68
Female	329	32
State		
Michigan	97	20
Minnesota	194	40
Pennsylvania	137	28
Wisconsin	53	11

Source: Mathematica analysis of Medicare enrollment data for February 2013 to June 2015.

* Less than 11 people were in the 18-34 year old group. Therefore we combined this group with individuals 35-44.

We analyzed service use for each participant in six-month intervals relative to the participant's enrollment month. Enrollment dates varied across the participants in our sample from February 2013 through June 2015. The baseline period (that is, the "pre-period") included data for up to 24 months prior to enrollment. The follow-up period (that is, the "post-period") included data for the month of enrollment and up to 23 additional months. We created periods representing 6-month intervals across the 24-month baseline period and the 24-month intervention period. If participants were not enrolled in Medicare for the full baseline period, they may not be represented in all baseline intervals. Similarly, since the data available for our analysis ended in June 2015, we did not have a full 24 months of intervention period data for individuals who enrolled after June 2013. Table A-V.3 shows the sample size by the various time periods used in the analysis. The sample size is notably smaller in the first baseline period and the last intervention period.

Table A-V.3. Sample sizes by time period

Time period	Sample size
Baseline	
19-24 months prior to enrollment	315
13-18 months prior to enrollment	475
7-12 months prior to enrollment	481
0-6 months prior to enrollment	481
Intervention	
0-6 months post-intervention start	481
7-12 months post-intervention start	461
13-18 months post-intervention start	422
19-24 months post-intervention start	299

Source: Mathematica analysis of Medicare claims data from February 2011 to June 2015.

3. EHR analysis population

The four ICSI sites (Table A-V.1) who provided the Medicare identifiers also submitted COMPASS registry data to Mathematica. Each partner submitted a full extract of their EHR data for each intervention participant (regardless of Medicare status) from enrollment through September 1, 2015. We reviewed and de-duplicated the data. Table A-V.4 shows the demographic data for the EHR sample. Characteristics of the population for which EHR data were provided differed only slightly from the overall participants served by the COMPASS program (see report Chapter Table VIII.2).

Table A-V.4. Demographics of participants in EHR analysis

Characteristic	Number of participants	Percentage of participants
Total	**2,182**	**100**
Age group[a]		
18-34	98	4
35-44	208	10
45–54	495	23
55–64	604	28
65 or older	768	35

Characteristic	Number of participants	Percentage of participants
Gender[b]		
Male	802	37
Female	1,379	63
Insurance		
Medicaid, non-dual	463	21
Medicare, non-dual	1,071	49
Dual	119	5
Commercial	473	22
Self-Pay/Unknown	56	3

Source: Mathematica analysis of COMPASS EHR data for February 2013 to June 2015.

[a] Age was not reported for 9 participants.

[b] Gender was unknown for one participant.

D. Specifications of measures

In this section we provide information on the specifications for measures for the FFS Medicare analysis. First we discuss core measure specifications (Section D.1). Then we discuss control variable specifications (Section D.2).

1. Core measures

CMMI requested that we calculate four standardized outcome measures for all awardees to the extent feasible. These measures are: total Medicare and/or Medicaid expenditures, all-cause inpatient hospitalizations, 30-day unplanned hospital readmissions, and outpatient ED visits. If it was possible to calculate the core measures identified by CMS and these measures were appropriate to the intervention, we analyzed them. For ICSI we were unable to analyze readmissions due to small sample size. Specifications for the other measures are discuss below.

a. Medicare Expenditures

Medicare FFS expenditures for the following claim types were included in this measure: carrier, DME, home health, hospice, inpatient, outpatient, and skilled nursing facility. Duplicate and denied claims were excluded. The total payment amount on each Medicare claim was summed across all file types to calculate total expenditures in each month. For services that extend beyond a single month (for example, an inpatient or long-term care stay) we assigned all Medicare payments to a single month. Expenditures for inpatient stays were assigned to the month of the discharge date. Expenditures for all other types of claims were assigned based on the claim from date. Expenditures were excluded from this analysis if they were assigned to a month during which the associated Medicare beneficiary was deemed ineligible for the analysis.

b. Hospitalizations

The specification for this variable aligns with the CMMI priority all-cause admissions per patient measure. We describe the steps to develop this measure here.

Step 1: Identify acute and psychiatric inpatient claims. Medicare inpatient hospital claims were identified based on claim type. We excluded rehabilitation and long-term care stays

based on provider identifiers. At the end of this step, only acute and psychiatric stays were included in the file.

Step 2: Eliminate duplicate or denied claims. We identified claims with the same information in all fields and only kept one of these claims. We also excluded denied claims from our analysis file.

Step 3: Combine claims that represent the same stay and combine transfer stays with initial stays. We identified and combined initial and interim claims into one discharge. Interim claims had:

4. the same admission date as the initial claim or
5. an admission date that was equal to the discharge date from the initial or another interim claim and the status on the other (previous) claim was "still a patient" or
6. a claim with an admission date that was equal to one day after the discharge date of the initial or another interim claim and the status on the other previous claim was "still a patient."

Such claims were combined to count as a single stay.

Next, we identified and combined claims associated with a transfer into a single stay. We identified claims indicating that the patient was transferred to either another short-term hospital, a CAH, another type institution for inpatient care, a federal hospital, or a psychiatric hospital or unit. Then we combined these claims with claims for the same beneficiary at a different facility where the admission date fell within one day of the discharge date of the first claim.

Step 4: Sum the number of discharges in each month. Once claims representing a single stay were combined, we summed the number of unique discharges for each enrollee for each month. Inpatient stays were counted in the month of the discharge date.

c. ED visits

Outpatient ED visit utilization is reflected in CMMI priority measure 62. This measure includes ED visits and observation stays that do not lead to an inpatient stay.

In the Medicare outpatient file, we identified outpatient ED claims as those with a revenue center code indicating an ED visit, excluding any claims that involved only lab or imaging services in the ED. We identified observation claims based on the combination of revenue center code indicating an observation unit and a unit count of greater than or equal to eight hours.

d. Calculating core outcome measures

Once we identified the services and expenditures for each core measure for each month, the monthly measures were summed to the annual analysis periods. For individuals with less than 6 months of data for a given analysis period, the sum of the eligible months was divided by the proportion of the analysis period for which they were eligible to create a full-time equivalent measure. Regressions were weighted by the proportion of period for which the individual was eligible.

2. Control variables for regression analysis

The control variables included in ICSI's pre-post analyses are listed in Table A-V.5 along with the specifications for the variables.

Table A-V.5. Pre-post analysis variable specifications—ICSI

Variable name	Specification
Intervention period	Categorical variable indicating time period of observation. Categories include: baseline period (pre-intervention; reference) and six month intervention intervals for the two year period
Race	Categorical variable indicating the individual's race. Categories include: White; Black; Other (reference)
Age	Continuous variable indicating age at enrollment
Sex	Categorical variable of member's sex. Categories include: female (reference); male
Length of enrollment	Continuous variable indicating the total length of time a participant has been enrolled in the intervention
Dually enrolled in Medicare and Medicaid	Indicator variable for dually enrolled in Medicare and Medicaid in one or more months during the observation period
Aged	Indicator variable for original reason for Medicare entitlement based on old age; (reference category includes beneficiaries with original reason for entitlement based on disability, end-stage renal disease (end-stage renal disease or disability and end-stage renal disease)
Pre-period Medicare enrolled	Indicator variable for availability of 12 months of FFS Medicare claims data prior to first day of observation period
HCC score	Continuous variable measuring HCC risk score calculated based on Medicare FFS claims data for 12 months prior to enrollment
Depression	Indicator variable for presence of depression prior to the start of the intervention period
Acute myocardial infarction (AMI)[a,b]	Indicator variable for presence of AMI prior to the start of the intervention period
Atrial fibrillation[a,b]	Indicator variable for presence of atrial fibrillation prior to the start of the intervention period
Heart failure (HF)[a,b]	Indicator variable for presence of HF prior to the start of the intervention period
Hypertension[a,b]	Indicator variable for presence of hypertension prior to the start of the intervention period
Ischemic heart disease (IHD)[a,b]	Indicator variable for presence of IHD prior to the start of the intervention period
Hyperlipidemia[a,b]	Indicator variable for presence of hyperlipidemia prior to the start of the intervention period
Stroke/Transient Ischemic Attack (TIA)[a,b]	Indicator variable for presence of stroke/TIA prior to the start of the intervention period
Diabetes[a,b]	Indicator variable for presence of diabetes prior to the start of the intervention period
Acquired Hypothyroidism[a,b]	Indicator variable for presence of acquired hypothyroidism prior to the start of the intervention period
Alzheimer's Disease, related disorders, or senile dementia[a,b]	Indicator variable for presence of Alzheimer's disease, related disorders or senile dementia prior to the start of the intervention period
Anemia[a,b]	Indicator variable for presence of anemia prior to the start of the intervention period
Asthma[a,b]	Indicator variable for presence of asthma prior to the start of the intervention period
Benign Prostatic Hyperplasia (PBH)[a,b]	Indicator variable for presence of BPH prior to the start of the intervention period

Variable name	Specification
Cataract[a,b]	Indicator variable for presence of cataract prior to the start of the intervention period
Chronic Kidney Disease (CKD)[a,b]	Indicator variable for presence of CKD prior to the start of the intervention period
Chronic Obstructive Pulmonary Disease (COPD)[a,b]	Indicator variable for presence of COPD prior to the start of the intervention period
Glaucoma[a,b]	Indicator variable for presence of glaucoma prior to the start of the intervention period
Hip/pelvic fracture[a,b]	Indicator variable for presence of Hip/pelvic fracture prior to the start of the intervention period
Osteoporosis[a,b]	Indicator variable for presence of osteoporosis prior to the start of the intervention period
Rheumatoid Arthritis/ Osteoarthritis (RA/OA)[a,b]	Indicator variable for presence of RA/OA prior to the start of the intervention period
Breast cancer[a,b]	Indicator variable for presence of breast cancer prior to the start of the intervention period
Endometrial cancer[a,b]	Indicator variable for presence of breast cancer prior to the start of the intervention period
Prostate cancer[a,b]	Indicator variable for presence of prostate cancer prior to the start of the intervention period
Lung cancer[a,b]	Indicator variable for presence of lung cancer prior to the start of the intervention period
Colorectal cancer[a,b]	Indicator variable for presence of colorectal cancer prior to the start of the intervention period

[a] This variable was based on specifications provided by the Chronic Conditions Warehouse (available at: https://www.ccwdata.org/web/guest/condition-categories). For each month of the baseline period, we flagged whether the beneficiary met the criteria for the condition based on the condition-specific look-back period.

[b] This variable included in the model for total Medicare expenditures only.

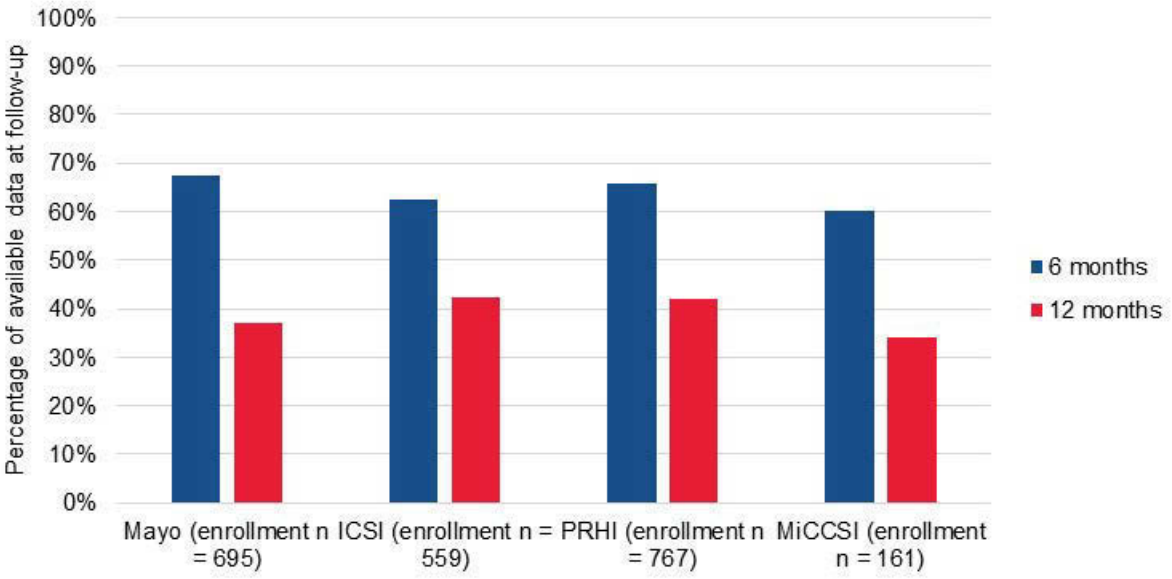
E. Supplemental EHR Analyses

We conducted additional analyses to examine potential outcome differences across the four sites and between the Medicare population (included in the analysis mentioned above) and non-Medicare participants not included in the analysis, but present in the EHR data.

1. Outcomes by site

Sample sizes and data availability varied by site. Figure A-V.1 shows the sample sizes and data availability for each site.

Figure A-V.1. Percentage of available data at 6 and 12 months by site

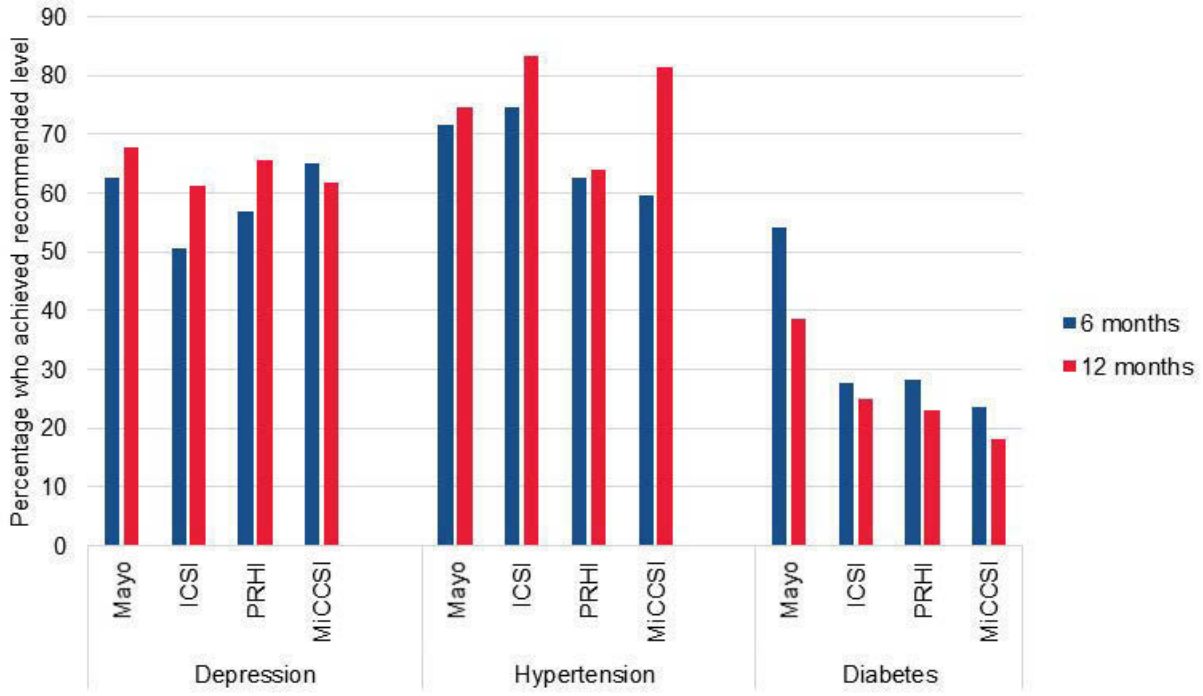


Source: Mathematica analysis of COMPASS Registry data.

Note: Sample sizes are provided for each health status measure at enrollment indicated by the "*enrollment n*=" text.

Outcomes at six months varied across the sites (Figure A-V.2). The percentage of participants with controlled diabetes was higher for Mayo Health System participants at both follow-up periods (when compared to the other three partners). Blood pressure control was also better for the ICSI and MiCCSI sites at 12 months. Depression levels also varied across sites with Mayo Health System and PHRI showing a higher rate of improvement at 12 months. It is important to note that these findings may be related to variations in the use of the COMPASS EHR system. Some sites reported that it was burdensome to enter information into their own EHR and the COMPASS system. Thus, it could be that some sites were more diligent at entering information into the EHR, resulting in difference in the documentation of participants' improvements. Additionally, sample size varied by site (Figure A-V.1). This variation may also reduce the accuracy of these results. These results should be interpreted with caution.

Figure A-V.2. Percentage of participants with suboptimal health status at enrollment who achieved recommended levels at six months by site

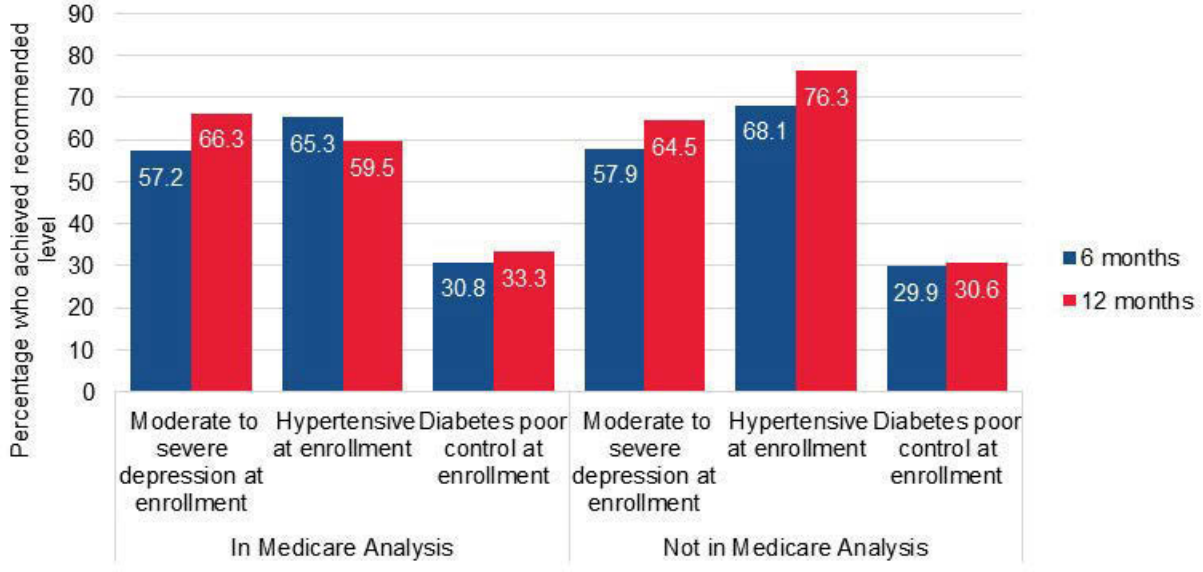


Source: Mathematica analysis of COMPASS Registry data.

2. Outcomes by Medicare analysis inclusion

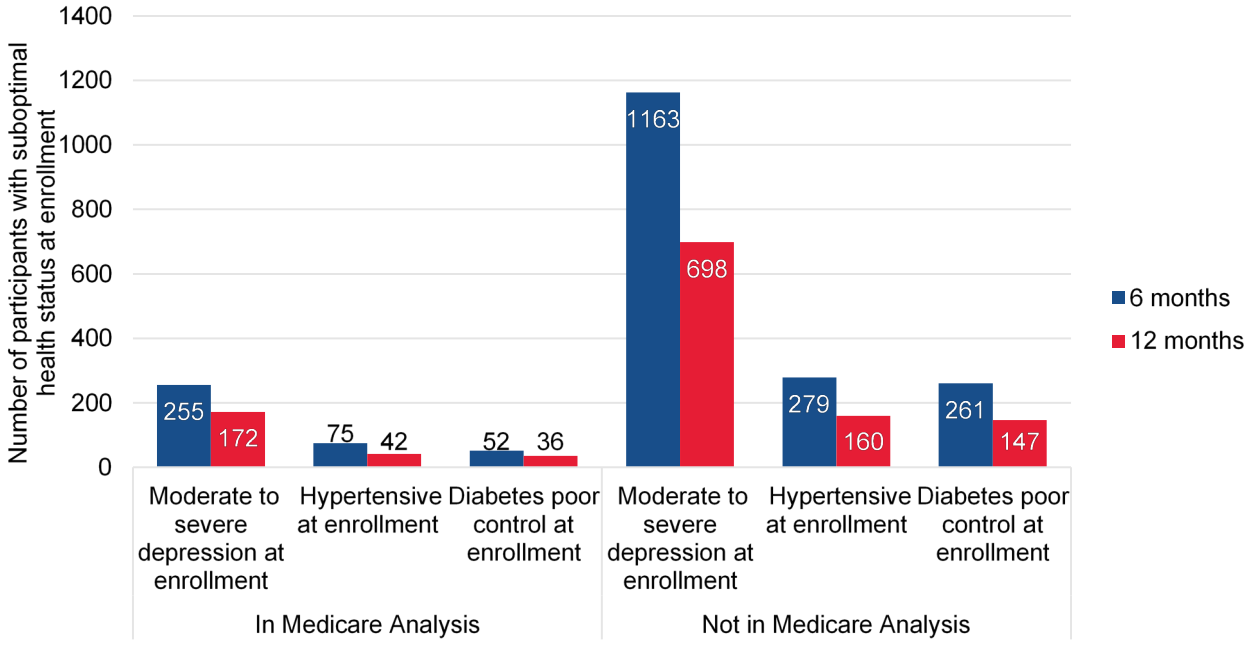
We conducted another analysis to assess whether there were substantial differences on health status measure outcomes between participants with and without FFS Medicare coverage (Figure A-V.3). Some minor differences were observed between the two groups. Specifically, patients without FFS Medicare coverage had a higher rate of improvement in hypertension at six months (68.1 percent versus 65.3 percent). The share of FFS Medicare participants with improved hypertension declined at 12 months from 65.3 percent to 59.5 percent while the share for the non-Medicare participants increased from 68.1 percent to 76.3 percent. It is unclear why this difference occurred and may be due to several factors (for example, participant characteristics, sample sizes at the 12-month timeframe, death, differential reporting). As a result, these results should be interpreted with caution. With the exception of the difference in 12-month improvement rate for hypertension, this analysis suggests that those included in our Medicare FFS analysis had similar outcomes to those who were not included. Figure A-V.4 shows the sample sizes for this analysis.

Figure A-V.3. Percentage of participants with suboptimal health status at enrollment who achieved recommended levels at follow-up by status of inclusion in the Medicare analysis



Source: Mathematica analysis of COMPASS Registry data.

Figure A-V.4. Sample sizes for Figure A-V.3



Source: Mathematica analysis of COMPASS Registry data.

VI. KMHS

A. Introduction

This appendix provides supplemental information on the quantitative analyses presented in Chapter VIII of the report. There, we presented findings from analyses of program impacts on Medicare and Medicaid utilization and expenditures and descriptive analyses of health status measures. Here, we first describe the data sources (Section B) and approach to selecting the analytic intervention and comparison populations for our analyses (Section C). Then, we specify how outcome and control variables were constructed (Section D). Finally, we present findings from supplemental analyses (Section E).

B. Description of data sources

In this section, we provide a general overview of the data sources used in the analyses for this awardee. In later sections of the appendix we provide more detail on how the data were used in the analyses.

1. CMS Medicare administrative data

Our analysis of impacts on Medicare utilization and expenditures used CMS Medicare administrative data. We obtained data files through the CMS's VRDC. We extracted all final action claims with dates of service from January 2009 through June 2015[96] for all individuals with a mental health visit billed by KMHS or by a comparison mental health provider (see Section C). We extracted standard analytic base and revenue-center/line-item claims datasets for the following claim types: carrier, DME, home health, hospice, inpatient, outpatient, and skilled nursing facility. To obtain information on beneficiary Medicare enrollment spans we used the MBSF. The MBSF includes information on date of birth, gender, most recent county of residence, enrollment in MA, and third party insurance coverage.

2. MAX and Alpha-MAX data for Washington state

Our analysis of impacts on Medicaid utilization used CMS Medicaid administrative data. We obtained data files through the CMS's VRDC. We extracted all final action claims with dates of service from January 2009 through June 2014[97] for all individuals in the Washington state files. We extracted standard analytic MAX (2009–2012) and Alpha-MAX (2013–2014) datasets for enrollment (PS files) and claim/encounter data including the following types: IP, LT, OT, and RX.[98] The Medicaid enrollment data included information on date of birth, gender, race and ethnicity, most recent county of residence, third party insurance coverage, and reason for eligibility.

[96] Claims were extracted in March 2016.

[97] Claims were extracted in October 2015.

[98] Our analyses used MAX and Alpha-MAX data for Washington state for the period from July 2011 through June 2014. Our analyses were limited to this period because managed care reporting for 2009 and 2010 was not comparable to the reporting for 2011 to 2014.

Since Washington state has high managed care penetration conducting our analysis required use of the managed care encounter data in the MAX and Alpha-MAX files. Since the quality and completeness of these data had not been validated, the first step in our analysis was to examine the usability of the encounter data in detail. First, we analyzed managed care enrollment data and then we analyzed the rate of claim submission by managed care plan and enrollee eligibility group (child, adult non-disabled, disabled, and aged).

Our analysis of managed care enrollment began by creating measures for each year for each patient indicating the number of months enrolled in a BHO and the number of months enrolled in an HMO. Based on review of BHO and HMO enrollment data, we found that BHO and HMO enrollment had ramped up over the proposed analysis period (2009–2014). Only a small share of the disabled population was enrolled in HMOs in 2011–2012; however, that share increased substantially in 2013. We also found that BHO enrollment ramped up between 2009 and 2010, and was constant between 2011 and 2013. Because we planned to use mental health visits to identify participants and comparison group members, we wanted consistency in mental health providers during the analysis period. We were concerned that including a period in which there was significant BHO enrollment ramp up would result in inconsistencies in mental health provider use and service utilization reporting. Therefore, we decided to analyze 2011–2014 Medicaid data, because managed care enrollment was more stable in this period. Further cementing our decision to use 2011 to 2014 data for analysis, we found through discussions with state and internal staff, and review of data that: 1) the state relied more on state-specific codes in 2009 and moved to more frequently use standard, national codes in later years, and 2) the state MMIS system changed in May 2010, creating data issues such as changes in provider IDs and plan IDs, as well as creating the possibility for other, unknown changes that might affect the analysis.

To further review the encounter data and determine whether or not the encounter data for each plan was complete enough to use for analysis, we analyzed claim submission rates by plan ID and enrollee eligibility group. We calculated the following measures for each analysis year (2011–2014) for each BHO and HMO plan: number of IP stays per 1,000 enrolled months, number of ED visits per 1,000 enrolled months, and number of mental health visits per 1,000 enrolled months. We reviewed the number of IP stays, ED visits, or MH visits by plan, as well as the average and standard deviation of the three measures. We excluded from analysis patients that, during any analysis year, were enrolled on one of three plans that did not report any encounters in either the IP or OT file in any analysis year (plan IDs 105010404, 105010405, and 105010406). We also excluded patients from the analysis who, during any analysis year, were in one plan that did not meet the data quality threshold[99] we set based on the average and standard deviations for the measures (plan ID 105020401).

[99] Our threshold was: for each plan for each year, plans' IP, ED, or mental health encounters per enrolled month must be within two standard deviations of the average of all plans' IP, ED, or mental health encounters per enrolled month.

3. Finder files

KMHS provided finder files listing the Medicare and Medicaid program identifiers for patients enrolled in these programs. These identifiers could be used to identify KMHS patients in the Medicare and Medicaid administrative data.

4. EHR data

For the KMHS health status measures analysis, we used EHR data that was provided by the awardee. These data included demographic information, the date and type of all services used, health status measures, diagnoses, and medication history.

C. Identification of intervention and comparison pool members

In this section we describe the steps taken to select the intervention and comparison pools for each of our analyses (Medicare analysis, Medicaid analysis, and health status measure analysis). For the Medicare and Medicaid analyses we used matching methods to select the comparison group. Propensity score matching and related matching methods are designed to create a comparison group of nonparticipants who are similar in observable characteristics to KMHS Medicare participants (Rosenbaum and Rubin 1983; Dehejia and Wahba 2002).

1. KMHS Medicare

Because all KMHS patients are considered intervention participants, we identified all patients who received services at KMHS as intervention group members and patients of other mental health treatment facilities in the state of Washington as the potential pool of comparison patients. Then from within the comparison pool we identified individuals most closely matched to KMHS patients to include in the comparison population. Constructing the matched comparison group involved several steps, which we detail below.

Step 1a: Identify facilities similar to KMHS in Washington state. Using the Substance Abuse and Mental Health Services Administration's mental health treatment facility locator, we identified all mental health treatment facilities in Washington state in 2014 with the following characteristics:

- Provides outpatient care
- Serves patients with Medicaid and Medicare
- Privately owned
- Serves adults
- Allows psychiatric emergency walk-in clients

Based on this set of characteristics, we identified 24 facilities. We considered requiring facilities to match additional characteristics of KMHS such as providing multiple levels of care, having special targeted programs,[100] or being in a geographic area of similar size; however, this

[100] KMHS provides multiple levels of care including residential and hospital care. KMHS also has special programs for individuals with severe mental illness and for individuals with mental health and substance abuse disorders.

would reduce the number of facilities from which to identify potential comparison group members to only five and would not allow for a sufficient number of potential comparison clients well-matched to KMHS clients. The current analysis period includes calendar years 2010 through June 2015. Of the 24 facilities initially identified, we excluded 7 facilities because they did not serve Medicare clients in all five and a half analysis years. We excluded one additional facility because multiple locations used the same NPI preventing us from identifying those services provided at the location that met the facility selection criteria. Thus, 16 comparison facilities were used in the analysis.

Step 1b: Identify additional facilities treating patients with dementia. When we compared the diagnoses reported on claims for KMHS patients to those for patients served by comparison facilities, we found substantial numbers of KMHS clients had a diagnosis of dementia; however few of the patients at the comparison facilities had a dementia diagnosis. Thus, in order to assure a sufficient number of comparison pool members well-matched to the KMHS clients with dementia, we identified additional facilities in the state that served at least 100 patients with a diagnosis of dementia on a psychiatric service claim. We included patients with dementia from these additional facilities in the pool of potential comparison group members, and only matched these patients with treatment group members with dementia.

Step 2: Identify treatment and potential comparison group members. Using Medicare data for calendar years 2010 through June 2015, we initially identified all individuals who receive a mental health service at KMHS or one of the potential comparison facilities.[101] We used CPT and ICD-9 diagnosis codes to identify mental health services. Individuals with a claim meeting any one of the three mental health service category definitions in Table A-VI.1 were selected for our initial analysis population. It should be noted that on January 1, 2013 the CPT codes used to bill psychiatric services changed. Providers began using new psychiatric visit codes 90791, 90792, and 90785 on that date. The psychiatric medication management code 90862 was not allowable beginning January 1, 2013. After this date providers billed appropriate evaluation and management codes with a mental health primary diagnosis. Each individual who received a mental health service was assigned to an intervention or comparison group based on the facility in which they initially received treatment.[102] Medicare enrollment and claims data for January 2009 through June 2015 were extracted for this population and used to develop measures of enrollment history, demographics, health conditions, and HCC score. Health conditions and HCC score were measured in the 12-month period prior to the month of the initial mental health visit at KMHS or a comparison facility in January 2010 or later. Mental health diagnosis at treatment initiation (in a category listed in Table A-VI.2) was measured in the initial month of mental health treatment and the two subsequent months. We allowed the two subsequent months because facilities commonly used a 799.9 (unknown or unspecified cause of morbidity) code during initial visits until they had specified a diagnosis.

[101] We include individuals with limited exposure to KMHS in both the pre- and post-period to reflect the general population treated at KMHS. The intervention may also increase the number of visits at KMHS, and therefore we did not want to include the number of visits as a selection criteria.

[102] Eighteen individuals were excluded because they were observed to receive services at more than one facility in their initial treatment month and could not be attributed to only one facility.

Table A-VI.1. Codes used to identify mental health services (KMHS Medicare)

Service category	CPT codes and additional requirements
1. Psychiatric visit	CPT-code = 90801 through 90899, 90791, 90792, and 90785 (psychiatric visit)
2. E&M visit with psych primary DX	CPT-code = any outpatient E&M visit (CPT=99201-99205, 99211-99215) with a MH primary diagnosis code listed in Table A.10
3. Psychiatric medication management visit	CPT-Code=M0064[a]

[a] M0064 was deleted from the HCPCS system December 31, 2014. Thus, this code was in use through the end of the period we used to identify patients for this analysis.

Table A-VI.2. ICD-9 Mental health diagnosis codes (KMHS Medicare)

Diagnosis group	ICD-9 Diagnosis code value
Schizophrenic disorders	295.xx including 295.00
Bipolar disorders	296.00, 296.01, 296.02, 296.03, 296.04, 296.05,296.06, 296.10, 296.11, 296.12, 296.13, 296.14,296.15, 296.16, 296.40, 296.41, 296.42, 296.43,296.44, 296.45, 296.46, 296.50, 296.51, 296.52, 296.53, 296.54, 296.55, 296.56, 296.60, 296.61, 296.62, 296.63, 296.64, 296.65, 296.66, 296.7, 296.80, 296.81, 296.82, 296.89, 296.90, 296.99
Depressive disorders	296.20, 296.22, 296.23, 296.24, 296.25, 296.26, 296.30, 296.32, 296.33, 296.34, 296.35, 296.36, 311
Persistent mental disorders due to conditions classified elsewhere	294.8x, 294.9x
Dementia	290.xx, 294.1x
Other psychotic disorders	297.xx-298.xx
Anxiety, dissociative, and somatoform	300.xx
Adjustment reaction	309.xx
Drug and alcohol indicator	292, 292.0, 292.1, 292.2, 292.8, 292.9, 304, 304.0, 304.1, 304.2, 304.3, 304.4, 304.5, 304.6, 304.7, 304.8, 304.9, 305, 305.2, 305.3, 305.4, 305.5, 305.6, 305.7, 305.8, 305.9 291, 291.0, 291.1, 291.2, 291.3, 291.4, 291.5, 291.8, 291.9, 303, 303.0, 303.9, 305.0
Other diagnosis not listed above	Everything not above (293.83, V62.84, V62.85, E950, E951, E952, E953, E954, E955, E956, E957, E958, E959, 301.0 to 301.9, 307.40 to 307.49, 312.0 to 312.23, 312.4 to 312.89, 313.81, 312.30 to 312.39, 302.0 to 302.9, 299.00 to 299.91, 307.1, 307.5, 307.51, 314.00 to 314.01, 307.20 to 307.3, 313.0 to 313.3, 313.82 to 316, 648.4, V65.2, V71.09, 780.09, V15.41, V15.42, V15.81, V17.0, V60.0, V62.29, V62.4, V62.81, V62.89) and all other codes in the range of 290.0-299.91 and 300.00-316 Also include 7999 in this category.
Any 294 diagnosis	294.xx

We restricted the analysis population to those residing in the local area of the analysis facilities to assure the patients had the potential to consistently access the facilities during the analysis period. We excluded individuals from the KMHS treatment group if they did not reside in Kitsap County or a contiguous county based on the most recent Medicare enrollment data available at the time they received their initial mental health service at KMHS. Potential comparison group members were similarly excluded if they did not reside in the county or a contiguous county for the mental health facility at which they initially received services.

Next, because of the limitations of the available Medicare data and to assure consistency in the expenditures observable for the analysis population, we required that during the full analysis period (1) the individual not be enrolled in Medicare Advantage (because we do not have access to managed care encounters), (2) have Medicare as their primary payer, and (3) be enrolled in Medicare Parts A and B (to ensure that we capture both inpatient and outpatient services). Applying these restrictions in a step-wise fashion resulted in the exclusion of 15 percent, 2 percent, and 1 percent of the analysis population, respectively. We also required that the individual have a value for the hierarchical behavioral health diagnosis variable; we excluded another four individuals due to this requirement.

When this step was complete the analysis population included 1,116 KMHS intervention participants and a pool of 12,017 individuals who received mental health services from comparison facilities.

Step 3: Match treatment participants at the individual level. The next step involved creating a matched comparison group. The matching process used metrics of individual-level characteristics identified based on pre-period Medicare data to select a subset of comparison pool members who were as similar as possible to the intervention group on observable characteristics. The matching algorithm first exact matched on the year an individual began treatment at KMHS or comparison mental health facility and a hierarchical variable of behavioral health diagnosis in the first three months of mental health treatment. The hierarchical variable included the following categories: dementia, schizophrenia, bipolar disorder, depression, or other condition. Then, within these cells, we used propensity score matching,[103] to match up to five comparison pool beneficiaries to each intervention group member. When a treatment beneficiary was difficult to match (that is, had few similar comparison beneficiaries in the same cell), the algorithm conducted a pairwise matching; when there were an abundance of comparisons for a treatment beneficiary, the algorithm matched multiple comparisons. The statistical goal is first to minimize bias and then, subject to that constraint, maximize the size of the comparison sample. The propensity score model included the following characteristics: age group (18–44, 45–54, 55–64, 65+), gender, disability status, year began treatment at KMHS or comparison mental health facility, whether the beneficiary was enrolled in Medicare for a full 12 months prior to receiving mental health treatment at KMHS or a comparison facility, Medicare/Medicaid dual enrollment status, flags for psychiatric conditions,[104] and HCC score.[105]

When this step was complete the analysis population included 1,116 KMHS intervention participants and 4,003 individuals in the comparison group. The reduction in the size of the

[103] A member's propensity score is the probability of belonging to the intervention group estimated based on a logistic regression model.

[104] We created flags to indicate that the patient had a diagnosis code for various conditions in the first three months of their claims during the intervention period. The diagnosis-related flags that we included in the matching included those for persistent mental disorder due to conditions classified elsewhere; dementia; anxiety, dissociative, or somatoform disorder; adjustment reaction disorder; alcohol- or drug-related diagnosis; "other" psychotic disorder; and "other" diagnosis.

[105] HCC score was used only for individuals enrolled in Medicare for 12 months prior to receiving a treatment at KMHS or a comparison facility because 12 months of claims history are required to calculate the score based on medical conditions.

comparison population relative to the previous step was due to individuals who were not matched to an individual attributed to KMHS.

Step 4: Assess the quality of the match. The following tests and procedures were used to verify that the treatment and comparison groups are similar or balanced. After we conducted matching, we examined the number of comparison beneficiaries matched to each treatment beneficiary (Table A-VI.3). A large number of 1:1 matches, or a large number of comparison beneficiaries that were excluded, could indicate that the matching was problematic. In this case, we examined the balance diagnostics described below to determine which variable(s) may be causing the difficulty. The number of 1:1 matches is generally related to the small number of potential comparison group members in a given exact matching cell with the same hierarchical behavioral health diagnosis. Although requiring an exact match on diagnosis category increased the number of pairwise matches, we believed it was important that the treatment and associated comparison group member match on this characteristic.

Table A-VI.3. Frequency table of ratio of treatment beneficiaries to comparison beneficiaries for each matched set (KMHS Medicare)

Ratio of treatment to comparison beneficiaries	1:1	1:2	1:3	1:4	1:5
Number of matched sets	292	81	56	54	633

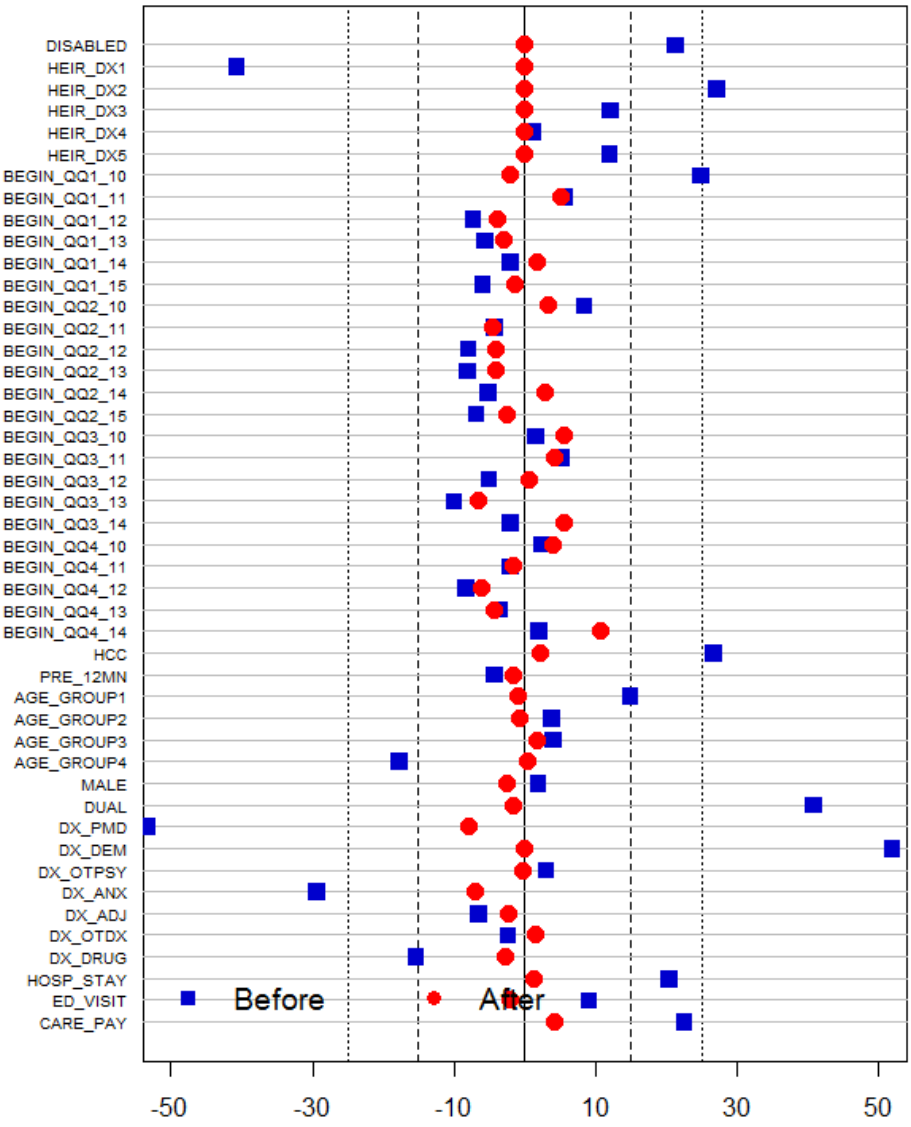
Note: Each cell indicates the number of treatment beneficiaries matched to the number of comparison beneficiaries indicated for that column. In this example, most of the treatment beneficiaries (633) were matched to 5 comparison beneficiaries.

Next, we examined the overall balance of the matched sample. We used an omnibus test that checks for covariate balance across the individuals in the treatment and matched comparison group (Hansen and Bowers 2008). The omnibus test is based on the differences between the individuals in the treatment and matched group across the matching variables; these differences are standardized by their variances and covariances and aggregated into a single number, a weighted mean. Standardization in this way implies that a matching variable whose difference across matched sets has a small variance is given more weight and that a matching variable whose difference across sets is highly correlated with other differences is given less weight. The advantages of the omnibus test are: (1) it generates a single probability statement through one *p*-value; (2) its distribution is roughly chi-square, which facilities the calculation of the *p*-value; and (3) it assesses balance on all linear combinations of the matching variables. However, a significant result from this chi-square test may be driven by a large sample rather than substantive differences between treatment and matched comparison groups. Alternatively, it could indicate that there may be some imbalance between the two groups on at least one of the matching variables. The results of this test were a chi-square statistic of 93.5 and a *p*-value of < 0.01 indicating an imbalance exists.

To further investigate imbalance between treatment and matched comparison groups, we evaluated how matching affected the balance on all matching variables (Figure A-VI.1) by comparing the absolute and standardized difference between the treatment and control groups for each variable before and after matching. The standardized difference measures the difference in means in *units* of the pooled standard deviation of treatment group and comparison group. The standardized difference measure is advantageous in that it allows us to compare all variables on

the same scale. We compared the standardized differences using plots with dashed lines at +/- 0.15 standardized differences to visually inspect whether we obtained good balance for each variable, and using a balance table that shows both absolute and standardized differences between treatment and comparison groups before and after matching.

Figure A-VI.1. Balance plot comparing the standardized difference for each matching variable before and after matching (KMHS Medicare)



Note: Blue markers show the standardized difference before matching; red markers show the standardized difference after exact matching and propensity score modeling. See Table A-VI.4 for descriptions of the variables included in this figure.

We provide more detail on the means and adjusted and standardized difference for the matching variables in Table A-VI.4 below.

Table A-VI.4. Balance table before and after matching (KMHS Medicare)

		Before matching					After matching				
	Variable description	Comparison	Treatment	adj.diff	std.diff	p	Comparison	Treatment	adj.diff	std.diff	p
DISABLED	Disability status	0.5811	0.6855	0.1044	0.2126	0	0.6855	0.6855	0	0	1
HEIR_DX1	Hierarchical variable of behavioral health diagnosis: dementia	0.3434	0.1532	-0.1902	-0.408	0	0.1532	0.1532	0	0	1
HEIR_DX2	Hierarchical variable of behavioral health diagnosis: schizophrenia	0.1673	0.2697	0.1025	0.2698	0	0.2697	0.2697	0	0	1
HEIR_DX3	Hierarchical variable of behavioral health diagnosis: bipolar disorder	0.136	0.1774	0.0414	0.1197	0.0001	0.1774	0.1774	0	0	1
HEIR_DX4	Hierarchical variable of behavioral health diagnosis: depression	0.2092	0.2133	0.0041	0.01	0.75	0.2133	0.2133	0	0	1
HEIR_DX5	Hierarchical variable of behavioral health diagnosis: other condition	0.1441	0.1864	0.0423	0.1191	0.0001	0.1864	0.1864	0	0	1
BEGIN_QQ1_10	Began treatment at KMHS or comparison mental health facility in first quarter of 2010	0.229	0.3342	0.1052	0.2477	0	0.3448	0.3342	-0.0106	-0.0217	0.2921
BEGIN_QQ1_11	Began treatment at KMHS or comparison mental health facility in first quarter of 2011	0.0352	0.0457	0.0105	0.0563	0.0721	0.0358	0.0457	0.0099	0.0515	0.0948
BEGIN_QQ1_12	Began treatment at KMHS or comparison mental health facility in first quarter of 2012	0.0349	0.0215	-0.0134	-0.074	0.018	0.0275	0.0215	-0.006	-0.0386	0.2471
BEGIN_QQ1_13	Began treatment at KMHS or comparison mental health facility in first quarter of 2013	0.0387	0.0278	-0.0109	-0.0573	0.0672	0.0325	0.0278	-0.0048	-0.0292	0.4061
BEGIN_QQ1_14	Began treatment at KMHS or comparison mental health facility in first quarter of 2014	0.0353	0.0314	-0.0039	-0.0214	0.4951	0.0286	0.0314	0.0028	0.0171	0.6185
BEGIN_QQ1_15	Began treatment at KMHS or comparison mental health facility in first quarter of 2015	0.0341	0.0233	-0.0108	-0.0604	0.0536	0.0256	0.0233	-0.0023	-0.0151	0.678
BEGIN_QQ2_10	Began treatment at KMHS or comparison mental health facility in second quarter of 2010	0.0582	0.078	0.0198	0.0834	0.0077	0.0695	0.078	0.0084	0.0332	0.2314
BEGIN_QQ2_11	Began treatment at KMHS or comparison mental health facility in second quarter of 2011	0.0318	0.0242	-0.0076	-0.0437	0.1624	0.0314	0.0242	-0.0072	-0.0448	0.1898
BEGIN_QQ2_12	Began treatment at KMHS or comparison mental health facility in second quarter of 2012	0.0364	0.0215	-0.0149	-0.0808	0.0099	0.0277	0.0215	-0.0062	-0.0423	0.2171
BEGIN_QQ2_13	Began treatment at KMHS or comparison mental health facility in second quarter of 2013	0.0387	0.0233	-0.0154	-0.0812	0.0095	0.03	0.0233	-0.0067	-0.0404	0.2196
BEGIN_QQ2_14	Began treatment at KMHS or comparison mental health facility in second quarter of 2014	0.0376	0.0278	-0.0098	-0.0523	0.0949	0.0236	0.0278	0.0042	0.0274	0.403

		Before matching					After matching				
	Variable description	Comparison	Treatment	adj.diff	std.diff	p	Comparison	Treatment	adj.diff	std.diff	p
BEGIN_QQ2_15	Began treatment at KMHS or comparison mental health facility in second quarter of 2015	0.0393	0.026	-0.0133	-0.0694	0.0266	0.03	0.026	-0.004	-0.0254	0.4304
BEGIN_QQ3_10	Began treatment at KMHS or comparison mental health facility in third quarter of 2010	0.0409	0.0439	0.003	0.0153	0.6238	0.0339	0.0439	0.0101	0.0543	0.0643
BEGIN_QQ3_11	Began treatment at KMHS or comparison mental health facility in third quarter of 2011	0.0328	0.0421	0.0093	0.0518	0.098	0.0346	0.0421	0.0075	0.0424	0.1867
BEGIN_QQ3_12	Began treatment at KMHS or comparison mental health facility in third quarter of 2012	0.0334	0.0242	-0.0092	-0.0517	0.0987	0.0234	0.0242	0.0008	0.0056	0.8805
BEGIN_QQ3_13	Began treatment at KMHS or comparison mental health facility in third quarter of 2013	0.0327	0.0152	-0.0175	-0.1005	0.0013	0.0248	0.0152	-0.0096	-0.0669	0.0607
BEGIN_QQ3_14	Began treatment at KMHS or comparison mental health facility in third quarter of 2014	0.0389	0.0349	-0.004	-0.0208	0.5072	0.0257	0.0349	0.0092	0.056	0.1053
BEGIN_QQ4_10	Began treatment at KMHS or comparison mental health facility in fourth quarter of 2010	0.0384	0.043	0.0046	0.0236	0.4502	0.0356	0.043	0.0074	0.04	0.1992
BEGIN_QQ4_11	Began treatment at KMHS or comparison mental health facility in fourth quarter of 2011	0.0334	0.0296	-0.0038	-0.0213	0.497	0.0324	0.0296	-0.0029	-0.0168	0.6064
BEGIN_QQ4_12	Began treatment at KMHS or comparison mental health facility in fourth quarter of 2012	0.0325	0.0179	-0.0146	-0.084	0.0073	0.0264	0.0179	-0.0085	-0.0605	0.0852
BEGIN_QQ4_13	Began treatment at KMHS or comparison mental health facility in fourth quarter of 2013	0.0354	0.0287	-0.0068	-0.0369	0.2379	0.036	0.0287	-0.0073	-0.0424	0.224
BEGIN_QQ4_14	Began treatment at KMHS or comparison mental health facility in fourth quarter of 2014	0.0325	0.0358	0.0034	0.019	0.5429	0.02	0.0358	0.0158	0.106	0.0037
HCC	HCC score	1.3122	1.5982	0.286	0.2664	0	1.5759	1.5982	0.0223	0.0226	0
PRE_12MN	Beneficiary was enrolled in Medicare for a full 12 months prior to receiving mental health treatment at KMHS or a comparison facility	0.8332	0.8172	-0.016	-0.0429	0.1707	0.824	0.8172	-0.0068	-0.0172	0.0547
AGE_GROUP1	Age group 18-44	0.2162	0.2778	0.0616	0.1484	0	0.2826	0.2778	-0.0048	-0.0102	0.2994
AGE_GROUP2	Age group 45-54	0.1663	0.1801	0.0138	0.0368	0.239	0.1831	0.1801	-0.003	-0.0074	0.5472
AGE_GROUP3	Age group 55-64	0.1292	0.1425	0.0132	0.0393	0.2089	0.1361	0.1425	0.0063	0.0177	0.2678
AGE_GROUP4	Age group 65+	0.4882	0.3996	-0.0886	-0.1775	0	0.3982	0.3996	0.0015	0.0032	0.7651
MALE	Gender	0.4372	0.4462	0.009	0.0182	0.5612	0.4587	0.4462	-0.0125	-0.025	0.076
DUAL	Medicare/Medicaid dual enrollment status	0.5476	0.7482	0.2007	0.4073	0	0.7549	0.7482	-0.0066	-0.0161	0.3613

		Before matching					After matching				
	Variable description	Comparison	Treatment	adj.diff	std.diff	p	Comparison	Treatment	adj.diff	std.diff	p
DX_PMD	Diagnosis of persistent mental disorders due to conditions classified elsewhere	0.2155	0.0054	-0.2102	-0.5335	0	0.0132	0.0054	-0.0078	-0.0799	0.0032
DX_DEM	Diagnosis of dementia	0.0448	0.1613	0.1165	0.5178	0	0.1613	0.1613	0	0	1
DX_OTPSY	Diagnosis of other psychotic disorder	0.0517	0.0582	0.0066	0.0295	0.3456	0.059	0.0582	-0.0007	-0.0033	0.8091
DX_ANX	Diagnosis of anxiety, dissociative, or somatoform disorder	0.1108	0.0215	-0.0893	-0.2944	0	0.032	0.0215	-0.0105	-0.0697	0.0016
DX_ADJ	Diagnosis of adjustment reaction disorder	0.0875	0.069	-0.0185	-0.0662	0.0345	0.0751	0.069	-0.0061	-0.0238	0.0172
DX_OTDX	Other behavioral health diagnosis	0.0492	0.0439	-0.0053	-0.0245	0.4338	0.0414	0.0439	0.0025	0.0144	0.3193
DX_DRUG	Drug and/or alcohol-related diagnosis	0.0312	0.0054	-0.0258	-0.154	0	0.0076	0.0054	-0.0022	-0.0278	0.3734
HOSP_STAY	Hospitalizations utilization outcome measure	0.3948	0.5923	0.1975	0.2029	0	0.5808	0.5923	0.0115	0.0116	0.7635
ED_VISIT	ED visits utilization outcome measure	1.2451	1.517	0.272	0.0897	0.0042	1.7656	1.6998	-0.0658	-0.0205	0.6159
CARE_PAY	Total expenditures outcome measure	10855.2202	15700.1065	4844.8863	0.2241	0	14694.8164	15601.1592	906.3428	0.0413	0.2298

Step 5: Create analysis weights. Weights were developed for each member of the analysis population. Weights for KMHS attributed individuals were set to one. Weights for comparison group members were set to one divided by the number of comparison group members assigned to the member's associated treatment person. An individual's participation in the analysis could be terminated as a result of a change in status before the end of the analysis period. An individual's weight was set to zero in analysis months following any of these status changes. There were four status changes for which individuals were dropped from the analysis: (1) to assure consistency of care within the treatment and comparison groups, we removed individuals from the analysis if they received services at a mental health facility other than their assigned facility; (2) we also removed individuals from the analysis if they moved out of the set of counties designated for their assigned facility, because they would have no or less access to the assigned facility; (3) individuals who were no longer enrolled in Medicare were dropped from our analysis because they were no longer included in the data available for analysis; and lastly, (4) individuals were dropped from our analysis if they died.

2. KMHS Medicaid

In this section we describe the steps taken to select the intervention and matched comparison groups for the KMHS Medicaid analysis and provide diagnostics to assess balance between the matched groups. We describe these steps below.

Step 1: Identify treatment and potential comparison group members. The method used to create the treatment and potential comparison group pool for the Medicaid analysis differs from the Medicare analysis since we were unable to identify mental health treatment facilities in the Medicaid administrative data as we did in the Medicare data. Using Medicaid data for calendar years 2011 through 2014, we identified the treatment group members by using a finder file from the KMHS EHR data, limiting to those patients who had an in-person visit at KMHS in the EHR data on or after January 1, 2013, and finally, limiting to those patients with a mental health service in the Medicaid data during 2013 or 2014.[106] We then identified the potential comparison group members as those not in the treatment group, who did not have an in-person visit at KMHS in the EHR data after January 1, 2011, and who did have a mental health service in the Medicaid data during 2013 or 2014. For the treatment and comparison pool, we used CPT and ICD-9 diagnosis codes to identify mental health services in the Medicaid data (Table A-VI.5). It should be noted that on January 1, 2013 the CPT codes used to bill psychiatric services changed. Providers began using new psychiatric visit codes 90791, 90792, and 90785 on that date. The psychiatric medication management code 90862 was not allowable beginning January 1, 2013. After this date providers billed appropriate evaluation and management codes with a mental health primary diagnosis. Medicaid enrollment and claims data for January 2011 through June 2014 were extracted for the treatment population and comparison pool and used to develop measures of enrollment history, demographics, health conditions, and CDPS flags.[107] Mental

[106] We include individuals with limited exposure to KMHS in both the pre- and post-period to reflect the general population treated at KMHS. The intervention may also increase the number of visits at KMHS, and therefore we did not want to include the number of visits as a selection criteria.

[107] The CDPS is a diagnosis-based risk adjustment model that was designed to allow Medicaid managed care organizations to adjust capitation payments for TANF and disabled beneficiaries for enrollee health status using administrative claims data. This is complemented by the Medicaid Rx (MRx) chronic disease classification, which

health diagnoses, identified by the codes in Table A-VI.6, at treatment initiation were measured in the initial month of mental health treatment and the two subsequent months. Physical health conditions, identified by the codes in Table A-VI.7, and CDPS flags were measured in the 12-month period prior to the month of the initial mental health visit in January 2013 or later.

Table A-VI.5. Codes used to identify mental health services (KMHS Medicaid)

Service category	CPT codes and additional requirements
1. Psychiatric visit	CPT-code = 90801 through 90899, 90791, 90792, and 90785 (psychiatric visit)
2. E&M visit with psych primary DX	CPT-code = any outpatient E&M visit (CPT=99201-99205, 99211-99215) with a MH, alcohol, or drug abuse primary diagnosis code listed in Table A.10
3. Psychiatric medication management visit	CPT-Code=M0064[a]
4. Additional psychiatric visit codes used by KMHS in Medicaid data	0143M, 0149M, 96153, 96154, H0001, H0002, H0031, H0036, H0040, H2011, H2012, H2015, H2022, S9484, S9485, T1001, T1005, T1017, T1023

[a] M0064 was deleted from the HCPCS system December 31, 2014. Thus, this code was in use through the end of the period we used to identify patients for this analysis.

Table A-VI.6. ICD-9 mental health diagnosis codes (KMHS Medicaid)

New variable name	Diagnosis group	ICD-9 Diagnosis code value
DX_SCHIZO	1. Schizophrenic disorders	295.xx including 295.00
DX_BIPOLAR	2. Bipolar disorders	296.00, 296.01, 296.02, 296.03, 296.04, 296.05,296.06, 296.10, 296.11, 296.12, 296.13, 296.14,296.15, 296.16, 296.40, 296.41, 296.42, 296.43,296.44, 296.45, 296.46, 296.50, 296.51, 296.52, 296.53, 296.54, 296.55, 296.56, 296.60, 296.61, 296.62, 296.63, 296.64, 296.65, 296.66, 296.7, 296.80, 296.81, 296.82, 296.89, 296.90, 296.99
DX_DEPRESS	3. Depressive disorders	296.20, 296.22, 296.23, 296.24, 296.25, 296.26, 296.30, 296.32, 296.33, 296.34, 296.35, 296.36, 311
DX_PMD	4. Persistent mental disorders due to conditions classified elsewhere	294.8x, 294.9x
DX_DEM	5. Dementia	290.xx, 294.1x
DX_OTPSY	6. Other psychotic disorders	297.xx-298.xx
DX_ANX	7. Anxiety, dissociative, and somatoform	300.xx
DX_ADJ	8. Adjustment reaction	309.xx
DX_DRUG	9. Drug and alcohol indicator	292, 292.0, 292.1, 292.2, 292.8, 292.9, 304, 304.0, 304.1, 304.2, 304.3, 304.4, 304.5, 304.6, 304.7, 304.8, 304.9, 305, 305.2, 305.3, 305.4, 305.5, 305.6, 305.7, 305.8, 305.9 291, 291.0, 291.1, 291.2, 291.3, 291.4, 291.5, 291.8, 291.9, 303, 303.0, 303.9, 305.0
DX_DRUG_ALT	10. Alternative definition for drug and alcohol indicator	292.xx, 304.xx, 305, 305.2-305.9

uses NDC codes for prescription drugs. We used these complementary systems to create flags for chronic diseases and their expected level of expenditures.

New variable name	Diagnosis group	ICD-9 Diagnosis code value
DX_OTDX	11. Other diagnosis not listed above	Everything not above (293.83, V62.84, V62.85, E950, E951, E952, E953, E954, E955, E956, E957, E958, E959, 301.0 to 301.9, 307.40 to 307.49, 312.0 to 312.23, 312.4 to 312.89, 313.81, 312.30 to 312.39, 302.0 to 302.9, 299.00 to 299.91, 307.1, 307.5, 307.51, 314.00 to 314.01, 307.20 to 307.3, 313.0 to 313.3, 313.82 to 316, 648.4, V65.2, V71.09, 780.09, V15.41, V15.42, V15,81, V17.0, V60.0, V62.29, V62.4, V62.81, V62.89) and All other codes in the range of 290.0-299.91 and 300.00-316

Table A-VI.7. ICD-9 primary care and substance abuse diagnosis codes (KMHS Medicaid)

Diagnosis	Diagnosis code (ICD-9)
Hypertension	362.11, 401.0, 401.1, 401.9, 402.00, 402.01, 402.10, 402.11, 402.90, 402.91, 403.00, 403.01, 403.10, 403.11, 403.90, 403.91, 404.00, 404.01, 404.02, 404.03, 404.10, 404.11, 404.12, 404.13, 404.90, 404.91, 404.92, 404.93, 405.01, 405.09, 405.11, 405.19, 405.91, 405.99, 437.2 (any DX on the claim)
Diabetes	249.00, 249.01, 249.10, 249.11, 249.20, 249.21, 249.30, 249.31, 249.40, 249.41, 249.50, 249.51, 249.60, 249.61, 249.70, 249.71, 249.80, 249.81, 249.90, 249.91,250.00, 250.01, 250.02, 250.03, 250.10, 250.11, 250.12, 250.13, 250.20, 250.21, 250.22, 250.23, 250.30, 250.31, 250.32, 250.33, 250.40, 250.41, 250.42,250.43, 250.50, 250.51, 250.52, 250.53, 250.60, 250.61, 250.62, 250.63, 250.70, 250.71, 250.72, 250.73, 250.80, 250.81, 250.82, 250.83, 250.90, 250.91, 250.92, 250.93, 357.2, 362.01, 362.02, 362.03, 362.04, 362.05, 362.06, 366.41 (any DX on the claim)
Drug Abuse	292.xx, 304.xx, 305, 305.2-305.9
Alcohol Abuse	291.xx, 303.xx, 305.0

To ensure consistency in Medicaid coverage and claims data availability across the analysis population, we required that during the full analysis period the individual (1) not be dually enrolled in Medicare, (2) have Medicaid as their primary payer, (3) not be a restricted-benefit enrollee, (4) not be an S-CHIP enrollee, (5) not have a missing enrollment record, and (6) have at least six months of Medicaid eligibility beginning in the month they first received mental health services during the intervention period through June 2014. Applying these restrictions resulted in the exclusion of 28 percent of the treatment group analysis population and 26 percent of the comparison pool analysis population, respectively.

When this step was complete the analysis population included 3,776 KMHS intervention participants and a comparison pool of 211,793 individuals who received mental health services from other facilities.

Step 2: Match treatment participants at individual level. The next step involved creating a matched comparison group. The matching process used metrics of individual-level characteristics identified based on pre-period Medicaid data to select a subset of comparison pool members who were as similar as possible to the intervention group on observable characteristics. The matching algorithm first exact matched on aid category, year began treatment at KMHS or another mental health facility, and a hierarchical variable of mental health diagnosis in the first three months of treatment including the following categories: schizophrenia; bipolar disorder; adjustment reaction disorder; depression; or anxiety, dissociative, or somatoform disorder. We were able to exactly match on the hierarchical diagnosis variable for all treatment group

members. Within the exact matching cells, we used propensity score matching,[108] to match up to five comparison pool beneficiaries to each intervention group member.

When a treatment beneficiary was difficult to match (that is, had few similar comparison beneficiaries in the same cell), the algorithm conducted a pairwise matching; when there were an abundance of comparisons for a treatment beneficiary, the algorithm matched multiple comparisons. The statistical goal is first to minimize bias and then, subject to that constraint, maximize the size of the comparison sample. The propensity score model included the following characteristics: age group (<18, 18–44, 45–54, 55–64, 65+), gender, disability status, whether the beneficiary was enrolled in Medicaid for a full 12 months prior to receiving mental health treatment during the intervention period, flags for psychiatric conditions,[109] and each of the CDPS diagnosis flags.

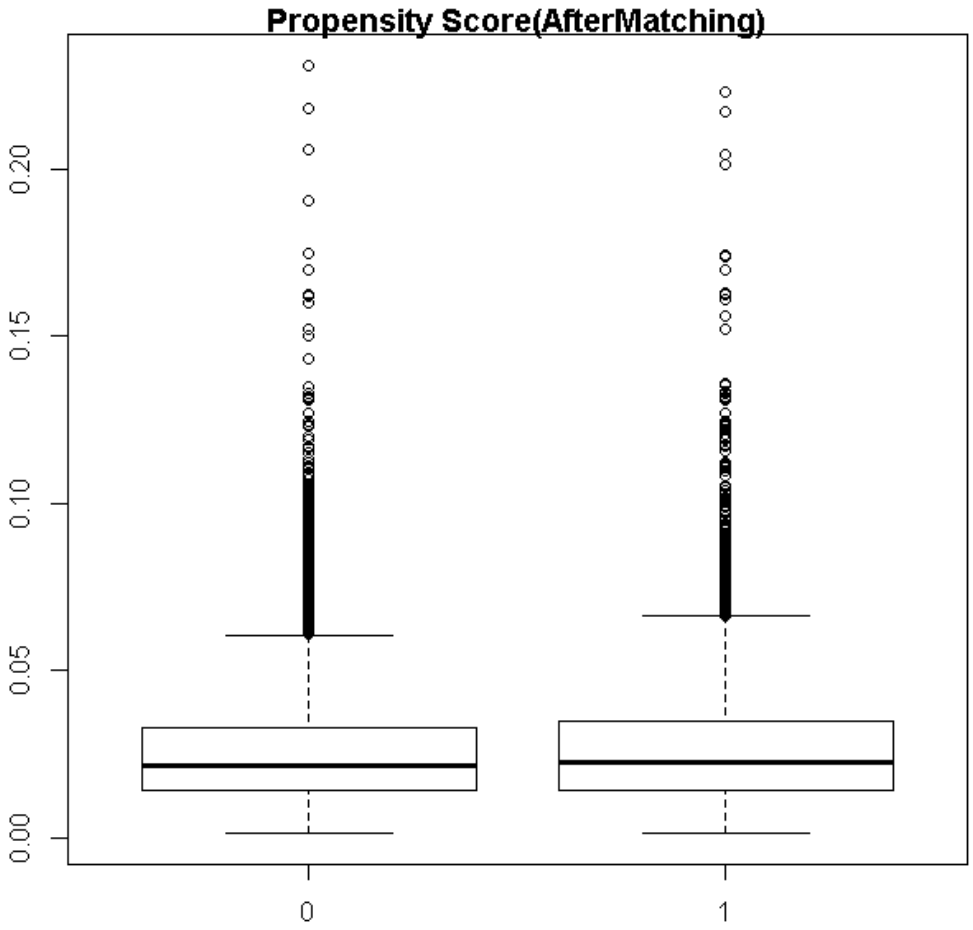
When this step was complete the analysis population included 3,776 KMHS intervention participants and 16,298 individuals in the comparison group. The reduction in the size of the comparison population relative to the previous step was due to individuals who were not matched to an intervention group member.

Step 3: Assess the quality of the match. The following tests and procedures were used to verify that the treatment and comparison groups were similar or balanced. First, we graphically compared the propensity score distributions for all treatment and comparison beneficiaries post-matching (Figure A-VI.2). This figure shows the propensity scores for the two groups looking very similar.

[108] A member's propensity score is the probability of belonging to the intervention group estimated based on a logistic regression model.

[109] We created flags to indicate that the patient had a diagnosis code for various conditions in the first three months of their claims during the intervention period. The diagnosis-related flags that we included in the matching included those for schizophrenia, bipolar disorder, depressive disorder, persistent mental disorder due to conditions classified elsewhere; dementia; "other" psychotic disorder; anxiety, dissociative, or somatoform disorder; adjustment reaction disorder; alcohol- or drug-related diagnosis; and "other" diagnosis.

Figure A-VI.2. Propensity score distributions (KMHS Medicaid)



Note: Figures present boxplots created using the estimated propensity scores for the comparison and intervention groups, the left and right panels, respectively. The line in the middle of each box represents the median score for the group. The lower and upper bounds of the box indicate the first and third quartile.

Next, we examined the number of comparison beneficiaries matched to each treatment beneficiary (Table A-VI.8). A large number of 1:1 matches, or a large number of comparison beneficiaries that were excluded, could indicate that the matching was problematic. We found five matches for the vast majority of participants (77 percent) indicating substantial overlap between the characteristics of the individuals in the comparison pool and intervention group.

Table A-VI.8. Frequency table of ratio of treatment beneficiaries to comparison beneficiaries for each matched set (KMHS)

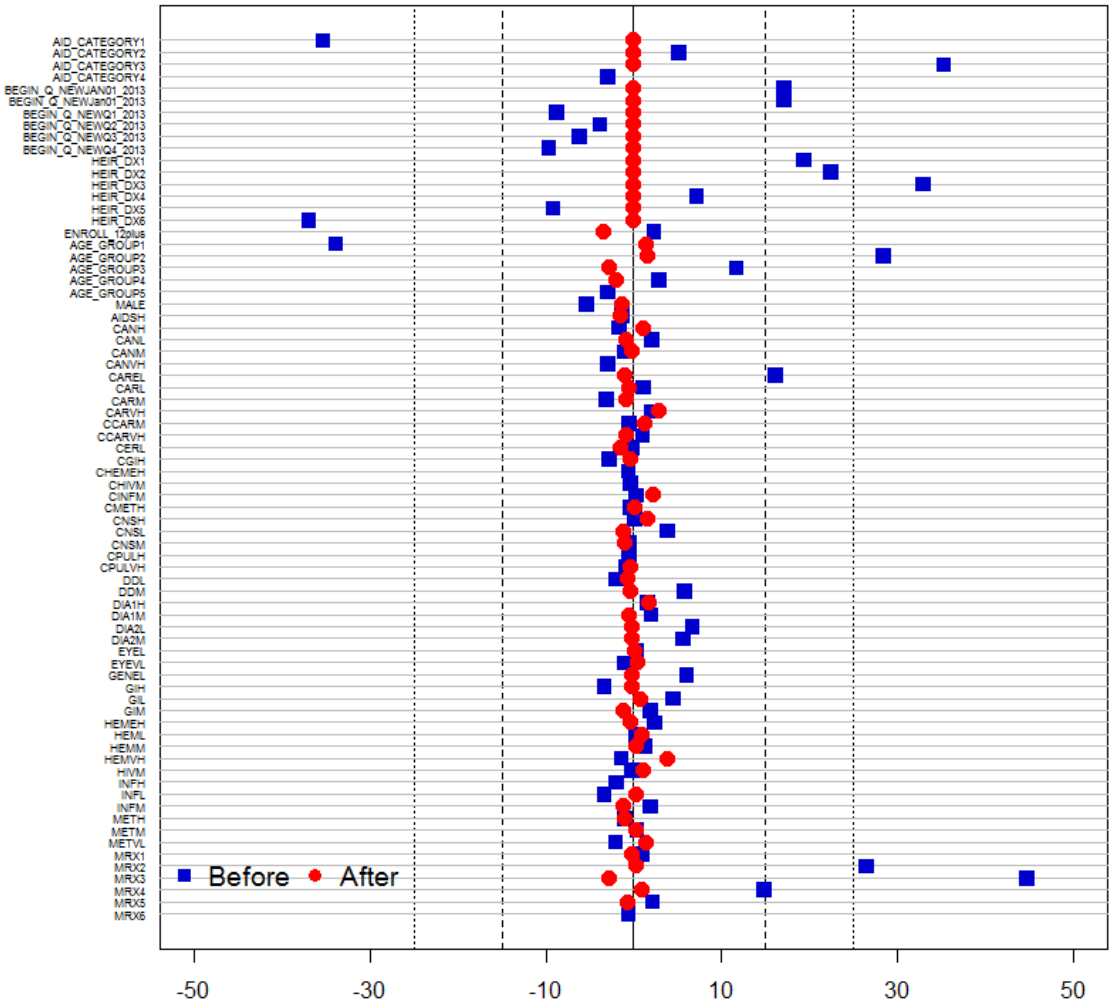
Ratio of treatment to comparison beneficiaries	1:1	1:2	1:3	1:4	1:5
Number of matched sets	425	148	149	140	2,914

Note: Each cell indicates the number of treatment beneficiaries matched to the number of comparison beneficiaries indicated for that column. In this example, most of the treatment beneficiaries (2,914) were matched to 5 comparison beneficiaries.

After evaluating the basic matching diagnostics above, we examined the overall balance of the matched sample. We used an omnibus test that checks for covariate balance across the individuals in the treatment and matched comparison group (Hansen and Bowers 2008). The omnibus test is based on the differences between the individuals in the treatment and matched group across the matching variables; these differences are standardized by their variances and covariances and aggregated into a single number, a weighted mean. Standardization in this way implies that a matching variable whose difference across matched sets has a small variance is given more weight and that a matching variable whose difference across sets is highly correlated with other differences is given less weight. The advantages of the omnibus test are: (1) it generates a single probability statement through one *p*-value; (2) its distribution is roughly chi-square, which facilities the calculation of the *p*-value; and (3) it assesses balance on all linear combinations of the matching variables. However, a significant result from this chi-square test may be driven by a large sample rather than substantive differences between treatment and matched comparison groups. Alternatively, it could indicate that there may be some imbalance between the two groups on at least one of the matching variables. The results of this test were a chi-square statistic of 182.6 and a *p*-value of < 0.01 indicating an imbalance exists.

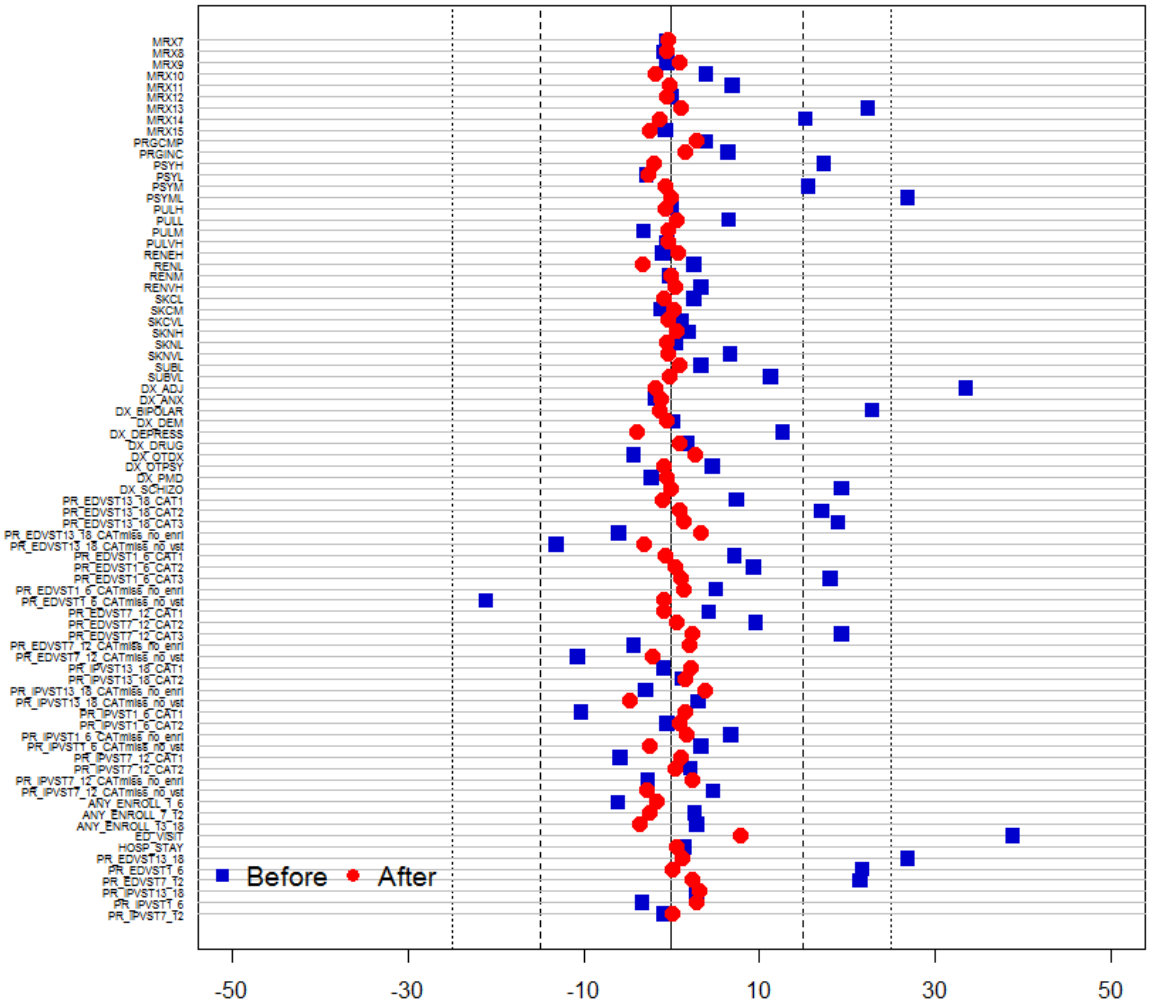
To further investigate imbalance between treatment and matched comparison groups, we evaluated how matching affected the balance on all matching variables (Figure A-VI.3) by comparing the absolute and standardized difference between the treatment and control groups for each variable before and after matching. The standardized difference measures the difference in means in *units* of the pooled standard deviation of treatment group and comparison group. The standardized difference measure is advantageous in that it allows us to compare all variables on the same scale. We compared the standardized differences using plots with dashed lines at +/- 0.15 standardized differences to visually inspect whether we obtained good balance for each variable, and using a balance table that shows both absolute and standardized differences between treatment and comparison groups before and after matching.

Figure A-VI.3.a Balance plot comparing the standardized difference for each matching variable before and after matching (KMHS), part 1



Note: See Table A-VI.9 for descriptions of the variables included in this figure.

Figure A-VI.3.b Balance plot comparing the standardized difference for each matching variable before and after matching (KMHS), part 2



Note: See Table A-VI.9 for descriptions of the variables included in this figure.

All variables were within 0.10 standard deviations indicating a strong balance. We provide more detail on the means and adjusted and standardized difference for the matching variables in Table A-VI.9 below.

Table A-VI.9. Balance table before and after matching (KMHS)

		Before matching					After matching				
	Variable description	Comparison	Treatment	adj.diff	std.diff	p	Comparison	Treatment	adj.diff	std.diff	p
AID_CATEGORY1	Aid category: child	0.5488	0.3729	-0.1759	-0.3537	0	0.3729	0.3729	0	0	1
AID_CATEGORY2	Aid category: adult	0.1784	0.1981	0.0197	0.0513	0.0018	0.1981	0.1981	0	0	1
AID_CATEGORY3	Aid category: disabled	0.2719	0.429	0.1571	0.3524	0	0.429	0.429	0	0	1
AID_CATEGORY4	Aid category: aged	0.0009	0	-0.0009	-0.0297	0.0705	0	0	0	#N/A	1
BEGIN_Q_NEWJAN01_2013	Quarter/year began treatment at KMHS or another mental health facility	0.6366	0.7185	0.0818	0.1703	0	0.7185	0.7185	0	0	1
BEGIN_Q_NEWJan01_2013	Quarter/year began treatment at KMHS or another mental health facility	0.6366	0.7185	0.0818	0.1703	0	0.7185	0.7185	0	0	1
BEGIN_Q_NEWQ1_2013	Quarter/year began treatment at KMHS or another mental health facility	0.1112	0.0837	-0.0275	-0.0876	0	0.0837	0.0837	0	0	1
BEGIN_Q_NEWQ2_2013	Quarter/year began treatment at KMHS or another mental health facility	0.0934	0.0821	-0.0113	-0.0389	0.0178	0.0821	0.0821	0	0	1
BEGIN_Q_NEWQ3_2013	Quarter/year began treatment at KMHS or another mental health facility	0.0785	0.0617	-0.0168	-0.0625	0.0001	0.0617	0.0617	0	0	1
BEGIN_Q_NEWQ4_2013	Quarter/year began treatment at KMHS or another mental health facility	0.0803	0.054	-0.0263	-0.0969	0	0.054	0.054	0	0	1
HEIR_DX1	Hierarchical variable of behavioral health diagnosis: schizophrenia	0.0276	0.0596	0.032	0.1938	0	0.0596	0.0596	0	0	1
HEIR_DX2	Hierarchical variable of behavioral health diagnosis: bipolar disorder	0.0653	0.121	0.0557	0.2242	0	0.121	0.121	0	0	1
HEIR_DX3	Hierarchical variable of behavioral health diagnosis: adjustment reaction disorder	0.0918	0.1875	0.0957	0.3293	0	0.1875	0.1875	0	0	1
HEIR_DX4	Hierarchical variable of behavioral health diagnosis: depression	0.1423	0.1674	0.025	0.0716	0	0.1674	0.1674	0	0	1
HEIR_DX5	Hierarchical variable of behavioral health diagnosis: anxiety, dissociative, or somatoform disorder	0.088	0.062	-0.026	-0.092	0	0.062	0.062	0	0	1
HEIR_DX6	Hierarchical variable of behavioral health diagnosis: any other	0.5851	0.4025	-0.1825	-0.3705	0	0.4025	0.4025	0	0	1

		Before matching					After matching				
	Variable description	Comparison	Treatment	adj.diff	std.diff	p	Comparison	Treatment	adj.diff	std.diff	p
ENROLL_12plus	Beneficiary was enrolled in Medicaid for a full 12 months prior to receiving mental health treatment during the intervention period	0.7047	0.715	0.0103	0.0226	0.1685	0.7306	0.715	-0.0155	-0.0348	0.0318
AGE_GROUP1	Age group <=17	0.5636	0.3954	-0.1682	-0.3393	0	0.3888	0.3954	0.0066	0.0137	0.0396
AGE_GROUP2	Age group 18-44	0.3001	0.4303	0.1303	0.2838	0	0.4226	0.4303	0.0078	0.0156	0.1655
AGE_GROUP3	Age group 45-54	0.0852	0.1178	0.0326	0.1164	0	0.1275	0.1178	-0.0096	-0.0287	0.0517
AGE_GROUP4	Age group 55-64	0.0502	0.0564	0.0062	0.0285	0.0824	0.0611	0.0564	-0.0047	-0.0193	0.1888
AGE_GROUP5	Age group 65+	0.0009	0	-0.0009	-0.0299	0.0689	0	0	0	#N/A	1
MALE	Gender	0.4652	0.4383	-0.0269	-0.0539	0.001	0.4454	0.4383	-0.0071	-0.0144	0.2836
AIDSH	Flag from CDPS+MRx algorithm	0.0026	0.0019	-0.0007	-0.0139	0.3962	0.0026	0.0019	-0.0008	-0.0148	0.4001
CANH	Flag from CDPS+MRx algorithm	0.003	0.0021	-0.0009	-0.0166	0.313	0.0016	0.0021	0.0005	0.0114	0.5071
CANL	Flag from CDPS+MRx algorithm	0.0038	0.005	0.0012	0.0203	0.2169	0.0057	0.005	-0.0007	-0.0091	0.6398
CANM	Flag from CDPS+MRx algorithm	0.0015	0.0011	-0.0004	-0.011	0.5019	0.0012	0.0011	-0.0001	-0.0029	0.8685
CANVH	Flag from CDPS+MRx algorithm	0.0009	0	-0.0009	-0.0295	0.0728	0	0	0	#N/A	1
CAREL	Flag from CDPS+MRx algorithm	0.0547	0.0916	0.0369	0.1614	0	0.0946	0.0916	-0.003	-0.0103	0.5647
CARL	Flag from CDPS+MRx algorithm	0.0344	0.0363	0.0019	0.0105	0.5226	0.0373	0.0363	-0.001	-0.0051	0.776
CARM	Flag from CDPS+MRx algorithm	0.0113	0.0079	-0.0034	-0.0318	0.0526	0.0087	0.0079	-0.0008	-0.0086	0.6231
CARVH	Flag from CDPS+MRx algorithm	0.0014	0.0021	0.0008	0.0206	0.2086	0.0011	0.0021	0.001	0.0282	0.1618
CCARM	Flag from CDPS+MRx algorithm	0.0004	0.0003	-0.0001	-0.0051	0.7544	0.0001	0.0003	0.0001	0.0119	0.5109
CCARVH	Flag from CDPS+MRx algorithm	0.0001	0.0003	0.0001	0.0098	0.5514	0.0004	0.0003	-0.0001	-0.0084	0.7737
CERL	Flag from CDPS+MRx algorithm	0.006	0.0058	-0.0002	-0.002	0.9053	0.007	0.0058	-0.0012	-0.0152	0.4116
CGIH	Flag from CDPS+MRx algorithm	0.0013	0.0003	-0.001	-0.0282	0.086	0.0003	0.0003	-0.0001	-0.0035	0.8309
CHEMEH	Flag from CDPS+MRx algorithm	0	0	0	-0.0062	0.7061	0	0	0	#N/A	1
CHIVM	Flag from CDPS+MRx algorithm	0	0	0	-0.0038	0.8174	0	0	0	#N/A	1
CINFM	Flag from CDPS+MRx algorithm	0.0002	0.0003	0	0.0026	0.8748	0.0001	0.0003	0.0002	0.0212	0.4652
CMETH	Flag from CDPS+MRx algorithm	0.0006	0.0005	-0.0001	-0.0035	0.8311	0.0005	0.0005	0	0.0017	0.9252
CNSH	Flag from CDPS+MRx algorithm	0.0018	0.0019	0	0.001	0.9521	0.0013	0.0019	0.0006	0.0155	0.3938
CNSL	Flag from CDPS+MRx algorithm	0.0486	0.0569	0.0083	0.0385	0.019	0.0599	0.0569	-0.003	-0.0126	0.4885
CNSM	Flag from CDPS+MRx algorithm	0.0079	0.0074	-0.0005	-0.0053	0.7474	0.0083	0.0074	-0.0009	-0.01	0.5753
CPULH	Flag from CDPS+MRx algorithm	0	0	0	-0.0058	0.7243	0	0	0	#N/A	1
CPULVH	Flag from CDPS+MRx algorithm	0.0004	0.0003	-0.0002	-0.0083	0.6135	0.0003	0.0003	-0.0001	-0.0041	0.8208
DDL	Flag from CDPS+MRx algorithm	0.0064	0.0048	-0.0016	-0.0203	0.217	0.0053	0.0048	-0.0006	-0.0074	0.6578
DDM	Flag from CDPS+MRx algorithm	0.0004	0.0016	0.0012	0.0574	0.0005	0.0017	0.0016	-0.0001	-0.0046	0.8695
DIA1H	Flag from CDPS+MRx algorithm	0.0005	0.0008	0.0003	0.0151	0.3566	0.0004	0.0008	0.0004	0.0178	0.3063
DIA1M	Flag from CDPS+MRx algorithm	0.0054	0.0069	0.0015	0.0197	0.229	0.0074	0.0069	-0.0005	-0.0058	0.7546
DIA2L	Flag from CDPS+MRx algorithm	0.0284	0.0395	0.011	0.0662	0.0001	0.0398	0.0395	-0.0003	-0.0017	0.9268
DIA2M	Flag from CDPS+MRx algorithm	0.0078	0.0127	0.0049	0.056	0.0006	0.0129	0.0127	-0.0002	-0.002	0.9124
EYEL	Flag from CDPS+MRx algorithm	0.0015	0.0016	0.0001	0.0035	0.8295	0.0015	0.0016	0.0001	0.0015	0.9345
EYEVL	Flag from CDPS+MRx algorithm	0.0115	0.0103	-0.0012	-0.0109	0.5056	0.01	0.0103	0.0004	0.0035	0.845
GENEL	Flag from CDPS+MRx algorithm	0.0241	0.0334	0.0093	0.0602	0.0002	0.0338	0.0334	-0.0004	-0.0024	0.9001
GIH	Flag from CDPS+MRx algorithm	0.0043	0.0021	-0.0022	-0.0339	0.0388	0.0022	0.0021	-0.0001	-0.0026	0.8869
GIL	Flag from CDPS+MRx algorithm	0.0721	0.0837	0.0115	0.0446	0.0066	0.0817	0.0837	0.0019	0.0071	0.7007
GIM	Flag from CDPS+MRx algorithm	0.014	0.0162	0.0022	0.0184	0.2636	0.0177	0.0162	-0.0015	-0.0117	0.5167
HEMEH	Flag from CDPS+MRx algorithm	0.0002	0.0005	0.0003	0.0233	0.1562	0.0006	0.0005	-0.0001	-0.0034	0.9136
HEML	Flag from CDPS+MRx algorithm	0.0067	0.0069	0.0002	0.0024	0.8822	0.0062	0.0069	0.0007	0.0087	0.6246

		Before matching					After matching				
	Variable description	Comparison	Treatment	adj.diff	std.diff	p	Comparison	Treatment	adj.diff	std.diff	p
HEMM	Flag from CDPS+MRx algorithm	0.0039	0.0048	0.0008	0.013	0.4274	0.0046	0.0048	0.0002	0.0023	0.9074
HEMVH	Flag from CDPS+MRx algorithm	0.0006	0.0003	-0.0004	-0.0142	0.3858	0	0.0003	0.0003	0.0375	0.0253
HIVM	Flag from CDPS+MRx algorithm	0.0006	0.0005	-0.0001	-0.0024	0.8829	0.0003	0.0005	0.0002	0.0106	0.5334
INFH	Flag from CDPS+MRx algorithm	0.0004	0	-0.0004	-0.0201	0.2217	0	0	0	#N/A	1
INFL	Flag from CDPS+MRx algorithm	0.0298	0.0241	-0.0057	-0.0337	0.04	0.0238	0.0241	0.0003	0.0021	0.9041
INFM	Flag from CDPS+MRx algorithm	0.0067	0.0082	0.0015	0.0183	0.2644	0.0093	0.0082	-0.0011	-0.0115	0.5462
METH	Flag from CDPS+MRx algorithm	0.0061	0.0053	-0.0008	-0.0103	0.5297	0.006	0.0053	-0.0008	-0.0097	0.5921
METM	Flag from CDPS+MRx algorithm	0.0095	0.0098	0.0003	0.0035	0.8332	0.0095	0.0098	0.0003	0.0033	0.8602
METVL	Flag from CDPS+MRx algorithm	0.0322	0.0286	-0.0036	-0.0206	0.209	0.0263	0.0286	0.0023	0.0142	0.4407
MRX1	Flag from CDPS+MRx algorithm	0.0169	0.018	0.0011	0.0089	0.5877	0.0182	0.018	-0.0002	-0.0016	0.9268
MRX2	Flag from CDPS+MRx algorithm	0.2327	0.3448	0.1121	0.2647	0	0.3437	0.3448	0.0011	0.0023	0.8881
MRX3	Flag from CDPS+MRx algorithm	0.4429	0.665	0.222	0.4474	0	0.6781	0.665	-0.0131	-0.0279	0.008
MRX4	Flag from CDPS+MRx algorithm	0.0532	0.0866	0.0334	0.1478	0	0.0843	0.0866	0.0023	0.0085	0.6409
MRX5	Flag from CDPS+MRx algorithm	0.0004	0.0008	0.0004	0.0213	0.1936	0.001	0.0008	-0.0002	-0.0076	0.7732
MRX6	Flag from CDPS+MRx algorithm	0	0	0	-0.0066	0.6892	0	0	0	N/A	1
MRX7	Flag from CDPS+MRx algorithm	0.0024	0.0021	-0.0003	-0.0062	0.7053	0.0023	0.0021	-0.0002	-0.0041	0.8151
MRX8	Flag from CDPS+MRx algorithm	0.0034	0.0029	-0.0005	-0.0087	0.5956	0.0032	0.0029	-0.0003	-0.005	0.777
MRX9	Flag from CDPS+MRx algorithm	0.0038	0.0034	-0.0003	-0.0051	0.7554	0.003	0.0034	0.0005	0.0083	0.6456
MRX10	Flag from CDPS+MRx algorithm	0.005	0.0077	0.0027	0.0387	0.0184	0.0093	0.0077	-0.0016	-0.0179	0.3569
MRX11	Flag from CDPS+MRx algorithm	0.009	0.0156	0.0066	0.0692	0	0.0158	0.0156	-0.0002	-0.0018	0.9319
MRX12	Flag from CDPS+MRx algorithm	0.0011	0.0011	0	-0.0006	0.9728	0.0013	0.0011	-0.0002	-0.0057	0.7403
MRX13	Flag from CDPS+MRx algorithm	0.0319	0.0715	0.0396	0.2233	0	0.0689	0.0715	0.0026	0.0106	0.5412
MRX14	Flag from CDPS+MRx algorithm	0.0772	0.1181	0.0409	0.1525	0	0.1226	0.1181	-0.0045	-0.0137	0.4476
MRX15	Flag from CDPS+MRx algorithm	0.0039	0.0034	-0.0004	-0.0069	0.6727	0.005	0.0034	-0.0016	-0.0241	0.1992
PRGCMP	Flag from CDPS+MRx algorithm	0.0381	0.0456	0.0075	0.0389	0.0179	0.04	0.0456	0.0056	0.0277	0.1048
PRGINC	Flag from CDPS+MRx algorithm	0.0142	0.0217	0.0076	0.0637	0.0001	0.0195	0.0217	0.0022	0.0162	0.3752
PSYH	Flag from CDPS+MRx algorithm	0.0348	0.0667	0.0319	0.173	0	0.0718	0.0667	-0.005	-0.0194	0.0892
PSYL	Flag from CDPS+MRx algorithm	0.0866	0.0784	-0.0082	-0.0292	0.0756	0.0859	0.0784	-0.0075	-0.0266	0.0985
PSYM	Flag from CDPS+MRx algorithm	0.0626	0.1004	0.0378	0.1553	0	0.1024	0.1004	-0.002	-0.0067	0.6993
PSYML	Flag from CDPS+MRx algorithm	0.3094	0.4338	0.1244	0.2688	0	0.4337	0.4338	0.0001	0.0002	0.9907
PULH	Flag from CDPS+MRx algorithm	0.0014	0.0013	0	-0.0007	0.9649	0.0016	0.0013	-0.0003	-0.0066	0.7351
PULL	Flag from CDPS+MRx algorithm	0.1147	0.1353	0.0206	0.0647	0.0001	0.1336	0.1353	0.0018	0.0052	0.7785
PULM	Flag from CDPS+MRx algorithm	0.0135	0.0098	-0.0037	-0.0324	0.0486	0.0101	0.0098	-0.0003	-0.0032	0.8573
PULVH	Flag from CDPS+MRx algorithm	0.0015	0.0013	-0.0002	-0.0051	0.7539	0.0015	0.0013	-0.0002	-0.0044	0.8052
RENEH	Flag from CDPS+MRx algorithm	0.0005	0.0003	-0.0002	-0.0107	0.5152	0.0002	0.0003	0.0001	0.0075	0.7518
RENL	Flag from CDPS+MRx algorithm	0.0328	0.0373	0.0045	0.0255	0.1202	0.044	0.0373	-0.0067	-0.0337	0.0708
RENM	Flag from CDPS+MRx algorithm	0.0012	0.0011	-0.0001	-0.0029	0.86	0.0011	0.0011	0	0	1
RENVH	Flag from CDPS+MRx algorithm	0.0053	0.0077	0.0024	0.033	0.0445	0.0073	0.0077	0.0003	0.0039	0.8344
SKCL	Flag from CDPS+MRx algorithm	0.0424	0.0474	0.005	0.025	0.1271	0.0494	0.0474	-0.002	-0.0094	0.6072
SKCM	Flag from CDPS+MRx algorithm	0.0251	0.0233	-0.0018	-0.0117	0.475	0.0229	0.0233	0.0004	0.0025	0.8864
SKCVL	Flag from CDPS+MRx algorithm	0.0502	0.0527	0.0025	0.0117	0.4771	0.0536	0.0527	-0.0009	-0.0039	0.8289
SKNH	Flag from CDPS+MRx algorithm	0.002	0.0029	0.0009	0.0197	0.2297	0.0026	0.0029	0.0003	0.0061	0.7529
SKNL	Flag from CDPS+MRx algorithm	0.004	0.0042	0.0003	0.004	0.8082	0.0046	0.0042	-0.0004	-0.0055	0.7651
SKNVL	Flag from CDPS+MRx algorithm	0.0691	0.0861	0.0169	0.0667	0	0.0872	0.0861	-0.0011	-0.0041	0.8247
SUBL	Flag from CDPS+MRx algorithm	0.0757	0.0845	0.0088	0.0332	0.0433	0.0821	0.0845	0.0024	0.0086	0.6221
SUBVL	Flag from CDPS+MRx algorithm	0.0366	0.0577	0.0212	0.1122	0	0.0581	0.0577	-0.0004	-0.0018	0.9246

		Before matching					After matching				
	Variable description	Comparison	Treatment	adj.diff	std.diff	p	Comparison	Treatment	adj.diff	std.diff	p
DX_ADJ	Diagnosis of adjustment reaction disorder	0.1	0.201	0.1011	0.3346	0	0.2083	0.201	-0.0073	-0.0182	0.0006
DX_ANX	Diagnosis of anxiety, dissociative, or somatoform disorder	0.1514	0.1446	-0.0068	-0.0191	0.245	0.1491	0.1446	-0.0045	-0.0124	0.3464
DX_BIPOLAR	Diagnosis of bipolar disorder	0.0695	0.1279	0.0584	0.2282	0	0.1324	0.1279	-0.0045	-0.0131	0.0077
DX_DEM	Diagnosis of dementia	0.0008	0.0008	0	0.0006	0.9686	0.0009	0.0008	-0.0001	-0.0047	0.805
DX_DEPRESS	Diagnosis of depressive disorder	0.1747	0.2227	0.048	0.1263	0	0.2395	0.2227	-0.0168	-0.0393	0.0001
DX_DRUG	Drug and/or alcohol-related diagnosis	0.0044	0.0056	0.0011	0.0169	0.3042	0.0049	0.0056	0.0006	0.0087	0.6415
DX_OTDX	Other behavioral health diagnosis	0.314	0.2937	-0.0203	-0.0438	0.0077	0.2814	0.2937	0.0123	0.0275	0.1031
DX_OTPSY	Diagnosis of other psychotic disorder	0.0191	0.0254	0.0063	0.0462	0.0049	0.0268	0.0254	-0.0014	-0.0085	0.6157
DX_PMD	Diagnosis of persistent mental disorders due to conditions classified elsewhere	0.0018	0.0008	-0.001	-0.023	0.1616	0.001	0.0008	-0.0002	-0.0055	0.7582
DX_SCHIZO	Diagnosis of schizophrenic disorder	0.0276	0.0596	0.032	0.1938	0	0.0596	0.0596	0	0	1
PR_EDVST13_18_CAT1	Pro-rated version of the ED counts measure for visits 13-18 months prior: Category 1	0.1271	0.1517	0.0247	0.074	0	0.1556	0.1517	-0.0038	-0.0106	0.5571
PR_EDVST13_18_CAT2	Pro-rated version of the ED counts measure for visits 13-18 months prior: Category 2	0.0587	0.099	0.0403	0.1705	0	0.0965	0.099	0.0025	0.0089	0.6304
PR_EDVST13_18_CAT3	Pro-rated version of the ED counts measure for visits 13-18 months prior: Category 3	0.022	0.0501	0.0281	0.1895	0	0.0474	0.0501	0.0027	0.0138	0.4595
PR_EDVST13_18_CATmiss_no_enrl	Pro-rated version of the ED counts measure for visits 13-18 months prior: No enrollment category	0.29	0.2627	-0.0273	-0.0601	0.0003	0.2484	0.2627	0.0144	0.0328	0.0419
PR_EDVST13_18_CATmiss_no_vst	Pro-rated version of the ED counts measure for visits 13-18 months prior: No visit category	0.5022	0.4364	-0.0658	-0.1316	0	0.4522	0.4364	-0.0157	-0.0316	0.0453
PR_EDVST1_6_CAT1	Pro-rated version of the ED counts measure for visits 1-6 months prior: Category 1	0.1736	0.2007	0.0271	0.0715	0	0.2038	0.2007	-0.0031	-0.0077	0.6727
PR_EDVST1_6_CAT2	Pro-rated version of the ED counts measure for visits 1-6 months prior: Category 2	0.0567	0.0784	0.0217	0.0935	0	0.0772	0.0784	0.0012	0.0047	0.804
PR_EDVST1_6_CAT3	Pro-rated version of the ED counts measure for visits 1-6 months prior: Category 3	0.0515	0.0916	0.0402	0.1806	0	0.0889	0.0916	0.0028	0.0102	0.5891

		Before matching					After matching				
	Variable description	Comparison	Treatment	adj.diff	std.diff	p	Comparison	Treatment	adj.diff	std.diff	p
PR_EDVST1_6_CATmiss_no_enrl	Pro-rated version of the ED counts measure for visits 1-6 months prior: No enrollment category	0.0444	0.0548	0.0104	0.0504	0.0022	0.0516	0.0548	0.0032	0.0143	0.3766
PR_EDVST1_6_CATmiss_no_vst	Pro-rated version of the ED counts measure for visits 1-6 months prior: No visit category	0.6738	0.5744	-0.0993	-0.2117	0	0.5785	0.5744	-0.0041	-0.0083	0.6141
PR_EDVST7_12_CAT1	Pro-rated version of the ED counts measure for visits 7-12 months prior: Category 1	0.1461	0.161	0.0149	0.0423	0.01	0.1644	0.161	-0.0033	-0.0091	0.6192
PR_EDVST7_12_CAT2	Pro-rated version of the ED counts measure for visits 7-12 months prior: Category 2	0.0694	0.0937	0.0243	0.0954	0	0.0919	0.0938	0.0019	0.0065	0.7228
PR_EDVST7_12_CAT3	Pro-rated version of the ED counts measure for visits 7-12 months prior: Category 3	0.0254	0.0561	0.0308	0.1937	0	0.0513	0.0561	0.0049	0.0239	0.2149
PR_EDVST7_12_CATmiss_no_enrl	Pro-rated version of the ED counts measure for visits 7-12 months prior: No enrollment category	0.1884	0.1713	-0.017	-0.0436	0.008	0.1636	0.1713	0.0078	0.0207	0.1958
PR_EDVST7_12_CATmiss_no_vst	Pro-rated version of the ED counts measure for visits 7-12 months prior: No visit category	0.5707	0.5177	-0.053	-0.1071	0	0.5289	0.5177	-0.0112	-0.0225	0.1836
PR_IPVST13_18_CAT1	Pro-rated version of the hospitalization counts measure for visits 13-18 months prior: Category 1	0.0324	0.0307	-0.0017	-0.0096	0.5576	0.027	0.0307	0.0037	0.0223	0.2191
PR_IPVST13_18_CAT2	Pro-rated version of the hospitalization counts measure for visits 13-18 months prior: Category 2	0.0092	0.0103	0.0012	0.0121	0.4605	0.0088	0.0103	0.0015	0.0161	0.3856
PR_IPVST13_18_CATmiss_no_enrl	Pro-rated version of the hospitalization counts measure for visits 13-18 months prior: No enrollment category	0.307	0.2932	-0.0139	-0.0301	0.0667	0.2762	0.2932	0.017	0.0376	0.0211
PR_IPVST13_18_CATmiss_no_vst	Pro-rated version of the hospitalization counts measure for visits 13-18 months prior: No visit category	0.6514	0.6658	0.0144	0.0303	0.0651	0.688	0.6658	-0.0222	-0.0476	0.0037
PR_IPVST1_6_CAT1	Pro-rated version of the hospitalization counts measure for visits 1-6 months prior: Category 1	0.0603	0.0358	-0.0245	-0.1034	0	0.0328	0.0358	0.0029	0.0158	0.3627
PR_IPVST1_6_CAT2	Pro-rated version of the hospitalization counts measure for visits 1-6 months prior: Category 2	0.0126	0.0119	-0.0006	-0.0058	0.7249	0.011	0.0119	0.001	0.0091	0.6125

		Before matching					After matching				
	Variable description	Comparison	Treatment	adj.diff	std.diff	p	Comparison	Treatment	adj.diff	std.diff	p
PR_IPVST1_6_CATmiss_no_enrl	Pro-rated version of the hospitalization counts measure for visits 1-6 months prior: No enrollment category	0.0479	0.0622	0.0143	0.0668	0	0.0583	0.0622	0.0039	0.0167	0.3027
PR_IPVST1_6_CATmiss_no_vst	Pro-rated version of the hospitalization counts measure for visits 1-6 months prior: No visit category	0.8792	0.8901	0.0108	0.0333	0.0424	0.8979	0.8901	-0.0078	-0.0255	0.1278
PR_IPVST7_12_CAT1	Pro-rated version of the hospitalization counts measure for visits 7-12 months prior: Category 1	0.0407	0.0291	-0.0116	-0.0587	0.0004	0.0273	0.0291	0.0019	0.0112	0.5136
PR_IPVST7_12_CAT2	Pro-rated version of the hospitalization counts measure for visits 7-12 months prior: Category 2	0.0103	0.0124	0.0021	0.0211	0.1997	0.0121	0.0124	0.0004	0.0034	0.8489
PR_IPVST7_12_CATmiss_no_enrl	Pro-rated version of the hospitalization counts measure for visits 7-12 months prior: No enrollment category	0.1979	0.187	-0.011	-0.0275	0.0936	0.1778	0.187	0.0092	0.0238	0.1408
PR_IPVST7_12_CATmiss_no_vst	Pro-rated version of the hospitalization counts measure for visits 7-12 months prior: No visit category	0.7511	0.7715	0.0204	0.0472	0.0041	0.7829	0.7715	-0.0115	-0.0275	0.0937
ANY_ENROLL_1_6	Number of months enrolled 1-6 months prior	0.9503	0.937	-0.0133	-0.061	0.0002	0.9408	0.937	-0.0039	-0.0163	0.3172
ANY_ENROLL_7_12	Number of months enrolled 7-12 months prior	0.7982	0.8088	0.0106	0.0264	0.1084	0.8185	0.8088	-0.0097	-0.0249	0.1237
ANY_ENROLL_13_18	Number of months enrolled 13-18 months prior	0.6874	0.7005	0.0131	0.0282	0.0861	0.717	0.7005	-0.0165	-0.0363	0.0259
ED_VISIT	Number of ED visits in 12 month claims/encounter history	3.438	6.0217	2.5837	0.3883	0	5.3618	6.0217	0.66	0.079	0
HOSP_STAY	Number of hospitalizations in 12 month claims/encounter history	0.4341	0.4531	0.0191	0.015	0.3604	0.4447	0.4531	0.0084	0.0061	0.757
PR_EDVST13_18	Continuous pro-rated version of the ED visits counts measure for visits 13-18 months prior	0.3932	0.7341	0.3408	0.2689	0	0.7133	0.7341	0.0207	0.0124	0.5315
PR_EDVST1_6	Continuous pro-rated version of the ED visits counts measure for visits 1-6 months prior	0.5171	0.8107	0.2936	0.2166	0	0.8088	0.8107	0.0019	0.0011	0.9558
PR_EDVST7_12	Continuous pro-rated version of the ED visits counts measure for visits 7-12 months prior	0.4545	0.7407	0.2862	0.2142	0	0.703	0.7407	0.0377	0.0233	0.1998
PR_IPVST13_18	Continuous pro-rated version of the hospitalization counts measure for visits 13-18 months prior	0.049	0.0573	0.0083	0.0284	0.0837	0.0477	0.0573	0.0096	0.0314	0.1035

		Before matching					After matching				
	Variable description	Comparison	Treatment	adj.diff	std.diff	p	Comparison	Treatment	adj.diff	std.diff	p
PR_IPVST1_6	Continuous pro-rated version of the hospitalization counts measure for visits 1-6 months prior	0.0792	0.0674	-0.0119	-0.0335	0.041	0.058	0.0674	0.0094	0.0277	0.1581
PR_IPVST7_12	Continuous pro-rated version of the hospitalization counts measure for visits 7-12 months prior	0.0593	0.0565	-0.0028	-0.0089	0.5893	0.0562	0.0565	0.0003	0.0009	0.9608

Note: We created CDPS flags following UCSD's CDPS + MRx methodology. Please see the CDPS website for programs for further information regarding the individual flags.

Step 5: Create analysis weights. Weights were developed for each member of the analysis population. Weights for KMHS attributed individuals were set to one. Weights for comparison group members were set to one divided by the number of comparison group members assigned to the member's associated treatment person. An individual's participation in the analysis could be terminated as a result of a change in status before the end of the analysis period. An individual's weight was set to zero in analysis months following any of these status changes. There were two status changes for which individuals were dropped from the analysis: (1) individuals who were no longer enrolled in Medicaid were dropped from our analysis because they were no longer included in the data available for analysis; and (2) individuals were dropped from our analysis if they died.

D. Specifications of measures

We used multiple types of measures in these analyses. CMMI requested that we calculate four standardized outcome measures for all awardees to the extent feasible. These measures are: total Medicare and/or Medicaid expenditures, inpatient hospitalizations, hospital readmissions, and ED visits. If it was possible to calculate the core measures identified by CMS and these measures were appropriate to the intervention, we used them.[110] Our specifications for these measures in Medicare and Medicaid administrative data are described in the first section below (Section D.1). For KMHS, we used multivariate regression models to adjust for differences across the analysis population in demographics, geography, socioeconomics, Medicaid/Medicare enrollment, and health status. We describe the specifications for the control variables in these models (Section D.2).

1. Core measures in Medicare and Medicaid administrative data

In this section, we provide detail on the data and analytic methods used to develop the core outcome measures. We begin by describing how we identified the spans of Medicare or Medicaid enrollment that were included in the analyses for each intervention or comparison group member. Then, we describe how we processed claims data and assigned expenditure and utilization information to months to develop each of the core measures. Finally, we discuss how we annualized and weighted the regressions models to adjust for individuals who were not observable for a full 12 months.

a. Identifying periods with observable data

In this section we describe the approach we used to identify the patients and periods of Medicare or Medicaid enrollment included in the analysis.

Define intervention start date. We assigned each intervention and comparison group member identified in Section C above an intervention start month. For the Medicare analysis individuals were assigned to the treatment facility at which they were first observed to receive mental health treatment. The Race to Health! program began on January 1, 2013. For the Medicare population individuals already in care at a mental health facility prior to this month had

[110] For the Medicare and Medicaid analyses we did not analyze readmissions due to the small number of patients with readmissions. For the Medicaid analysis we did not analyze total expenditures because expenditure information was not available for most Medicaid enrollees who were enrolled in a managed care plan.

January 2013 assigned as their intervention start month. Individuals who initiated care at KMHS or a comparison mental health facility during or after January 2013 had the first month they received mental health treatment at their assigned facility assigned as their intervention start month. Because individuals could not be assigned to a facility in the Medicaid analysis, we assigned intervention start month as the first month after December 2012 when the intervention or comparison group member was observed to receive a mental health treatment service in the Medicaid claims data.

Define baseline and intervention periods. Baseline and intervention periods were defined for each intervention participant or comparison group member relative to their intervention start month. The first intervention period was defined as the intervention start month and the appropriate number of months following that month.[111] Where applicable the second intervention period was defined starting in the months following the last month in the first intervention period. The first baseline period started in the month prior to the intervention start month and moved backward. For each individual included in the analysis the proportion of each baseline and intervention period for which the individual was eligible for the analysis was calculated. This proportion was used to pro-rate the expenditure and utilization measures for individuals enrolled for less than the full analysis period. It was also used to weight observations in the regression analysis.

b. Summarizing monthly expenditures and utilization

Once the individuals and periods eligible for the analysis were identified as described above, expenditures and utilization associated with each core measure were aggregated for the periods during which the individual was deemed eligible for the analysis. In this section, we define the specifications for identifying total Medicare expenditures, hospitalizations, and ED visits, and total Medicaid hospitalizations and ED visits. We summarized each of these measures monthly for each individual in the analysis population. Then, we aggregated sets of months to the analysis period.

i. Expenditures

For Medicare, the following claim types were included in this analysis: carrier, DME, home health, hospice, inpatient, outpatient, and skilled nursing facility. Only FFS data were included in this analysis. Part D services were excluded. Duplicate and denied claims were excluded. The total payment amount on each Medicare claim was summed across all file types to calculate total expenditures. For services that extend beyond a single day (for example, an inpatient or long-term care stay) we counted all Medicare payments recorded based on a single date. Inpatient stays expenditures were counted in the month of the discharge date. For other types of claims all expenditures were assigned based on the claim from date. Expenditures were excluded from this analysis if they were assigned to a month during which the associated Medicare beneficiary was deemed ineligible for the analysis.

[111] This was five months for the descriptive analysis of outcomes and up to 29 months for the impact analysis.

ii. Hospitalizations

The specifications for the hospitalization measures were developed to align with the CMMI priority all-cause admissions per patient measure. For this measure only acute stays or psychiatric stays were included in the analysis. We describe the steps to develop these counts here.

Step 1: Identify hospitalization claims. For Medicare administrative data inpatient hospital claims were identified by claim type. Then, we identified and excluded rehabilitation and long-term care stays from the Medicare data based on provider identifier codes. At the end of this step, only acute and psychiatric stays were included in the file. For Medicaid MAX and Alpha-MAX, we analyzed all claims in the IP file.[112]

Step 2: Eliminate duplicate or denied claims. For Medicare, we identified claims with the same information in all fields and only kept one of these claims. We also excluded denied claims from our analysis. For MAX and Alpha-MAX the data files included final paid claims, so no additional adjustment was necessary.

Step 3: Combine claims that represent the same stay and combine transfer stays with initial stays. For all data types, we identified and combined initial and interim claims into one discharge. Interim claims had (1) the same admission date as the initial claim, (2) an admission date that was equal to the discharge date from the initial or another interim claim and the status on the other (previous) claim was "still a patient", or (3) a claim with an admission date that was equal to one day after the discharge date of the initial or another interim claim and the status on the other previous claim was "still a patient." Such claims were combined to count as a single stay.

Next, we identified and combined claims associated with a transfer into a single stay. We identified claims indicating that the patient was transferred to either another short-term hospital, a CAH, another type institution for inpatient care, a federal hospital, or a psychiatric hospital or unit. Then we combined these claims with claims for the same beneficiary at a different facility where the admission date fell within one day of the discharge date of the first claim.

Step 4: Sum the number of discharges in each month. Once claims representing a single stay were combined, we summed the number of unique discharges for each enrollee for each month. Inpatient stays were counted in the month of the discharge date.

iii. ED visits

Outpatient ED visit utilization is reflected in CMMI priority measure 62. This measure includes ED visits and observation stays that do not lead to an admission.

In the Medicare outpatient file, we identified outpatient ED claims as those with a revenue center value indicating an ED visit, excluding any claims that involved only lab or imaging

[112] LT file claims were not included in this analysis. Psychiatric hospital services may be reported in the LT file. We will assess reporting and update to include psychiatric hospitalization services excluded from the IP file in the addendum to the current report.

services in the ED. We identified observation claims based on the combination of revenue center code, CPT-code and a unit count of greater than or equal to eight hours.

In addition to the codes identified above, for Medicaid data, we reviewed claims not identified as inpatient and considered them as ED visits if the procedure code indicated ED visit (CPT code = 99281-99285) or a combination of the procedure code and place of service code indicated ED visit. If the entire claim only included lab and imaging codes based on CPT codes = 70000-79999 or 80000-89999, we did not count the claim as an ED visit.

ED visits that led to inpatient stays (i.e., ones that share the same start date with an inpatient stay) were excluded. If two or more ED visits or observation stays had the same patient identifier and date of service, we counted them as one visit.

c. Calculating outcome measures

Once we identified the services and expenditures for each core measure for each month, the monthly measures were summed to the appropriate analysis periods. Only services in a month where a person was eligible for analysis were included in the sums.[113] For individuals eligible for less than the full analysis period, the sum for the eligible months was divided by the proportion of the analysis period for which they were eligible to create a full-time equivalent measure. Regressions were weighted by the proportion of period for which the individual was eligible.[114]

2. Other measures

In this section we describe the methods for creating other dependent and independent variables included in our analyses.

a. Office visits

For Medicare, we identified well-care, primary care, and preventive care visits in the carrier, outpatient hospital, SNF and HHA files based having line items with any codes listed in Table A-VI.10. For each intervention and comparison group member we summed the number of visits in each month that were well care, preventive care or office visit. If there were multiple claims with the same date of service they were count as only one visit even if the providers were different.

Table A-VI.10. Office visit services

Variable name	Specification
Primary care visit	CPT Code = 99201-99205, 99211-99215, 99304-99310, 99315-99316, 99318, 99324-99328, 99334-99350
Well-visit	CPT Code = 99381-99387, 99391-99397, 99432, 99461; ICD-9 = V20.2,V20.3, V70.0, V70.3,V70.5, V70.6, V70.8, V70.9; HPCPS = G0438, G0439
Preventive care	CPT Code = 99401-99404, 99406-99409, 99411-99412, 99420, 99429

[113] For example, if a person had third party insurance coverage in a particular month, they were not counted as eligible for the analysis in that month. In parallel any services provided in that month were excluded from the analysis.

[114] For KMHS, weights for comparison group members were also based on the number of comparison group members associated with the same participant.

b. Analysis control variables

The control variables included in the KMHS regression models are listed in Table A-VI.11 (Medicare) and Table A-VI.12 (Medicaid) along with the specifications for the variables. Variables were derived based on the program enrollment data provided by KMHS, and Medicare or Medicaid administrative data.

Table A-VI.11. Impact analysis model control variable specifications—KMHS Medicare

Variable name	Specification
Treatment indicator	Categorical variable indicating treatment status. Categories include: comparison group member (reference); KMHS intervention participants.
Cohort indicator	Categorical variable indicating cohort group status. Categories include: non-cohort group member (reference); cohort group member.
Interaction between treatment and cohort	Interaction between treatment and cohort indicator variables.
Intervention period	Categorical variable indicating time period of observation. Categories include: six-month increments of the pre-intervention and post-intervention periods.
Interaction between treatment and intervention period	Interaction between treatment and intervention period indicator variables.
Interaction between cohort and intervention period	Interaction between cohort and intervention period indicator variables.
Interaction between treatment, cohort, and intervention period	Interaction between treatment, cohort, and intervention period indicator variables.
Black non-Hispanic race	Indicator variable for individual's race categorized as Black non-Hispanic.
Hispanic ethnicity	Indicator variable for individual's ethnicity categorized as Hispanic.
Unknown race	Indicator variable for individual's race categorized as unknown.
Age	Continuous variable indicating age when first used mental health service at KMHS or a comparison facility in the analysis period (Medicare), or first used mental health service in the analysis period (Medicaid).
Age squared	Continuous variable measuring age as defined above squared.
Sex	Categorical variable of member's sex. Categories include: female (reference); male.
Mental health diagnosis indicators	Indicator variables for mental health diagnoses in first three months in analysis period receiving services at KMHS or comparison facility. Indicators included: schizophrenia; bipolar disorder; depression; persistent mental disorders due to conditions classified elsewhere; dementia; other psychotic disorders; anxiety, dissociative, and somatoform disorders; adjustment reaction disorder; drug or alcohol-related disorder; other mental health diagnosis.
Dually Enrolled in Medicare and Medicaid	Indicator variable for dually enrolled in Medicare and Medicaid based on Medicare enrollment database at time of first mental health visit in analysis period at KMHS or comparison facility.
Disabled	Indicator variable for original reason for Medicare entitlement.
Pre-Period Medicare enrolled	Indicator variable for availability of 12 months of FFS Medicare claims data prior to month of first mental health visit during analysis period at KMHS or comparison facility. Individual must have Medicare as primary insurer, be enrolled in Parts A&B and not be enrolled in Medicare Advantage during the pre-period.
Length of time in mental health treatment	Continuous variable of length of time in mental health treatment.
Length of time in mental health treatment squared	Continuous variable of length of time in mental health treatment squared.
HCC conditions	Flags for conditions from HCC algorithm calculated based on 12 months of Medicare FFS claims data from 12 months prior to first mental health visit in analysis period at KMHS or comparison facility. We excluded conditions with less than 2% of the population.

Table A-VI.12. Impact analysis model control variable specifications—KMHS Medicaid

Variable name	Specification
Treatment indicator	Categorical variable indicating treatment status. Categories include: comparison group member (reference); KMHS intervention participants.
Disabled	Indicator variable for current Medicaid entitlement based on disability
Interaction between treatment and disabled	Interaction between treatment and disabled indicators variables.
Intervention period	Categorical variable indicating time period of observation. Categories include: six-month increments of the pre-intervention and post-intervention periods.
Interaction between treatment and intervention period	Interaction between treatment and intervention period indicator variables.
Interaction between disabled and intervention period	Interaction between disabled and intervention period indicator variables.
Interaction between treatment, disabled, and intervention period	Interaction between treatment, disabled, and intervention period indicator variables.
Sex	Categorical variable of member's sex. Categories include: female (reference); male.
Black non-Hispanic race	Indicator variable for individual's race categorized as Black non-Hispanic.
Hispanic ethnicity	Indicator variable for individual's ethnicity categorized as Hispanic.
Unknown race	Indicator variable for individual's race categorized as unknown.
Pre-Period Medicaid enrolled	Indicator variable for availability of 12 months of Medicaid claims or encounter data prior to month of first mental health visit during analysis period.
Age	Continuous variable indicating age when first used mental health service in the analysis period (Medicaid).
Age squared	Continuous variable measuring age as defined above squared.
First month in observation period	Variables indicating the first month in the observation period. Options included: first month in the observation period occurred before the individual enrolled in Medicaid, first month in the observation period occurred between 1 and 6 months of the individual enrolling in Medicaid, and first month occurred after the first 6 months of the individual enrolling in Medicaid.
Mental health diagnosis indicators	Indicator variables for mental health diagnoses in first three months in analysis period receiving services at KMHS or another facility. Indicators included: schizophrenia; bipolar disorder; depression; other psychotic disorders; anxiety, dissociative, and somatoform disorders; adjustment reaction disorder; drug or alcohol-related disorder; other mental health diagnosis.
CPDS Score	Categorical variables of natural log of CDPS score, calculated based on 12 months of Medicaid claims or encounter data from 12 months prior to first mental health visit in analysis period, interacted with aid category.

E. Supplemental analyses

We conducted several supplemental analyses of the Race to Health! program that were not included in the main body of the report. We present supplemental analysis of program impacts on Medicare utilization and expenditures by cohort status (Section E.1) and health status measures (Section E.2).

1. Analyses by cohort status

As referenced in Section VIII.C of the main body of the report, a component of Race to Health! involved periodically identifying and selecting groups of patients, known as cohorts, based on their physical health comorbidities. While intervention services were used by all KMHS patients, cohort patients were specifically targeted to receive services due to their high

level of need. KMHS staff identified both adult and child cohorts.[115] Impact analysis results suggest that the program reduced overall Medicare expenditures for the non-cohort subgroup, but those for the cohort were unchanged. The reductions in Medicare expenditures for the non-cohort subgroup may be driven by substantial declines in hospitalizations identified in this subgroup. In contrast, we found no change in hospitalizations among the cohort subgroup. ED and office visits declined relative to the comparison group for both the cohort and non-cohort subgroup. Notable findings are as follows:

- During the first two and a half years of the program, we found that the total Medicare expenditures decreased $346 per enrolled month for non-cohort subgroup members relative to the comparison group (p-value < 0.01). Meanwhile, we found no change in expenditures for the cohort subgroup.
- Hospitalizations decreased 0.022 per enrolled month for non-cohort subgroup members relative to the comparison group (p-value < 0.01). In the cohort subgroup we found no significant change in hospitalizations.
- ED visits declined for participants relative to the comparison group in both the cohort (0.020 with p-value = 0.03) and non-cohort (0.050 with p-value = 0.04) subgroups.
- Mean office visits were higher for KMHS patients in the cohort and non-cohort subgroups relative to their comparison groups throughout the analysis period. However, the difference between KMHS patients and their respective comparison groups declined in the post-implementation period for both subgroups resulting in a significant negative estimated impact for the program in both subgroups.

2. Additional health status measures

We conducted analyses of three health status measures in addition to those presented in Section VIII.C of the main body of the report. These three measures related to diabetes control, LDL cholesterol, and smoking status.

Diabetes control. We selected diabetes control for our analysis because, in addition to being a focus of the Race to Health! monitoring and wellness activities, diabetes prevalence among individuals with SMI may be up to two to three times higher than general population prevalence (De Hert et al. 2011a). Our analysis measure is based on the HEDIS measure for comprehensive diabetes care. According to this measure diabetes is considered to be well controlled when HbA1c level is less than 8.0 percent, and poorly-controlled when HbA1c is greater than 9.0 percent.

For our analyses, we identified individuals with who were screened for HbA1c control between January and December 2014. For a subset of individuals who were screened in the first six months of 2014 (to allow for enough time in the data to have 12 months of possible follow-

[115] KMHS staff chose the adult cohorts by using information from the state's PRISM data system and KMHS' EHR. PRISM is a web-based application that integrates data on Medicaid enrollees from multiple sources and provides risk assessment tools such as the chronic disability illness system, which assigns risk scores to Medicaid enrollees based on the severity of their health care needs. For the child cohorts, staff asked providers for recommendations and then analyzed EHR data to search for comorbidities.

up)[116] and had an initial screening value indicating poor HbA1c control, we planned to report on the percentage of these individuals who 1) received a follow-up screening within 12 months following their initial screening, and 2) had a follow-up value indicating HbA1c control.

In 2014, we identified 112 KMHS patients with diabetes. Thirty-three percent of these patients were screened for diabetes control in 2014. However, we are unable to report follow-up screening and control for this measure due to small sample size.

LDL cholesterol. Cholesterol control was another focus of the Race to Health! monitoring activities. We adapted criteria developed by the National Heart, Lung, and Blood Institute at NIH[117] to assign intervention group members to the following risk categories defined by risk factors and LDL levels: 1) no risk – no risk factors and low cholesterol, no statin prescription needed; 2) at-risk – less than two risks factors and LDL level 130-189 mg/dL, statin prescription may be used; and 3) high-risk – two or more risk factors and LDL level 100-189 mg/dL, statin prescription recommended. Risk factors include hypertension, low HDL cholesterol levels, BMI, age, and smoker status. We were not able to specifically analyze statin use in the Race to Health! population due to lack of data.

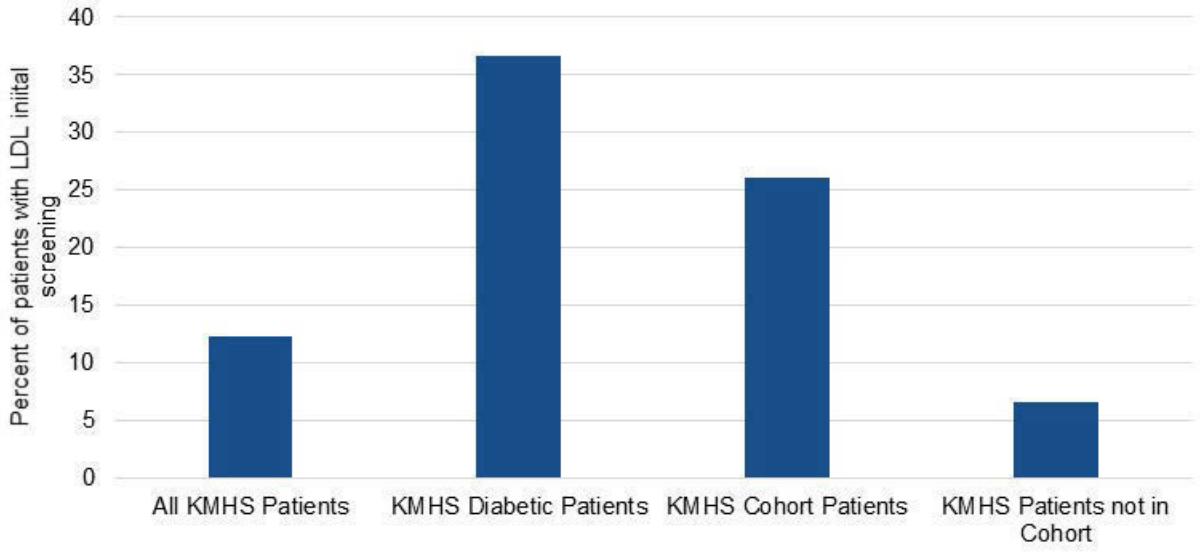
For this analysis, we identified patients who had an LDL screening between January and December 2014 and then categorized patients who had data needed to assign patients to the three risk groups described above. Then, we calculated the proportion of individuals who were at risk or high risk at their initial screening in the first six months of 2014 who had a follow-up LDL cholesterol screening within the following 12 months. Finally, of those who were categorized as at risk or high risk at their initial screening in the first six months of 2014, we calculated the proportion of individuals who had a value for LDL below 130 mg/dL at the follow-up within 12 months of the initial screening.

Only 12 percent of KMHS patients had LDL levels screened at their initial screening between January and December 2014 (Figure A-VI.4). Patients who were diabetic or part of the cohort subgroup had higher rates of screening during 2014 (37 percent and 26 percent, respectively). Among all patients screened during this period, 79 percent were categorized as not at risk, 6 percent were categorized as at risk, and 15 percent were categorized as high risk. We were not able to report further analysis of follow-up due to small sample size.

[116] For this and other health status measures, we first calculate the number and percent of patients who had an initial screening during calendar year 2014 to assess initial screening rates. Then for our follow-up analyses, we focus on the subset of patients who had their initial screening in the first six months of 2014 because that time period allows enough time in the data to have 12 months of possible follow-up.

[117] http://www.nhlbi.nih.gov/files/docs/guidelines/atglance.pdf

Figure A-VI.4. Percent of patients with LDL screening between January and December 2014



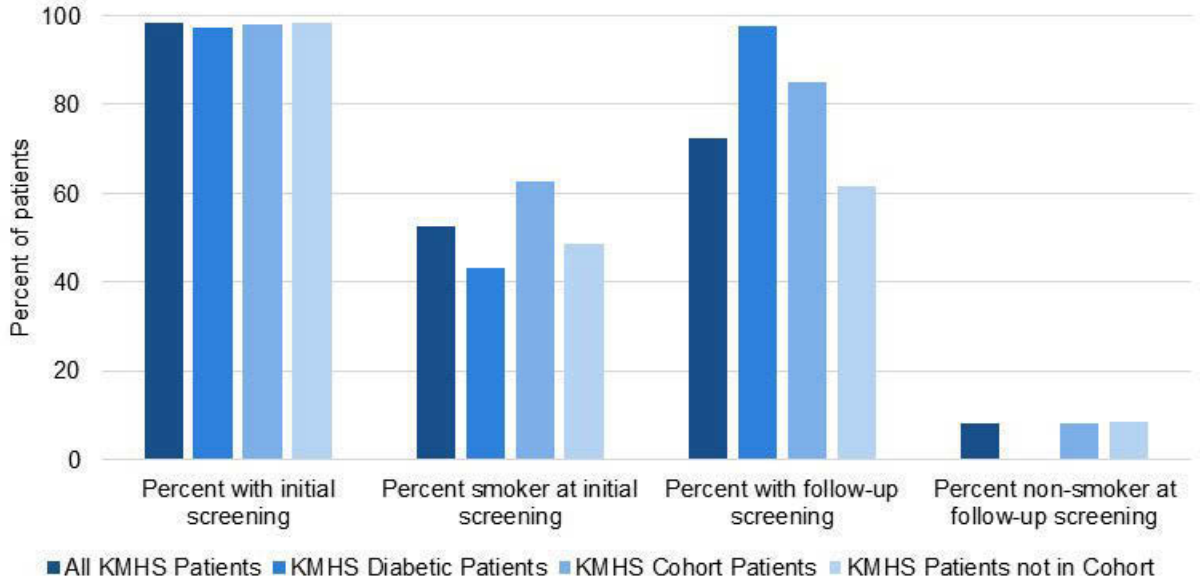
Source: Mathematica analysis of KMHS' EHR data for January 2014-June 2015.

Note: The denominator for each bar is the number of patients in the noted patient group who were an active patient in 2014 and were age 18 or older as of January 1, 2014.

Smoking status. Race to Health! aimed to reduce smoking through improved monitoring of patients' smoking habits, as well as through the introduction of new wellness activities. For our analysis, we calculated the percent of patients who had a smoking screening in 2014 and identified current smokers at this initial screening. Then, we calculated the proportion of individuals who were a smoker at their initial screening in the first six months of 2014 who had a follow-up smoking screening within 12 months and we calculated the proportion of these individuals whose most recent reported value of smoking status indicated they were no longer a smoker.

Nearly all KMHS patients and each subgroup were screened for smoking status in 2014 (Figure A-VI.5). Over half of KMHS patients were smokers at initial screening in 2014; fewer diabetic patients were smokers (43 percent) and more cohort patients were smokers (63 percent). Among patients initially screened in the first six months of 2014 who were smokers, 72 percent of all KMHS patients had a follow-up smoking screening; follow-up screening rates were higher for diabetic and cohort patients (98 percent and 85 percent, respectively). Of these patients, consistently fewer than 10 percent were non-smokers at their follow-up.

Figure A-VI.5. Smoking status screening and follow-up



Source: Mathematica analysis of KMHS' EHR data for January 2014-June 2015.

Note: The denominator for "percent with initial screening" and "percent smoker at initial screening" is those patients meeting the criteria of each patient group, and who were active patients in 2014 and were age 18 or older as of January 1, 2014. The denominator for "percent with follow-up screening" and "percent non-smoker at follow-up screening" is those patients meeting the criteria of each patient group, and who were active patients in 2014 and were age 18 or older as of January 1, 2014, and who were smokers at an initial screening in the first six months of 2014.

VII. MAIMONIDES MEDICAL CENTER

A. Introduction

We conducted two analyses of the MMC program. To measure the impact of the MMC program on key outcomes for MMC's Medicare participants, we conducted an impact analysis. We also conducted a pre-post analysis to measure the trends in outcomes over time for MMC's Medicaid participants. In the following sections, we describe our data sources (Section B), how we identified the intervention groups for the Medicare and Medicaid analyses (Section C), the steps to construct the comparison group for the Medicare impact analysis (Section D), and how we specified the measures for both analyses (Section E).

B. Description of data sources

In this section, we provide a general overview of the data sources used in the analyses for MMC:

- **Finder files.** MMC provided files with participant SSN, program enrollment date, demographic information (date of birth and gender), and Medicaid and Medicare identifiers (HICs) for participants enrolled in these programs. The SSNs and Medicaid and Medicare program identifiers were used to identify program participants in Medicaid and Medicare administrative data.
- **CMS Medicare administrative data.** Our analysis of MMC program impacts for the Medicare population used CMS Medicare administrative data. We obtained data files through the CMS's VRDC. We extracted all final action claims with dates of service from January 1, 2009 through June 30, 2015[118] for individuals for whom the HIC or SSN included in the MMC finder files matched to a BENE_ID in the VRDC cross-reference files. We extracted standard analytic base and revenue-center/line-item claims datasets for the following claim types: carrier, DME, home health, hospice, inpatient, outpatient, and skilled nursing facility. To obtain information on beneficiary Medicare enrollment spans we used the MBSF. The MBSF includes information on date of birth, gender, most recent county of residence, enrollment in MA, and third party insurance coverage.
- **NYS Medicaid claims and enrollment data.** We obtained claims and enrollment data from the NYSDOH Medicaid Data Warehouse for the period from January 1, 2009 through June 30, 2015. The claims data provided information on FFS and managed care payment amounts. When service level payment information was not available for managed care covered services, these payment amounts were estimated based on FFS payment guidelines. Claims fields relevant to this analysis also included service type, provider type, and procedure and diagnosis codes. The enrollment data provided monthly Medicaid enrollment and demographic information. Participants in MMC's health home were selected for this extract based on Medicaid identifiers provided by MMC.

[118] Claims for the intervention group members were extracted from February 29 to March 1, 2016. Claims for the comparison group were extracted from May 8–11, 2016.

C. Identification of the intervention populations

In this section, we discuss how we identified the intervention populations used in our Medicare and Medicaid analyses. We could not include all program participants in either analysis due to data source limitations as well as specific exclusion criteria we applied.

1. MMC - Medicare

As described above, MMC provided Mathematica with HIC numbers for their Medicare enrolled participants who enrolled in the MMC program between February 2013 and June 2015. We used these identifiers to extract Medicare enrollment and claims data for participants from CMS's VRDC. We started with 2,138 MMC Medicare participants and applied a number of exclusions in the following order:

- **CBC enrollees.** We excluded 764 participants who received care management services through Coordinated Behavioral Care (CBC), another Medicaid health home with which MMC collaborated beginning in 2014. We excluded CBC participants from the analysis because they primarily received services outside of the MMC Medicaid health home program and these services were not funded by HCIA.
- **Missing enrollment information.** We excluded 195 participants from our analysis because they were missing MBSF data in their MMC enrollment year.
- **Lack of Medicare FFS enrollment.** We excluded 563 participants from the sample because they were not continuously enrolled with FFS Medicare as their primary payer for six months around their program enrollment month (three months before their enrollment month, the month of enrollment, two months after enrollment). We excluded three participants because they were enrolled in FFS Medicare Parts A and B with Medicare as primary payer for less than 6 months in the 12 months before enrollment.
- **Geographic location.** We excluded 25 participants from the sample because they were not physically located in the MMC service area in Brooklyn, NY (as measured by zip code) during the month of enrollment.
- **Lack of SMI diagnosis in claims.** Beneficiaries were required to have a diagnosis of at least one of the targeted SMI condition categories: schizophrenia and related disorders, bipolar disorders, depressive disorders, other psychotic disorders, and childhood emotional disturbance. Beneficiaries were defined as having a condition if they had at least one inpatient or two or more outpatient claims (not including prescription drugs) with the relevant diagnoses in the two years prior to enrollment. We excluded 120 participants from the sample because they did not have a diagnosis of at least one of the targeted SMI condition categories in the claims data in the two years prior to enrollment. Then, we dropped participants with diagnoses of other psychotic disorders (n=7) and childhood emotional disturbance (n=0) due to small sample size. See Table A-VII.1 below for the diagnosis codes used to identify participants with the three remaining qualifying SMI condition category.

Although our exclusions were processed in the order above, it should be noted that most participants who were excluded, were excluded for multiple reasons. Our final pre-matching Medicare intervention group included 464 participants.

Table A-VII.1. Diagnoses codes used to identify qualifying condition categories in treatment and comparison groups (MMC)

Schizophrenia and related disorders	295.XX including 295.00
Bipolar disorders	296.00, 296.01, 296.02, 296.03, 296.04, 296.05, 296.06, 296.10, 296.11, 296.12, 296.13, 296.14, 296.15, 296.16, 296.40, 296.41, 296.42, 296.43, 296.44, 296.45, 296.46, 296.50, 296.51, 296.52, 296.53, 296.54, 296.55, 296.56, 296.60, 296.61, 296.62, 296.63, 296.64, 296.65, 296.66, 296.7, 296.80, 296.81, 296.82, 296.89, 296.90, 296.99
Depressive disorders	296.20, 296.22, 296.23, 296.24, 296.25, 296.26, 296.30, 296.32, 296.33, 296.34, 296.35, 296.36

2. MMC – Medicaid

As described above, MMC provided Mathematica with HIC numbers for their Medicaid enrolled participants who enrolled in the MMC program between February 2013 and June 2015. We obtained claims and enrollment data from the NYSDOH Medicaid Data Warehouse, using these identifiers to obtain data for MMC participants. We started with 8,946 MMC Medicaid participants who matched to NY data and applied two exclusions in the following order:

7. **Medicare-Medicaid dual enrollees.** We excluded 1,349 participants who were dually enrolled in both Medicare and Medicaid.
8. **CBC enrollees.** We excluded an additional 2,079 participants who received care management services through Coordinated Behavioral Care (CBC), another Medicaid health home with which MMC collaborated beginning in 2014. As for the Medicare analytic population, we excluded CBC participants from the analysis because they primarily received services outside of the MMC Medicaid health home program and these services were not funded by HCIA.

Our final Medicaid intervention group included 5,518 participants.

D. Identification of the comparison population

We used matching techniques to develop a comparison group for MMC's FFS Medicare participants. Propensity score matching and related matching methods are designed to create a comparison group that is similar in observable characteristics to the treatment group (Rosenbaum and Rubin 1983; Dehejia and Wahba 2002). Limiting the comparison group to a matched subsample of Medicare beneficiaries—closely matching on observed characteristics of the participants—may also reduce differences between participants and comparison group members in terms of unobserved characteristics if those characteristics are correlated with matching variables. We identified Medicare enrollees residing in three comparison cities (Philadelphia, Pennsylvania; Pittsburg, Pennsylvania; and Chicago, Illinois) with schizophrenia and related disorders, bipolar disorders, and/or depressive disorders[119] as the potential group of comparison patients. We focused on this subset of qualifying condition diagnoses because only seven

[119] We excluded participants with diagnoses of other psychotic disorders or childhood emotional disturbance only from our analysis due to small sample size.

treatment group members in Medicare FFS had other psychotic disorders diagnoses and no treatment group members had diagnoses of childhood emotional disturbance.

We chose a comparison group outside of NYS for several reasons. First, NYS's health home program was implemented across NYS and individuals enrolled in Medicare who are dually enrolled in Medicaid with eligibility for full Medicaid benefits are able to enroll. Thus, potential comparison group members for participants who were dually eligible for full Medicaid benefits may have been matched to individuals enrolled in other health homes, making them an inappropriate comparison group because they would be affected by a similar intervention. In addition, we were unable to obtain provider identifiers for all of MMC's many partners. Thus, we could not exclude patients from the comparison group who were not participating in the intervention but who received services from an MMC-participating provider, and thus may have indirectly benefited from the intervention. By going outside of NYS to choose the comparison group, this potential contamination was avoided. To identify the most appropriate cities to use as comparison sites to MMC's service area in Brooklyn, NY, we conducted a comprehensive analysis of the relevant demographic, socioeconomic, and health care factors of approximately 20 of the largest urban centers in the country that are also located in states that did not implement a Medicaid health home program. We compared locations by examining the following characteristics: total Medicare spending per beneficiary, Medicare enrollee hospital discharge rates, the number of all physicians and primary care physicians per 100,000 residents, city poverty rate, and median household income. These characteristics are listed in Table A-VII.2 below.

Table A-VII.2. Characteristics of major metropolitan areas nation-wide without Medicaid Health Home Program

Candidates	City pop. rank	Region	Price, age, sex, and race adjusted total Medicare spending, 2012	Hospital discharges per 1,000 Medicare enrollees, 2012	All physicians per 100,000 residents, 2011	Primary care physicians per 100,000 residents, 2011	City poverty rate, 2010[a]	City median household income, 2010[a]	Total number of Medicare enrollees, 2012
Brooklyn, NY	#1 (Part of NYC)	NE	$11,371	323.1	267	94	23.4%	$46,958	153,548
Boston, MA	#10	NE	$9,632	286.5	325	117	21.9%	$54,485	70,962
San Antonio, TX	#25	South	$10,330	262.7	195	63	20.1%	$46,317	141,885
San Bernardino, CA	#13	West	$10,508	289.1	173	65	33.0%	$38,774	11,852
Pittsburgh, PA	#22	NE	$10,725	313.5	226	76	22.8%	$40,009	61,122
Philadelphia, PA	#5	NE	$10,554	316.1	279	91	26.7%	$37,460	120,007
Joliet, IL	#3 (Part of Chicago)	Midwest	$10,958	375.7	194	64	12.5%	$62,008	52,860
Chicago, IL	#3	Midwest	$11,017	321.3	260	102	22.7%	$47,831	205,812
Dallas, TX	#4	South	$11,039	255.2	208	66	24.1%	$43,359	149,774
Tampa, FL	#19	South	$11,427	303.4	203	70	22.0%	$43,740	58,072
Houston, TX	#6	South	$11,535	262.4	175	50	22.9%	$45,728	234,994
Fort Lauderdale, FL	#8	South	$11,808	293.9	238	84	21.2%	$48,898	37,549
Fort Worth, TX	#4 (Part of Dallas)	South	$11,905	304.9	165	57	19.3%	$52,492	72,874
Gary, IN	#3 (Part of Chicago)	Midwest	$12,260	416.8	192	62	38.7%	$27,458	24,945
Los Angeles, CA	#2	West	$12,907	313.2	236	79	22.4%	$49,682	109,206
Thresholds around the Brooklyn, NY mean			+/- 10%	+/- 20%	+/- 35%	+/- 35%	+/-20%	+/- 21%	n.a.

Source: "The Dartmouth Atlas of Health Care." 2016. The Dartmouth Institute for Health Policy and Clinical Practice. Available at http://www.dartmouthatlas.org/data/region/.

Note: Gray shading indicates the Brooklyn, NY mean to which other cities are compared. Red shading indicates values that fall outside the thresholds around the Brooklyn, NY mean. Blue shading indicates cities for which all values fall within the thresholds.

[a] US Census Bureau Community Facts, 2010-2014 American Community Survey 5-Year Estimates. Available at http://factfinder.census.gov/faces/nav/jsf/pages/community_facts.xhtml.

n.a. = not applicable.

Chicago, Pittsburg, and Philadelphia were the only three cities that were similar to Brooklyn on all measures of interest, and were also comparable in regard to region and population composition. Like Brooklyn, the three comparison cities are major east coast or Midwest metropolitan areas. Unlike NYS, neither Pennsylvania nor Illinois started a health home program. In addition, we did not identify any major changes to the Medicaid program during the relevant time period in either state; thus, care was more stable in Philadelphia, Pittsburg, and Chicago, making them good comparisons.

From within the general pool of FFS Medicare enrollees with the targeted SMI diagnoses in the three comparison cities, we matched individuals to MMC's Medicare participants. Constructing the matched comparison group involved several steps, which we detail below.

Step 1: Identify potential comparison pool members. We initially extracted claims and enrollment information for all Medicare beneficiaries residing in the three comparison cities who had claims indicating one or more of the three qualifying conditions[120] during a 24-month period prior to any month between February 2013 and March 2015. For each potential comparison pool member, we needed to create a pseudo-enrollment month that reflected the month when the member likely would have enrolled in the intervention if they had been a participant. The pseudo-enrollment month allows us to define the pre- and post-intervention periods for the analysis. For each person in the potential comparison pool we identified all the months between February 2013 and March 2015 which had an eligible SMI diagnosis in the prior 24 months and an office visit claim in the current month. These criteria aim to ensure potential comparison beneficiaries have at least one target condition during the 24-month pre-period and that they have some engagement with the healthcare system, as measured by a primary care visit. Comparison pool members who did not have any months meeting this criteria were dropped. For remaining potential comparison pool members, we randomly selected one of their eligible months to be the pseudo-enrollment month, weighting the probability of selecting a given month based on the proportion of intervention participants who enrolled in the same month relative to the proportion of comparison pool members for which the month could be selected.

Step 2: Apply exclusion criteria. Once each potential comparison pool member had a pseudo-enrollment month assigned, we applied exclusion criteria parallel to the exclusion criteria for intervention group members discussed above. We excluded comparison pool members who were not continuously enrolled with FFS Medicare as their primary payer for six months around their pseudo-enrollment month (three months before their enrollment month, the month of enrollment, two months after enrollment). We also required at least six months of FFS Medicare data in the year prior to pseudo-enrollment. We excluded any potential comparison group members whose current or original eligibility is by ESRD.

Our potential comparison group included 48,067 beneficiaries who met all inclusion criteria and for whom we were able to set a pseudo-enrollment date.

Step 3: Match treatment participants at the individual level. The matching process used metrics of individual-level characteristics identified based on pre-period data to select a subset of comparison pool members who were as similar as possible to the intervention group on

[120] At least one inpatient or two outpatient claims with the indicated diagnoses.

observable characteristics. The matching algorithm first exact matched on diagnoses of schizophrenia and related disorders, bipolar disorders, and/or depressive disorders, and disability status. We then fit a propensity score model. A beneficiary's propensity score is the probability of belonging to the treatment group estimated from this model. We included the following characteristics in the model: age, sex, race (White, Black, and other), dual status, HCC condition indicators (created as part of creating the HCC score),[121] the number of months the beneficiary was Medicare FFS eligible during the year prior to enrollment (or pseudo-enrollment), year and month of enrollment (or pseudo-enrollment), and number of qualifying condition categories (i.e., one, two, or three of the qualifying condition categories). We also included pre-period levels of two of the core outcomes measures (hospitalizations and ED visits), as well as the number of primary care visits, broken out into categories for the 12 months prior to enrollment. We chose not to include total expenditures due to potential differences in Medicare geographic adjustments to payment in New York City versus the comparison sites that might lead to different levels of spending for individuals with the same acuity. We also did not include readmissions due to the small number of participants who had readmissions.

We matched up to seven comparison group beneficiaries to each treatment group beneficiary. When a treatment beneficiary was difficult to match (that is, had few similar comparison beneficiaries), the algorithm conducted a pair match. When there was an abundance of comparisons for a treatment beneficiary, the algorithm matched multiple comparisons. The statistical goal was first to minimize bias and then, subject to that constraint, maximize the size of the comparison sample. The optimal matching algorithm that we used selected comparison beneficiaries without replacement and minimized the overall differences between treated and matched comparison beneficiaries so that they were similar, on average, on observed characteristics in the pre-period.

Step 4. Assess the quality of the match. This section describes diagnostic tests that we used to assess the quality of the matches.

We began by examining the ratio of potential comparison beneficiaries to treatment beneficiaries by exact matching strata prior to matching in order to understand how difficult it might be to match at least one comparison beneficiary to each participant. For example, if many strata had low ratios of potential comparison beneficiaries to participants prior to matching, we might have considered reducing the number of variables used for exact matching. We found a sufficient number of comparison group individuals for each treatment group person in each strata.

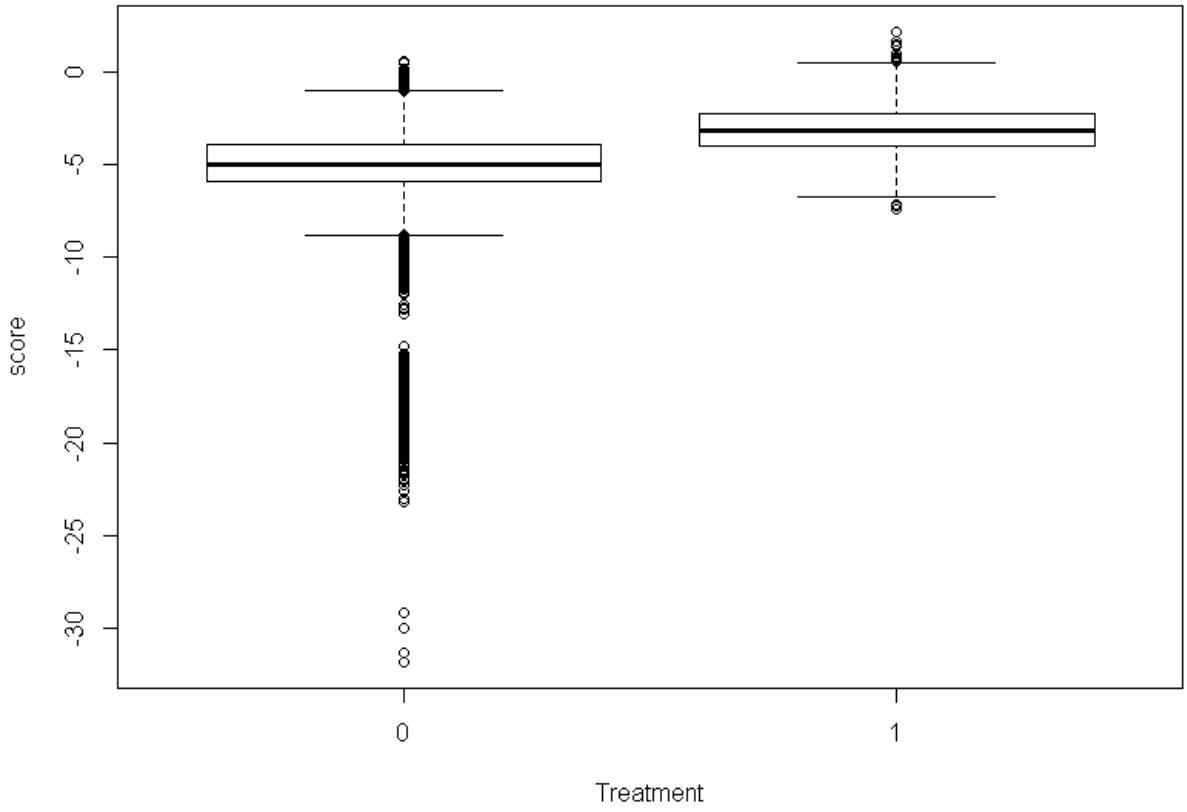
We graphically compared the propensity score distributions[122] for all treatment and comparison beneficiaries prior to matching, looking for overlap in the propensity scores for the treatment and comparison groups (Figure A-VII.1). The distribution of propensity scores differed

[121] HCC score = Hierarchical Condition Category Score. The HCC model was developed to risk adjust Medicare payments to Medicare Advantage plans by assessing expected expenditures of enrollees. The HCC score provides a proxy of overall health status, as sicker individuals are expected to cost more than healthier individuals.

[122] We calculated the log-odds of the propensity score rather than the more common probability scale, because log odds provided better overlap.

between the two groups, which is to be expected; however, there appeared to be sufficient overlap to find comparison group members for each treatment group member.

Figure A-VII.1. Log-odds of propensity score distributions for treatment and comparison pool members (MMC)



Note: Figure presents boxplots created using the log-odds of the estimated propensity scores for the comparison and treatment groups, in the left and right panels respectively. The width of the boxplots corresponds to the amount of data that contributed to the plots.

After we conducted matching, we examined the number of comparison beneficiaries matched to each treatment beneficiary (Table A-VII.3). A large number of 1:1 matches, or a large number of comparison beneficiaries that were excluded, could indicate that the matching was problematic. This was not an unexpected problem for MMC, as we knew that treatment group members were actively recruited and that selection criteria could not be mimicked for the comparison group. However, we found the number and distribution of matches to be acceptable. After the matching, there were 4,923 matched beneficiaries in the comparison group.

Table A-VII.3. Frequency table of ratio of treatment beneficiaries to comparison beneficiaries for each matched set (MMC)

Ratio of treatment to comparison beneficiaries	1:1	1:2	1:3	1:4	1:5	1:6	1:7	0:1
Number of matched sets	126	24	23	24	28	22	616	43,144

Note: Each cell indicates the number of treatment beneficiaries matched to the number of comparison beneficiaries indicated for that column.

After evaluating the basic matching diagnostics above, we examined the overall balance of the matched sample. We used an omnibus test that checks for covariate balance across the treatment and matched comparison beneficiaries (Hansen and Bowers 2008). The omnibus test is based on the differences between treatment and matched comparison beneficiaries across the matching variables; these differences are standardized by their variances and covariances and aggregated into a single number, a weighted mean. Standardization in this way implies that a matching variable whose difference across matched sets has a small variance is given more weight and that a matching variable whose difference across sets is highly correlated with other differences is given less weight. The advantages of the omnibus test are: (1) it generates a single probability statement through one *p*-value; (2) its distribution is roughly chi-square, which facilities the calculation of the *p*-value; and (3) it assesses balance on all linear combinations of the matching variables. However, a significant result from this chi-squared test may be driven by a large sample size rather than substantive differences between treatment and matched comparison groups. Alternatively, it could indicate that there may be some imbalance between the two groups on at least one of the matching variables. The results of this test were a chi-square statistic of 132.83 and associated *p*-value of 0.02.

To further investigate imbalance between treatment and matched comparison groups, we evaluated how matching affected the balance on all matching variables by comparing the absolute and standardized differences between the treatment and comparison groups for each variable before and after matching (Figure A-VII.2). The standardized difference measures the difference in means in units of the pooled standard deviation. The standardized difference measure is advantageous in that it allows us to compare all variables on the same scale. We compared the standardized differences using plots with dashed lines at +/- 0.10 and +/- 0.20 standardized differences to visually inspect whether we obtained good balance for each variable.

Figure A-VII.2. Balance plot comparing the standardized difference for each matching variable before and after matching (MMC)

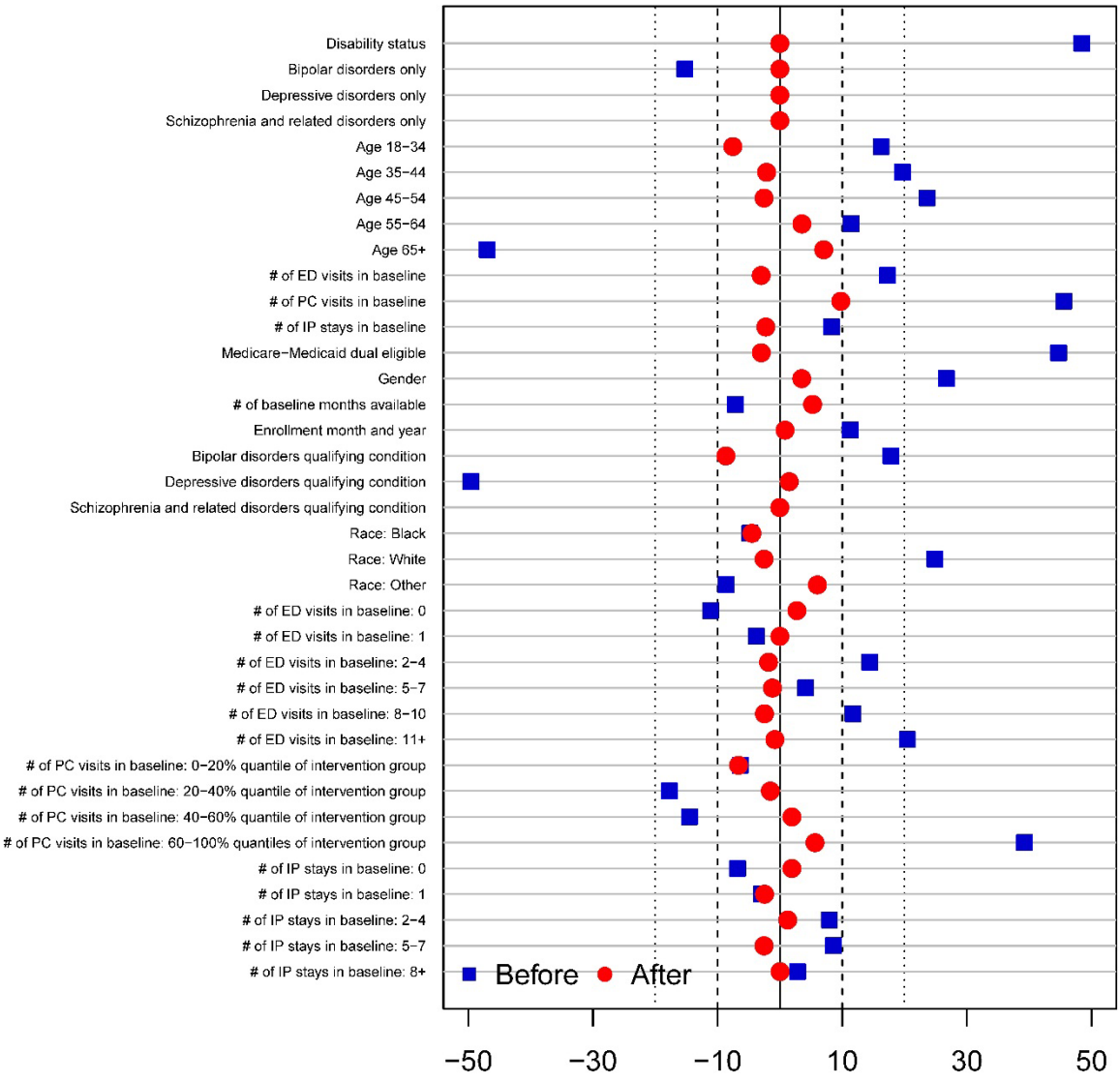
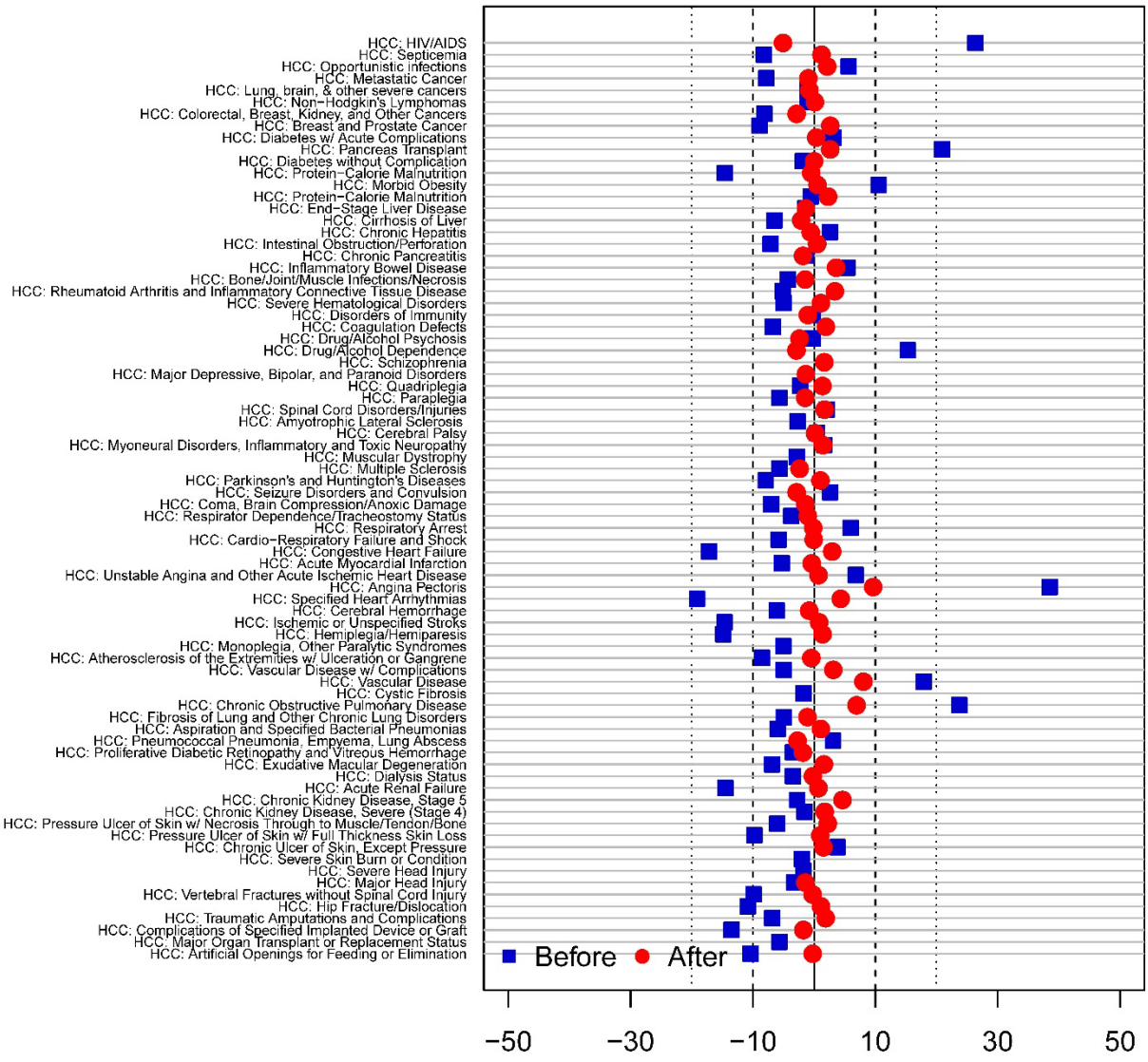


Figure A-VII.2 *(continued)*



In addition to the exact match variables (with zero absolute and standardized differences), we ideally wanted all variables to fall within +/- 0.10 standardized differences. All of the matching variables met these conditions. Although we did not include total expenditures as a matching variable, we examined the standardized differences between the treatment and comparison group for this measure. This difference fell within +/- 0.25 which we felt was acceptable given the known difference in Medicare spending between NYS and the comparison sites. The absolute mean difference between the intervention and comparison groups on the expenditure measure was $3,915.

E. Specifications of measures

We analyzed program impact on four of CMMI's core outcome measures: total Medicare or Medicaid[123] expenditures, inpatient hospitalizations, hospital readmissions,[124] and ED visits. These outcomes are appropriate for evaluating the MMC program because improvements in care coordination and management were anticipated to reduce acute care service use and thereby reduce overall expenditures.[125] Our specifications for these measures in Medicare and Medicaid administrative data are described in Section E.1 below. Our analyses also used several other types of measures, as described in Section E.2 below.

1. Core measures in Medicare and Medicaid administrative data

In this section, we provide detail on the data and analytic methods used to develop the core outcome measures in Medicare and Medicaid administrative data. We begin by describing how we identified the patient population and the associated spans of Medicare or Medicaid enrollment that were included in the analyses. Then, we describe how we processed claims data and assigned expenditure and utilization information to months to develop each of the core measures. Finally, we discuss how we annualized and weighted the regressions models to adjust for individuals who were not observable for a full 12 months.

a. Identifying periods with observable data

In this section we describe the approach we used to identify the patients and periods of Medicare or Medicaid enrollment included in the analyses. When an individual was not enrolled in Medicare/Medicaid, if Medicare/Medicaid was not their primary insurance, if they were not covered by Medicare Part B, or if they were enrolled in MA, their health expenditures and utilization were not consistently observable in the administrative data available for this analysis.

[123] Medicaid expenditures include both FFS and managed care payments. When service level payment information was not available for managed care covered services, these payment amounts were estimated based on FFS payment guidelines.

[124] We were unable to estimate the readmission measure for the Medicaid population.

[125] Because of data limitations, Medicaid costs and service utilization for dual Medicare-Medicaid enrollees are not included in the analyses, even though dual enrollees are included in the Medicare analytic population. Although Medicare is the primary payer, the exclusion of Medicaid costs for dual enrollees means that specialized services for people with serious mental illness covered under Medicaid options and waivers provided to dual enrollees are not reflected in the analyses.

Thus, we limited our analysis to patients and time periods during which sufficient data was available to calculate the core measures.

i. CMS Medicare administrative data

Identifying the patients and periods of enrollment to include in the analysis for CMS Medicare administrative data required several steps.

Step 1: Link awardee identifiers to CMS administrative files. MMC provided us with a finder file including HIC numbers and SSNs for all participants. We first matched the HIC numbers to the VRDC BENE_ID crosswalk. Individuals who did not match to the crosswalk by HIC number were then matched by SSN. Matches by HIC and SSN were verified by comparing the date of birth, gender, SSN, and HIC to the data from the matched record. Records that matched on all of these variables or that had only a discrepancy in one component of these variables were retained in the analysis. For example, if HIC, SSN, gender, year of birth, and month of birth matched but day of birth was discrepant, the record was retained in the analysis. Where discrepant information was identified, the information from the Medicare record was used for the remainder of the analysis because this information was deemed more reliable than the information included in the patient record.

Step 2: Exclude months where FFS Medicare is not the primary payer. In order to be included in the analysis, the potential analysis months had to meet the following requirements: (1) the person had to be enrolled in Medicare Part A and B during the month; and (2) the person could not: a) be enrolled in MA, b) have a primary insurer that was not Medicare, c) be a railroad retiree, or d) have a date of death prior to the enrollment month. Based on the criteria for identifying intervention patients and the criteria for excluding months from the analysis based on Medicare enrollment information, we created a variable for each month from January 2010 to June 2015 indicating whether or not the month was eligible for analysis. This indicator was used to identify enrolled months to include in the analysis as well as to assure that services were only included when the associated service month was eligible for the analysis. See Section C.1 above for additional exclusion criteria that were applied during the development of the intervention participant group.

Step 3: Define baseline and intervention periods. Baseline and intervention periods were defined for each intervention participant or comparison group member, relative to their enrollment month (or pseudo-enrollment month).[126] The first intervention period was defined as the enrollment month and five months following that month. Where applicable the second intervention period was defined starting in the months following the last month in the first intervention period. The first baseline period started in the month prior to the enrollment month and moved backward five months. For each individual included in the analysis the proportion of each baseline and intervention period for which the individual was eligible for the analysis was calculated. This proportion was used to pro-rate the expenditure and utilization measures for individuals enrolled for less than the full analysis period. It was also used to weight observations in the regression analysis.

[126] Pseudo-enrollment was defined for comparison group members as described in Section D.1.

ii. MMC Medicaid administrative data

Identifying the patients and periods of enrollment to include in the analysis for MMC required several steps.

Step 1: Link awardee provided Medicaid identifiers to administrative data. We identified intervention group members based on Medicaid identifiers provided by MMC. We first matched the Medicaid identifiers in the program administrative data to those in the NYS Medicaid enrollment data extract. Identifiers associated with 44 (less than 0.5 percent) individuals did not match to the extract. For records that did match, we compared the gender, day of birth, month of birth, and year of birth listed in the Medicaid enrollment extract to the same information in the program administrative data. Matches were excluded from the analysis if there was a discrepancy in more than one of these measures. For example, if gender, year of birth, and month of birth matched but day of birth was discrepant, the record was retained in the analysis. Where discrepant information was identified, the information from the Medicaid extract record was used for the remainder of the analysis because this information was deemed more reliable than the information included in the patient record. Because of discrepant information, 1,500 matches (14.4 percent) were excluded, resulting in 8,946 individuals moving to the next analytic step.

Step 2: Exclude Medicare-Medicaid dual enrollees and enrollees in CBC. To ensure a consistent set of benefits were represented in the Medicaid administrative claims for the analysis population, we required full benefit Medicaid enrollment and no third party coverage. Based on this restriction, 1,349 individuals who were dually eligible for Medicare and Medicaid were excluded. We also excluded 2,079 participants who were in the CBC program, as mentioned earlier. These exclusions reduced the sample to 5,518 MMC participants.

Step 3: Define baseline and intervention periods. Six-month baseline and intervention periods were defined for each intervention participant relative to their enrollment month in the same manner as described in Step 3 above for CMS Medicare administrative data. To ensure that there was sufficient Medicaid enrollment in each analysis period, we limited the analysis to include eight 6-month baseline periods (counting back from the enrollment date) and four 6-month intervention periods (counting forwards from the enrollment date).

b. Summarizing monthly expenditures and utilization

Once the individuals and periods eligible for the Medicare and Medicaid analyses were identified as described above, expenditures and utilization associated with each core measure were aggregated for the periods during which the individual was deemed eligible for the analysis. In this section, we define the specifications for identifying total Medicare or Medicaid[127] expenditures, hospitalizations, hospital readmissions,[128] and ED visits. We

[127] Medicaid expenditures include both fee-for-service and managed care payments. When service level payment information is not available for managed care covered services, these payment amounts are estimated based on fee-for-service payment guidelines.

[128] We were unable to estimate the readmission measure outcome for the Medicaid population.

summarized each of these measures monthly for each individual in the analysis population. Then, we aggregated sets of months for annual impact analysis for Medicare.

i. Expenditures

For Medicare, the following claim types were included in this analysis: carrier, DME, home health, hospice, inpatient, outpatient, and skilled nursing facility. Only FFS data were included in this analysis. Part D services were excluded. Duplicate and denied claims were excluded. The total payment amount on each Medicare claim was summed across all file types to calculate total expenditures. For services that extend beyond a single day (for example, an inpatient or long-term care stay) we counted all Medicare payments recorded based on a single date. Inpatient stays expenditures were counted in the month of the discharge date. For other types of claims, all expenditures were assigned based on the claim from date. Expenditures were excluded from this analysis if they were assigned to a month during which the associated Medicare beneficiary was deemed ineligible for the analysis.

All claim types in the NY Medicaid administrative data were included in the analysis. Duplicate and denied claims were excluded. For claims with services spanning more than one day, expenditures were counted based on the service begin date. Expenditures included both FFS and managed care payments, but excluded capitation payments. When service level payment information was not available for managed care covered services, these payment amounts were estimated based on FFS payment guidelines.

ii. Hospitalizations

The specifications for the hospitalization measures were developed to align with the CMMI priority all-cause admissions per patient measure. For this measure, only acute stays or psychiatric stays were included in the analysis. We describe the steps to develop these counts here.

Step 1: Identify hospitalization claims. For Medicare administrative data, we identified inpatient hospital claims by claim type. Then, we identified and excluded rehabilitation and long-term care based on provider identifier codes. At the end of this step, only acute and psychiatric stays were included in the file.

For NY State Medicaid data, inpatient hospital claims were identified by using the Medicaid Managed Care Operating Report code (MMCOR_CD) values of 01 (“Inpatient Psych, Acute Detox Subabuse”) or 04 (“Medical/surgical”), Surveillance and Utilization Review System Category of Service code (SURS_SUBSYSTEM_COS_CD) value of 11 (“Inpatient”), and the eMedNY claim type code (CLAIM_TYPE_CD) value of “I” (“Inpatient”).

Step 2: Eliminate duplicate or denied claims. For Medicare and Medicaid, we identified claims with the same information in all fields and only kept one of these claims. We also excluded denied claims from our analysis.

Step 3: Combine claims that represent the same stay and combine transfer stays with initial stays. For Medicare and Medicaid data, we identified and combined initial and interim claims into one discharge. Interim claims had (1) the same admission date as the initial claim, (2) an admission date that was equal to the discharge date from the initial or another interim

claim and the status on the other (previous) claim was "still a patient", or (3) a claim with an admission date that was equal to one day after the discharge date of the initial or another interim claim and the status on the other previous claim was "still a patient." Such claims were combined to count as a single stay.

Next, we identified and combined claims associated with a transfer into a single stay. We identified claims indicating that the patient was transferred to either another short-term hospital, a CAH, another type institution for inpatient care, a federal hospital, or a psychiatric hospital or unit. Then combined these claims with claims for the same beneficiary at a different facility where the admission date fell within one day of the discharge date of the first claim.

Step 4: Sum the number of discharges in each month. Once claims representing a single stay were combined, we summed the number of unique discharges for each enrollee for each month. Inpatient stays were counted in the month of the discharge date.

iii. Readmissions

Hospital readmissions were only counted for the Medicare analysis. The approach to calculating hospital readmissions in the Medicare claims data required several steps. We describe these steps below.

Step 1: Select stays qualifying as index stays. We began with the stays identified above for the hospitalization measure. Then we excluded stays that ended in death, had a principal diagnosis of pregnancy or condition originating in the perinatal period, or for which the patient was not continuously enrolled in Medicaid for the 30 days following the discharge date.

Step 2: Identify stays qualifying as readmissions. The remaining discharges were designated as index discharges. We identified readmissions for the same patients in the 30-day window following the discharge date. Then we excluded planned readmissions following HEDIS specifications.

Step 3: Sum index stays and readmissions by month. For each patient and calendar month, we summed the index stays with a discharge date in the month and any associated readmissions. To be included in our analysis the patient had to be continuously eligible for our analysis during the 30-day period following discharge from the index stay.

iv. ED visits

Outpatient ED visit utilization is reflected in CMMI priority measure 62. This measure includes ED visits that do not lead to an inpatient stay, as well as observation stays that do not lead to an admission.

In the Medicare outpatient file, we identified outpatient ED claims as those with a revenue center value indicating and ED visit, excluding any claims that involved only lab or imaging services in the ED. We identified observation claims based on the combination of revenue center code, CPT-code and a unit count of greater than or equal to eight hours.

In addition, for Medicaid data, we reviewed claims not identified as inpatient and considered them as ED visits if the procedure code, cost center revenue code, or managed care operating report code indicated ED visit.

ED visits that led to inpatient stays (i.e., ones that overlapped with or were adjacent to an inpatient stay) were excluded. If two or more ED visits or observation stays had the same patient identifier and beginning date of service, we counted them as one visit.

c. Calculating outcome measures

Once we identified the services and expenditures for each core measure for each month, the monthly measures were summed to the appropriate analysis periods. Only services in a month where a person was eligible for analysis were included in the sums.[129] For individuals eligible for less than the full analysis period, the sum for the eligible months was divided by the proportion of the analysis period for which they were eligible to create a full-time equivalent measure. Regressions were weighted by the proportion of period for which the individual was eligible.[130]

2. Other measures

In this section we describe the methods for creating the control variables included in our analyses. Our analyses used multivariate regression models to adjust for differences across the analysis populations in demographics, geography, socioeconomic characteristics, Medicaid/Medicare enrollment, and health status.

The control variables included in the MMC Medicare impact regression models are listed in Table A-VII.4 along with the specifications for the variables. The control variables included in the MMC Medicaid pre-post analysis are listed in Table A-VII.5 along with the specifications for the variables. Note that when HCC and CDPS categorical variables had means of less than two percent, we did not include them as control variables.

Table A-VII.4. Impact analysis model control variable specifications—MMC

Variable name	Specification
Intervention period	Categorical variable indicating time period of observation. Categories include: baseline period (pre-enrollment; reference category); nine months post enrollment
Treatment indicator	Categorical variable indicating treatment status. Categories include: comparison group member (reference category); MMC intervention participants
Interaction between intervention period and treatment	Interaction between intervention period and treatment indicator variables
Time period	Categorical variable indicating the calendar quarter of the initial month of observation period. Categories range from: 1Q2012 (reference category) to 1Q2014
Race	Categorical variable indicating the individual's race. Categories include: White (reference category); Black; and Hispanic

[129] For example, if a person had third party insurance coverage in a particular month, they were not counted as eligible for the analysis in that month. In parallel, any services provided in that month were excluded from the analysis.

[130] Weights for comparison group members in the Medicare analysis were also based on the number of comparison group members associated with the same participant.

Variable name	Specification
Age	Continuous variable indicating age on the first day of the observation period
Age squared	Continuous variable measuring age as defined above squared
Sex	Categorical variable of member's sex. Categories include: female (reference category); male
Dually enrolled in Medicare and Medicaid	Indicator variable for dually enrolled in Medicare and Medicaid based on Medicare enrollment database indicator for dual status indicating dual status in one or more months during the observation period
Disabled	Indicator variable for original reason for Medicare entitlement based on disability
Pre-period Medicare enrolled	Indicator variable for availability of 12 months of FFS Medicare claims data prior to first day of observation period
HCC score	Indicator variables for HCC conditions in Medicare FFS claims data for 12 months prior to enrollment date
Bipolar disorder	Indicator variable for schizophrenia diagnosis on one inpatient or two or more outpatient claims (not including prescription drugs) with the relevant diagnoses in the two years prior to enrollment
Schizophrenia	Indicator variable for depression disorder diagnosis on one inpatient or two or more outpatient claims (not including prescription drugs) with the relevant diagnoses in the two years prior to enrollment
Depression	Indicator variable for bipolar disorder diagnosis on one inpatient or two or more outpatient claims (not including prescription drugs) with the relevant diagnoses in the two years prior to enrollment
Health service area	Indicator of health service area of member's residence. Categories include: Brooklyn (reference category), Philadelphia, Chicago, Pittsburg

Table A-VII.5. Pre-post model control variable specifications—MMC Medicaid analysis

Variable name	Specification
Intervention or baseline period	Categorical variable for each six-month intervention and baseline period except the first baseline period (the omitted category)
Age	Continuous variable indicating age as of the first month of each six-month analysis period
Sex	Categorical variable of member's sex. Categories include: female (reference category); male
Disabled	Categorical variable indicating whether member was eligible for Medicaid based on disability
Race	Categorical variable of member's race. Categories include: Black, Hispanic, Asian, other, African-American, and White (reference category)
Continuous Medicaid enrollment	Categorical variable indicating whether the enrollee had continuous enrollment in Medicaid for the 12 months prior to enrollment in the program
Calendar month flags	Vector of categorical variables that index the calendar month during which the first month of each six-month baseline or intervention period falls
CDPS flags	Flags indicating member's conditions based on the CDPS
AIDS, high	Categorical variable indicating whether member had AIDS, pneumocystis pneumonia, cryptococcosis, or Kaposi's sarcoma
Metabolic, high	Categorical variable indicating whether member had panhypopituitarism, pituitary dwarfism, non-HIV immunity deficiencies
Metabolic, medium	Categorical variable indicating whether member had kwashiorkor, marasmus, and other malnutrition, parathyroid, and adrenal gland disorders
Metabolic, very low	Categorical variable indicating whether member had other pituitary disorders, gout
Hematological, medium	Categorical variable indicating whether member had other hereditary hemolytic anemia, aplastic anemia, splenomegaly, agranulocytosis

Variable name	Specification
Hematological, low	Categorical variable indicating whether member had other white blood cell disorders, purpura, other coagulation defects
Substance abuse, low	Categorical variable indicating whether member had opioid, barbiturate, cocaine, amphetamine abuse or dependence, drug psychoses
Substance abuse, very low	Categorical variable indicating whether member had alcohol abuse, dependence, or psychosis
Infectious, high	Categorical variable indicating whether member had staphylococcal or pseudomonas septicemia, cytomegaloviral disease
Infectious, low	Categorical variable indicating whether member had poliomyelitis, oral candida, herpes zoster, parasitic intestinal infections
Cancer, medium	Categorical variable indicating whether member had mouth, breast or brain cancer, malignant melanoma, radiation or chemotherapy
Cancer, high	Categorical variable indicating whether member had lung cancer, ovarian cancer, secondary malignant neoplasms, leukemia, multiple myeloma
Diabetes Type 1, medium	Categorical variable indicating whether member had type 1 diabetes without complications or with neurological or ophthalmic complications
Diabetes Type 2, medium	Categorical variable indicating whether member had type 2 or unspecified diabetes with complications, proliferative diabetic retinopathy
Diabetes Type 2, low	Categorical variable indicating whether member had type 2 or unspecified diabetes without complications
Eye, low	Categorical variable indicating whether member had retinal detachment, choroidal disorders, vitreous hemorrhage
Eye, very low	Categorical variable indicating whether member had cataract, glaucoma, congenital eye anomaly, corneal ulcer
Cerebrovascular, low	Categorical variable indicating whether member had intracerebral hemorrhage, precerebral occlusion, hemiplegia, cerebrovascular accident
Cardiovascular, medium	Categorical variable indicating whether member had congestive heart failure, cardiomyopathy, tricuspid and pulmonary valve disease
Cardiovascular, low	Categorical variable indicating whether member had endocardial disease, myocardial infarction, angina, coronary atherosclerosis, or dysrhythmias
Cardiovascular, extra low	Categorical variable indicating whether member had hypertension
Gastrointestinal, medium	Categorical variable indicating whether member had regional enteritis and ulcerative colitis, chronic liver disease and cirrhosis, enterostomy
Nervous system, medium	Categorical variable indicating whether member had paraplegia, muscular dystrophy, multiple sclerosis
Nervous system, low	Categorical variable indicating whether member had epilepsy, Parkinson's disease, cerebral palsy, migraine, or cerebral degeneration
Genital, extra low	Categorical variable indicating whether member had uterine and pelvic inflammatory disease, endometriosis, or hyperplasia of prostate
Gastrointestinal, low	Categorical variable indicating whether member had ulcer, hernia, GI hemorrhage, intestinal infectious disease, or intestinal obstruction
Psychiatric, high	Categorical variable indicating whether member had schizophrenia
Psychiatric, medium	Categorical variable indicating whether member had bipolar affective disorder
Psychiatric, low	Categorical variable indicating whether member had other depression, panic disorder, or phobic disorder
Developmental disability, low	Categorical variable indicating whether member had mild or moderate mental retardation, Down's syndrome
Pregnancy, complete	Categorical variable indicating whether member had normal delivery, multiple delivery, delivery with complications
Pulmonary, medium	Categorical variable indicating whether member had other bacterial pneumonias, chronic obstructive asthma, adult respiratory distress syndrome

Variable name	Specification
Pulmonary, low	Categorical variable indicating whether member had viral pneumonias, chronic bronchitis, asthma, COPD, or emphysema
Renal, very high	Categorical variable indicating whether member had chronic renal failure, kidney transplant status or complications
Renal, low	Categorical variable indicating whether member had kidney infection, kidney stones, hematuria, urethral stricture, bladder disorders
Skeletal, medium	Categorical variable indicating whether member had chronic osteomyelitis, aseptic necrosis of bone
Skeletal, low	Categorical variable indicating whether member had rheumatoid arthritis, osteomyelitis, systemic lupus, or traumatic amputation of foot or leg
Skeletal, very low	Categorical variable indicating whether member had osteoporosis, musculoskeletal anomalies, thoracic and lumbar disc degeneration
Skin, low	Categorical variable indicating member had other chronic ulcer of skin
Skin, very low	Categorical variable indicating member had Cellulitis, burn, lupus erythematosus
Alzheimer's	Categorical variable indicating use of cholinesterase inhibitors, NMDA receptor antagonists
Anti-coagulants	Categorical variable indicating use of coumarin, heparin
Folate deficiency	Categorical variable indicating use of folic acid
CMV Retinitis	Categorical variable indicating use of eye antivirals
ICD-9 diagnosis category	Categorical variable indicating whether member had one of the following conditions based on ICD-9 diagnoses codes in the 24 months prior to enrollment in the program
Psychotic disorders	Categorical variable indicating claim diagnosis of: 293.81, 293.82, 293.83
Schizophrenia and related disorders	Categorical variable indicating claim diagnosis of: 295.XX including 295 with no digits after or 295.00
Bipolar disorders	Categorical variable indicating claim diagnosis of: 296.00, 296.01, 296.02, 296.03, 296.04, 296.05, 296.06, 296.10, 296.11, 296.12, 296.13, 296.14, 296.15, 296.16, 296.40, 296.41, 296.42, 296.43, 296.44, 296.45, 296.46, 296.50, 296.51, 296.52, 296.53, 296.54, 296.55, 296.56, 296.60, 296.61, 296.62, 296.63, 296.64, 296.65, 296.66, 296.7, 296.80, 296.81, 296.82, 296.89, 296.90, 296.99
Depressive disorders	Categorical variable indicating claim diagnosis of: 296.20, 296.22, 296.23, 296.24, 296.25, 296.26, 296.30, 296.32, 296.33, 296.34, 296.35, 296.36,
Disturbance of emotions specific to childhood and adolescence	Categorical variable indicating claim diagnosis of: 301.13

REFERENCES

Babyak, Michael A., C.R. Snyder, and Lauren Yoshinobu. "Psychometric Properties of the Hope Scale: A Confirmatory Factor Analysis." *Journal of Research in Personality*, vol. 27, no. 2, June 1993, pp. 154–169.

Bernstein, Rebecca S., Linda N. Meurer, Ellen J. Plumb, and Jeffrey L. Jackson. "Diabetes and Hypertension Prevalence in Homeless Adults in the United States: A Systematic Review and Meta-Analysis." *American Journal of Public Health*, vol. 105, no. 2, February 2015, pp. e46–e60.

Brazier, John E., and Jennifer Roberts. "The Estimation of a Preference-Based Measure of Health from the SF-12." *Medical Care*, vol. 42, no. 9, September 2004, pp. 851–859.

Brouwer, Danny, Rob R. Meijer, Anke M. Weekers, and Joost J. Baneke. "On the Dimensionality of the Dispositional Hope Scale." *Psychological Assessment*, vol. 20, no. 3, September 2008, pp. 310–315.

Creamer, Mark, Meaghan L. O'Donnell, Isla Carboon, Virginia Lewis, Konstancja Densley, Alexander McFarlane, Derrick Silove, and Richard A. Bryant. "Evaluation of the Dispositional Hope Scale in Injury Survivors." *Journal of Research in Personality*, vol. 43, no. 4, August 2009, pp. 613–617.

Daeppen, Jean-Bernard, Marc-Antoine Krieg, Bernard Burnand, and Bertrand Yersin. "MOS-SF-36 in Evaluating Health-related Quality of Life in Alcohol-Dependent Patients." *The American Journal of Drug and Alcohol Abuse*, vol. 24, no. 4, 1998, pp. 685–694.

Dehejia, Rajeev H., and Sadek Wahba. "Propensity Score-Matching Methods for Nonexperimental Causal Studies." *The Review of Economics and Statistics*, vol. 84, no. 1, February 2002, pp. 151–161.

Dozois, David J.A., Henny A. Westra, Kerry A. Collins, Tak S. Fung, and Jennifer K.F. Garry. "Stages of Change in Anxiety: Psychometric Properties of the University of Rhode Island Change Assessment (URICA) Scale." *Behaviour Research and Therapy*, vol. 42, no. 6, June 2004, pp. 711–729.

Field, Craig A., Bryon Adinoff, T. Robert Harris, Samual A. Ball, and Kathleen M. Carroll. "Construct, Concurrent, and Predictive Validity of the URICA: Data from Two Multi-Site Clinical Trials." *Drug Alcohol Dependence*, vol. 101, no. 1-2, 2009, pp. 115–123.

Findler, Marianne, Joshua Cantor, Lisa Haddad, Wayne Gordon, and Teresa Ashman. "The Reliability and Validity of the SF-36 Health Survey Questionnaire for Use with Individuals with Traumatic Brain Injury." *Brain Injury*, vol. 15, no. 8, 2001, pp. 715–723.

Franz, Marion J., Jackie L. Boucher, Stephanie Rutten-Ramos, and Jeffrey J. VanWormer. "Lifestyle Weight-Loss Intervention Outcomes in Overweight and Obese Adults with Type 2 Diabetes: A Systematic Review and Meta-Analysis of Randomized Clinical Trials." *Journal of the Academy of Nutrition and Dietetics,* vol. 115, no. 9, September 2015, pp. 1447–1463.

Henderson, Melinda J., Karen K. Saules, and Luke W. Galen. "The Predictive Validity of the University of Rhode Island Change Assessment Questionnaire in a Heroin-Addicted Polysubstance Abuse Sample." *Psychology of Addictive Behaviors*, vol. 18, no. 2, June 2004, pp. 106–112.

Koh, Katherine A., Jessica S. Hoy, James J. O'Connell, and Paul Montgomery. "The Hunger-Obesity Paradox: Obesity in the Homeless." *Journal of Urban Health*, vol. 89, no. 6, December 2012, pp. 952–964.

McHorney, Colleen A., John E. Ware, Jr., and Anastasia E. Raczek. "The MOS 36-Item Short-Form Health Survey (SF-36): II. Psychometric and Clinical Tests of Validity in Measuring Physical and Mental Health Constructs." *Medical Care*, vol. 31, no. 3, 1993, pp. 247–263.

National Committee for Quality Assurance (NCQA). The State of Health Care Quality 2015. Washington, DC: NCQA, 2015.

National Institutes of Health. "Classification of Overweight and Obesity by BMI, Waist Circumference, and Associated Disease Risks." Accessed August 29, 2016 at [http://www.nhlbi.nih.gov/health/educational/lose_wt/BMI/bmi_dis.htm].

Nidecker, Melissa, Carlo C. DiClemente, Melanie E. Bennett, and Alan S. Bellack. "Application of the Transtheoretical Model of Change: Psychometric Properties of Leading Measures in Patients with Co-Occurring Drug Abuse and Severe Mental Illness." *Addictive Behaviors,* vol. 33, no. 8, August 2008, pp. 1021–1030.

Pahwa, Rohini, Maanse Hoe, Shelley Levin, and John S. Brekke. (January 2012). "Cross-Ethnic and Cross-National Measurement Invariance of Brief Symptom Inventory-18 in Severe Mental Illness." Presented at the Sixteenth Annual Meeting of the Society for Social Work Research, Washington, DC, January 2012.

Petrou, Stavros, and Christine Hockley. "An Investigation into the Empirical Validity of the EQ-5D and SF-6D Based on Hypothetical Preferences in a General Population." *Health Economics,* vol. 14, no. 11, November 2005, pp. 1169–1189.

Rosenbaum, Paul R., and Donald B. Rubin. "The Central Role of the Propensity Score in Observational Studies for Causal Effects." *Biometrika*, vol. 70, no. 1, 1983, pp. 41–55.

Snyder, C. R., Cheri Harris, John R. Anderson, Sharon A. Holleran, Lori M. Irving, Sandra T. Sigmon, Lauren Yoshinobu, June Gibb, Charyle Langelle, and Pat Harney. "The Will and the Ways: Development and Validation of an Individual-Differences Measure of Hope." *Journal of Personality and Social Psychology*, vol. 60, no. 4, April 1991, pp. 570–585.

Tsai, Jack, and Robert A. Rosenheck. "Obesity Among Chronically Homeless Adults: Is it a Problem?" *Public Health Reports*, vol. 128, no. 1, January-February 2013, pp. 29–36.

University of Maryland, Baltimore County. "URICA – Readiness Score: Calculating and Understanding the Readiness Score." Accessed September 8, 2016, at [http://habitslab.umbc.edu/urica-readiness-score/].

Wang, Jichuan, Brian C. Kelly, Brenda M. Booth, Russel S. Falck, Carl Leukefeld, and Robert G. Carlson. "Examining Factorial Structure and Measurement Invariance of the Brief Symptom Inventory (BSI)-18 among Drug Users." *Addictive Behaviors*, vol. 35, no. 1, 2010, pp. 23–29.

www.mathematica-mpr.com

Improving public well-being by conducting high quality, objective research and data collection

PRINCETON, NJ ■ ANN ARBOR, MI ■ CAMBRIDGE, MA ■ CHICAGO, IL ■ OAKLAND, CA ■ WASHINGTON, DC

MATHEMATICA
Policy Research