MATHEMATICA
Policy Research

REPORT

SECOND ANNUAL REPORT

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

March 1, 2016

Henry Ireys
Jonathan Brown
Ellen Bouchery
Liz Babalola
Michael Barna
Crystal Blyler
Dan Friend
Christine Fulton
Matt Kehn
Tricia Higgins
Jasmine Little
Jennifer Lyons
Jessica Nysenbaum
Laura Ruttner
Allison Siegwarth
Kate Stewart
Michaela Vine

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

FIGURES

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 Innovations Award (HCIA) initiative employed 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 on innovative care models for individuals with mental health and substance use disorders (Table ES.1). The three-year award period began in July 2012. For three awardees (Felton, ICSI, and KMHC), the project period ended on June 30th, 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 allowed the awardees to complete their own evaluations and transition their projects to more sustainable sources of funding.

The 10 projects in this group involved some of the same activities (for example, training staff to coordinate care and using information technology to monitor care) but they focused on different subgroups within this 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 housing the homeless. The number of participants enrolled in these projects varied widely, depending on the specific objectives of the awardees.

A rigorous, multifaceted evaluation of these interventions will 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 because individuals with mental health and substance use disorders are among the most vulnerable and costly groups of Medicare, Medicaid, and CHIP beneficiaries. Any enduring solution to improving health care and lowering costs cannot ignore them.

Table ES.1. Behavioral health and substance abuse awardees

Awardee[a]	Overview of intervention	Intervention population	Dollars awarded	Enrollment goal (percent of goal reached)
Center for Health Care Services (CHCS)	Integrate 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)	Implement an integrated treatment model to improve treatment 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)	Improve treatment for schizophrenia through training, care management, and new technology	Patients with schizophrenia who were recently discharged from the hospital and are receiving care at community treatment centers in one of eight states	$9,380,855	770 (66)
Fund for Public Health in New York (FPHNY)	Provide crisis intervention services to facilitate early engagement and continuity of care, combining community-based care with access to primary care	Individuals in Manhattan, Brooklyn, the Bronx, and Queens who have been diagnosed with psychosis	$17,608,085	2,232 (63)
HealthLinkNow (HLN)	Provide 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)	Implement collaborative care management for patients with depression and diabetes or cardiovascular disease	High-risk adult patients with Medicare or Medicaid coverage in one of seven states who have active depression and uncontrolled diabetes or cardiovascular disease or both	$17,999,635	2,704 (100)
Kitsap Mental Health Services (KMHS)	Integrate 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)	Coordinate mental and physical health care through advanced health information technology	Adults with serious mental illness living in southwest Brooklyn	$14,842,826	500[d] (100)

Awardee[a]	Overview of intervention	Intervention population	Dollars awarded	Enrollment goal (percent of goal reached)
ValueOptions (ValueOptions)	Provide care coordination	Plan members in Massachusetts with two or more detoxification admissions	$2,760,737	1,492[b] (82)
Vinfen Corporation (Vinfen)	Integrate 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 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 usually 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 methods

In September 2013, CMMI contracted with Mathematica Policy Research to evaluate these 10 projects. To conduct our evaluation, we are using a mixed-methods approach that involves collecting and analyzing quantitative and qualitative data. Overall, we designed our evaluation to achieve three interrelated goals:

1. Use quantitative and qualitative data to address evaluation questions developed by CMMI, with a focus on four quantitative measures of program effectiveness (use of emergency department (ED) services, rates of hospitalization and hospital readmission, and total Medicare and Medicaid expenditures)
2. "Tell the story" of each awardee through the development of narratives describing the proposed plan, implementation processes, and project outcomes
3. Derive cross-cutting lessons learned about successful projects based on a synthesis of findings across awardees

As we work toward these goals, we are addressing a series of key research questions, some of which apply to all HCIA awardees and some of which are focused specifically on this group. Examples of such questions include:

- How do the projects integrate physical and mental health services?
- Which components of care coordination are most important and effective?
- What role does organizational leadership play in the success of the interventions and why?
- How does the staffing turnover rate for certain roles (such as patient navigators or peer support specialists) compare with that of other health care workers?

In this report, we focus on questions selected in collaboration with our contracting officer representative as the most feasible and important to examine at this stage of the evaluation (Table ES.2). The questions are grouped according to the evaluation domains developed by CMMI; we also indicate the sources of data we are using to address the questions.

The questions we address in this report provide the structure for presenting:

- Final enrollment figures for all awardees
- Preliminary findings about the effectiveness of the implementations as perceived by key stakeholders
- Staff perceptions of the interventions' impacts on the health of participants
- For five awardees, preliminary quantitative estimates of the interventions' impacts on service use and costs of care
- Insights into the contextual factors that shaped the implementation of the awardees' interventions

To address the selected questions, we analyzed (1) relevant extracts of information from our qualitative database, which was generated from interviews conducted during our first and second round of site visits with the awardees (March through June, 2014 and 2015) and (2) quantitative information on service use and costs of care for participants in five awardees' programs and for members of comparison groups, which allowed us to develop preliminary impact estimates.[1]

Estimating program impacts requires a rigorous evaluation design to isolate the effects of the intervention from effects that could be caused by other factors. Ideally, all awardees would have used a randomized experimental design in which providers or participants were randomly assigned to an intervention or comparison group. Only CHCS used such a design. For the other awardees, we sought to identify comparison groups, primarily by using propensity score matching techniques. We were able to use these techniques to develop comparison groups for three awardees (KMHS, MMC, and ValueOptions).

To conduct our impact analyses for these four awardees (that is, CHCS, KMHS, MMC, and ValueOptions), we used data that covered a pre-intervention period and a post-intervention period that was long enough to allow the intervention to have its intended effects. Pre-and post-intervention data for both participants in the intervention group and members of a comparison group allow us to estimate impacts using a "difference-in-differences" (DD) design, which can account for differences between the intervention and comparison groups that might otherwise bias our estimates.[2]

[1] We received the analytic files at different times for different awardees based on the availability of data. We obtained Medicare data extracts for KMHS, as well as ValueOptions data, in February 2015. We obtained Medicare data extracts for the MMC intervention group members from late March through early April, 2015, and obtained data for the comparison groups in May 2015. CHCS provided data in April 2015.

[2] To the extent possible, we used CMMI's specifications for four core outcome measures: total expenditures, hospitalizations, hospital readmissions, and emergency department (ED) visits. Data for these measures were not available for CHCS; instead, we used scores on several standard symptom checklists as our outcome measures. We

Table ES.2. Research questions and data sources

Research question	Data source(s)			
	Qualitative	Enrollment reporting	Medicare and/or Medicaid claims	Outcome data collected by awardees
Implementation effectiveness				
What was the target population and who was enrolled?	X	X		
To what extent was implementation timely, conducted as planned, and responsive to site-level constraints?	X	X		
How did unexpected events support or conflict with successful implementation of the innovation?	X			
Program effectiveness				
What has been the impact of the program on total cost of care?			X	
What has been the impact of the program on hospitalization rates?			X	
What has been the impact of the program on hospital readmission rates?			X	
What has been the impact of the program on ED visit rates?			X	
What has been the impact of the program on awardee-specific outcomes?			X[a]	X[b]
From the perspective of key informants, to what extent does the intervention improve desired health outcomes for participants?	X			
Context				
To what extent did organizational features support or conflict with implementation?	X			
To what extent did the policy and political environment support or conflict with implementation?	X			

[a] For ValueOptions, we analyze short-term residential treatment stays and days of intensive day treatment.

[b] For CHCS, we analyze measures of psychological stress, capacity for self-management, feelings of hope, and capacity for life change.

Our impact estimates are quite preliminary because for most awardees they are based on only a portion of their total enrolled population, including some individuals who have not been in the program for long. In addition, with the exception of CHCS and ValueOptions, the participants included in the intervention group are limited to a subsample of participants for whom fee-for-service Medicare or Medicaid data were available; the outcomes of these participants may not reflect the outcomes of all the awardees' participants. Detailed descriptions

investigated carefully the possibility of generating a comparison group for ICSI but, for reasons noted in Chapter VIII, elected not to do so.

of our methods for generating comparison groups and conducting our impact estimates are provided in the individual awardee chapters and included in Appendix A.

For six awardees (Feinstein, HLN, ICSI, FPHNY, Felton, and Vinfen), the data we need to construct comparison groups and estimate program impacts are not available at this time. Challenges to conducting impact analyses include the small number of participants in some programs and incomplete, unreliable, or substantially delayed Medicaid data for some states. In Chapter I and in the individual awardee chapters, we provide detailed descriptions of the circumstances that prevent us from conducting rigorous quantitative analyses at this point in the evaluation. For one of these awardees (ICSI), we were able to conduct pre-post analyses for the intervention group.

In the upcoming year, we expect to conduct impact analyses for FPHNY as we obtain new data. We hope also to conduct such analyses for Feinstein and HLN, although analyses may be difficult because participants in these awardees' programs (1) are distributed across several states, not all of which have complete and reliable Medicaid claims data; (2) are covered by a variety of insurance types, which limits the number of participants with any one type of insurance; and (3) have wide cross-state variation in their demographics, which can limit the opportunities to identify suitable comparison groups. If we cannot use a DD design with these awardees, we will conduct pre-post analyses without a comparison group.

We are working with Felton to obtain reliable data to conduct pre-post analyses of its program's participants, but we do not anticipate obtaining data that would allow us to identify a comparison group. We do not anticipate being able to acquire from Vinfen the individual-level data needed to support rigorous quantitative analyses.

Several awardees plan to conduct analyses of data that they have collected on program participants (and, in some cases, comparison groups) and to publish reports about their programs' outcomes. Awardees' own analyses may focus on evaluation questions and use data sources and measures different from those included in this report. For example, several awardees may examine 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 are focusing primarily on service use and cost outcomes—outcomes that can be assessed similarly across the awardees and that are priorities for CMMI. Because the awardees are using different sources of data and are focusing on different outcome variables, they may reach conclusions that differ substantially from the ones we present in this and future reports.

We developed this report, the second of three annual reports we will submit to CMS, during June and July 2015—just as most awardees were ending their projects. The results presented here should be considered preliminary because of the following reasons:

- The impact and pre-post analyses are limited to a subset of the full population of the awardee's participants, and they cover a limited period of time; analytic results could change as we obtain additional data and as participants have more exposure to the program.
- Substantial lags in the availability of Medicaid claims and administrative data mean that we do not have the quantitative information we need to estimate the full impacts of the interventions for Medicaid beneficiaries in most of the awardees' programs.

- In May 2015, we collected extensive qualitative data and conducted a survey of the awardee workforce; future reports will include findings from analyses of these data.
- The awardees' final reports were not submitted to CMS in time for us to use them as a source of data for this report.

In our third annual report, which we will submit to CMMI in August 2016, we will provide a more comprehensive set of findings and update the preliminary impact estimates reported here. Specifically, we will be able to:

- Analyze Medicare and Medicaid data available in late 2015 and early 2016
- Conduct further analyses of recently collected qualitative and survey data
- Review final reports from the awardees

Selected findings

In Chapter II, we synthesize the information presented in the awardee-specific chapters to present cross-cutting findings. Here, we highlight findings that are particularly noteworthy. *As noted, the results are preliminary and may change in the future when more data are available.* Because of their preliminary nature, we have refrained from speculating on the reasons for these outcomes. In our next annual report, when we expect our analyses to yield more certain outcomes, we will integrate qualitative data into our analyses to describe key factors contributing to the outcomes.

- **Although the majority of awardees implemented their interventions relatively close to their planned schedule, all experienced challenges that affected program implementation. These challenges included:**
  - Hiring and retaining behavioral health specialists.
  - Clearing the significant bureaucratic and legal hurdles involved in implementing a program encompassing multiple partners and providers at various sites; these hurdles included developing business associate agreements, obtaining multiple IRB reviews, and devising efficient reporting mechanisms.
  - Adopting technology—such as a new platform for electronic health records or a telehealth device—which proved to be more resource- and labor-intensive than some awardees expected.
  - Balancing program standardization with some allowances for customizing the design based on site-specific needs.
  - Incorporating the potential for continuous adaptation of program design in response to challenges and lessons learned.
- **Based on results for four awardees, their interventions had varying effects on Medicaid and Medicare expenditures.**
  - For the first year post-intervention, impact estimates for KMHS, MMC, and ValueOptions indicate no significant effect of the intervention on expenditures. Pre-post

analyses for ICSI indicate a statistically significant (p-value <.01) increase in Medicare expenditures during the first intervention year relative to baseline.

- Data for impact estimates for the second year post-intervention were available only for KMHS; analyses indicated a decrease in total Medicare expenditures per patient when comparing trends over time for both program participants in the intervention group and members of the comparison group (p-value <.01).
- Impact estimates for expenditures are not available for CHCS.

- **Based on results for three awardees, their interventions had varying effects on hospitalizations.**
  - For the first year post-intervention, pre-post analyses for ICSI indicate a statistically significant (p-value = .03) increase of 0.5 hospitalizations per patient.
  - For the first year post-intervention, the impact estimate for MMC indicates no statistically significant difference in hospitalizations.
  - For both the first and second intervention year, the estimates for KMHS indicate a marginally significant (p-value = 0.07) decrease of 0.1 hospitalizations per patient.
  - Impact estimates for hospitalizations are not available for the other awardees.
- **KMHS' intervention had no impact on hospital readmissions in either the first or second year post-intervention.**
  - KMHS was the only awardee for which it was possible to estimate hospital readmissions.
- **Based on results for four awardees, their interventions either decreased ED visit rates or had no effects.**
  - For the first year post-intervention, the impact estimate for ValueOptions indicates a statistically significant (p-value = 0.02) decrease of 2.1 ED visits per patient.
  - For the first year post-intervention, impact estimates for KMHS and MMC indicate no statistically significant impact on ED visits. Likewise, there was no impact on ED visits for KMHS in the second year post-intervention. Pre-post analyses for ICSI showed no significant differences between baseline and the first year.
  - Impact estimates for ED visits were not available for the other awardees.
- **Our analyses found no impacts of CHCS' or ValueOptions' interventions on awardee-specific outcomes.**
  - For CHCS, there were no impacts on measures of psychological stress, hopefulness, and participant capacity for self-management and behavioral change.
  - For ValueOptions, we there were no impacts on intensive day treatment and residential treatment services.
- **Organizational culture and experience with intervention models played an important role in implementing nearly all awardee programs.**

- Respondents from Felton, KMHS, MMC, and Vinfen identified the mission and values of their organizations as key factors supporting implementation.
- According to staff at ICSI, sites that were experienced with their intervention model were able to implement it more efficiently than other sites did.
- At times, the prevailing organizational culture slowed some awardees' implementation progress (for example, at CHCS, FPHNY, ICSI, and MMC), particularly when programs required frontline staff to adjust how they approached patient care.
- Some clinical staff in the FPHNY and ICSI programs initially resisted following the practices of the new program models because they conflicted with staff's training in more traditional models of care.
- Awardees overcame these barriers by providing more training to program staff and creating more opportunities for communication and feedback.

- **Successful program implementation often meant engaging with community stakeholders such as social service providers, county health staff, and local hospitals.**
  - Felton established an effective referral system that included local school districts and county social service providers, some of whom were initially apprehensive.
  - Some of ICSI's sites leveraged existing partnerships—including relationships with local YMCAs, community colleges, and fire departments—to better implement the intervention in their communities.
  - At least one awardee (FPHNY) had difficulty generating appropriate referrals at some sites, and struggled to engage external stakeholders such as hospitals and emergency departments and raise their awareness of the program.
- **State policies and programs affected program implementation for many awardees.**
  - Felton respondents highlighted California's Mental Health Services Act of 2004, which provides funding for prevention and early intervention services, as creating an incentive for counties to adopt or collaborate with programs like the one they implemented under the HCIA initiative.
  - FPHNY respondents noted that New York State's Medicaid redesign provides an opportunity to sustain the FPHNY intervention.
  - HLN respondents cited several recent changes to state policies as potentially supporting the use of telemedicine, a key component of their intervention.
  - MMC leaders noted that their HCIA-funded program developed alongside New York State's health homes program, which will continue to support program services for most MMC participants after the award period.

Conclusion

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. This report provides preliminary findings about the implementation and impact of these awardees' programs—findings that may

help CMS consider new policies or programs for improving care and decreasing costs for this priority population. These findings also can provide a context for understanding results that awardees may report through their own analyses.

We expect to build on our findings with further analyses of both quantitative and qualitative data in the upcoming year. Specifically, we anticipate obtaining additional Medicaid data for several awardees, which will allow us to conduct more comprehensive analyses than were possible for this report. Furthermore, we will be analyzing qualitative data pertaining to awardees' intentions for sustaining key elements of their programs. In addition, we expect to integrate qualitative and survey data to answer evaluation questions related to awardees' workforce development and training efforts. Finally, in our third annual report, we expect to synthesize findings from these diverse analyses to provide a final summary of the implementation lessons and outcomes of these programs.

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. The models address groups of beneficiaries with poor clinical outcomes or heavy utilization of services. Based on evidence about the implementation and outcomes, CMMI may promote replication of the most promising models of care. Consequently, rigorous evaluation of the HCIA initiative is critical to achieving CMMI's mission.

B. Overview of the behavioral health awardees

Throughout the health care system, leaders are responding to the need for cost efficiencies, care coordination for high-risk populations, and the growing emphasis on quality measurement and public reporting. In the first round of the HCIA initiative, 10 awardees implemented interventions that focus on individuals with mental health and substance use disorders (Table I.1). The projects share some cross-cutting themes (for example, training staff to coordinate care and using information technology to monitor care) but they focused on different subgroups within this broad 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, and there was a wide variation in the number of participants enrolled in the programs.

Most awards began in early July of 2012. For three awardees (Felton, ICSI, and KMHS) the project period ended on June 30th, 2015. Four awardees (Feinstein, HLN, MMC, and Vinfen) received four-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 will allow the awardees to complete their own evaluations and transition their projects to more sustainable sources of funding.

Several awardees plan to conduct analyses of data that they have collected on program participants (and, in some cases, comparison groups) and to publish reports about their programs' outcomes. Awardees' own analyses may focus on evaluation questions and use data sources and measures different from those included in this report. For example, several awardees may examine 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 are focusing primarily on service use and cost outcomes—outcomes that can be assessed similarly across the awardees and that are priorities for CMMI. Because the awardees are using different sources of data and are focusing on different outcome variables, they may reach conclusions that differ substantially from the ones we present in this and future reports.

Table I.1. Behavioral health and substance abuse awardees

Awardee (name abbreviation used in report)	Overview of intervention (dollars awarded[a])	Intervention population (target number of direct participants[b])
Center for Health Care Services (CHCS)	Integrate primary care clinic into behavioral health service setting ($4,557,969)	Adults in San Antonio, Texas who are homeless (260)
The Felton Institute (Felton)	Implement 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)	Improve intervention for 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)	Provide 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 (2,232)
HealthLinkNow (HLN)	Provide behavioral care services via telemedicine to individuals in rural areas lacking 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)	Implement collaborative care management model for patients with active depression and uncontrolled diabetes or cardiovascular disease ($17,999,635)	High-risk adult patients with Medicare or Medicaid coverage in one of seven states who have active depression and uncontrolled diabetes or cardiovascular disease or both (2,704)
Kitsap Mental Health Services (KMHS)	Integrate primary health care for individuals with severe mental illness ($1,858,437)	Adults with severe mental illness and one comorbidity; children with severe emotional disturbance and one physical comorbidity; Kitsap County, Washington (not applicable[c])
Maimonides Medical Center (MMC)	Coordinate mental and physical health care through advanced health information technology ($14,842,826)	Adults with serious mental illness living in southwest Brooklyn (500)
ValueOptions (ValueOptions)	Provide care coordination ($2,760,737)	Plan members in Massachusetts with two or more detoxification admissions (1,492)
Vinfen Corporation (Vinfen)	Integrate health care services into existing behavioral health outreach teams in community ($2,942,962)	Individuals in Boston with serious mental illness (400)

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

[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 the Lewin Group.

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

C. Evaluation goals, data, and methods

In September 2013, CMMI contracted with Mathematica Policy Research to evaluate the 10 projects in Table I.1. To conduct the evaluation, we are using a mixed-methods approach that synthesizes qualitative data on implementation with quantitative data on outcomes for each awardee. Overall, we designed the evaluation to achieve three interrelated goals:

1. Use quantitative and qualitative data to address about 100 evaluation questions developed by CMMI, with a focus on four core quantitative measures of program effectiveness (use of emergency department services, rates of hospitalization and hospital readmission, and total Medicare and Medicaid expenditures)
2. "Tell the story" of each awardee through the development of narratives describing the proposed plan, implementation processes, and project outcomes
3. Derive cross-cutting lessons learned about successful projects based on a synthesis of findings across awardees

Our evaluation draws on four types of data. The first is enrollment numbers, which we obtained primarily from the website operated by the Lewin Group for CMMI. We used these numbers to determine each awardee's progress toward its final enrollment goal. Awardees specified an enrollment goal in their early reports; many revised these goals as they implemented their projects.

The second type is Medicare and Medicaid claims data, which we use to calculate CMMI's four core measures and other outcomes in order to examine program effectiveness and impacts. We were not able to use these claims data for all awardees because of lags in the availability of data, small sample sizes, and incomplete or unreliable data that would threaten the reliability and/or generalizability of the findings.

The third category of data includes assessment, survey, and administrative data obtained from the awardee. To the extent possible, we use these data to calculate other measures of program effectiveness, such as changes in participants' symptoms and functional status.

The final data source includes qualitative information derived from interviews and focus groups conducted during site visits to awardees in the spring of 2014 and 2015. During these sites visits, we met with awardee leaders and staff, program participants, and other stakeholders to gather information about the implementation process and their experiences with various components of the programs. During the site visits, we conducted both semi-structured interviews (with program leaders, frontline staff, and other stakeholders) and focus groups (with staff and, where possible, both program participants and non-participants to understand differences in their receipt of care and experiences with care).

D. Overview of report

This second annual report addresses the research questions listed in Table I.2. These questions are about the context and effectiveness of the implementation and the overall effectiveness of the program.

Table I.2. Research questions and data sources

Research question	Data source(s)			
	Qualitative	Enrollment reporting	Medicare and/or Medicaid claims	Outcome data collected by awardees
Implementation effectiveness				
What was the target population and who was enrolled?	X	X		
To what extent was implementation timely, conducted as planned, and responsive to site-level constraints?	X	X		
How did unexpected events support or conflict with successful implementation of the innovation?	X			
Program effectiveness				
What has been the impact of the program on total cost of care?			X	
What has been the impact of the program on hospitalization rates?			X	
What has been the impact of the program on hospital readmission rates?			X	
What has been the impact of the program on ED visit rates?			X	
What has been the impact of the program on awardee-specific outcomes?			X[a]	X[b]
From the perspective of key informants, to what extent does the intervention improve desired health outcomes for participants?	X			
Context				
To what extent did organizational features support or conflict with implementation?	X			
To what extent did the policy and political environment support or conflict with implementation?	X			

[a] For ValueOptions, we analyze short-term residential treatment stays and days of intensive day treatment.

[b] For CHCS, we analyze measures of psychological stress, capacity for self-management, feelings of hope, and capacity for life change.

In collaboration with our contracting officer representative, we selected these research questions to provide CMMI, HCIA awardees, and the broader behavioral health community with information that may be useful for refining and replicating these and similar interventions. The findings in this report build on our past quarterly reports (submitted in March, June, September, and December 2014, and March, June, and August 2015) and our first annual report (submitted in October 2014).

As we have noted, this report is the second of three that we will provide to CMMI. As more data become available, we expect to update our analyses of the programs' effectiveness. We will

submit a third annual report in August 2016 and revise that report, with any updated findings, for submission in August 2017.

For our analyses of the programs' effectiveness, we drew on quantitative data obtained in the late winter and spring of 2015, which we used to assess program effects on service utilization, expenditures, and other outcomes.[3] Estimating program impacts requires a rigorous evaluation design to isolate the effects of the intervention from the effects of other factors. In the absence of an experimental design, in which providers or participants are randomly assigned to the intervention or comparison groups, we sought to identify alternative comparison groups for our evaluation. Estimating impacts also requires drawing on data that cover the time period before the intervention was implemented (or before participants were enrolled) as well as data that cover a follow-up period long enough to allow the intervention to have its intended effects. Such data facilitate the estimation of impacts by allowing us to use a "difference-in-differences" design, which can account for differences between the intervention and comparison groups that otherwise might bias estimates of the interventions' impacts.

At this point in the evaluation, we cannot estimate impacts for all the awardees, because we do not have complete and reliable quantitative data available for many of them. In addition, data are not available to construct comparison groups for all awardees. The challenges vary somewhat by awardee, but common challenges include small numbers of participants enrolled in Medicare and/or Medicaid, incomplete or unreliable Medicaid data for some states, and considerable lags in the availability of Medicaid data for some states.

After carefully assessing the strengths and limitations of the quantitative data currently available for each awardee, we decided in collaboration with CMMI that we would not report quantitative outcomes on program effectiveness for Feinstein, HLN, FPHNY, Felton, or Vinfen at this stage of the evaluation. Table I.3 describes key data limitations for these awardees. In the future, we plan to report these quantitative outcomes for FPHNY because we expect to obtain data that are based on a larger and more representative sample. Although we hope to also report quantitative data on program effectiveness for Feinstein and HLN in the future, it may be difficult to do so because of constraints in the availability of data, sample sizes, and variation in the insurance coverage and geographic distribution of participants. As a result, the generalizability of our findings could be severely limited. We are working with Felton to obtain reliable data to conduct pre-post analyses of the intervention participants, but we do not anticipate obtaining data that would allow us to identify a comparison group.[4] We do not

[3] We received the analytic files at different times for different awardees based on the availability of data. We obtained Medicare data extracts for KMHS, as well as ValueOptions data, in February 2015. We obtained Medicare data extracts for the ICSI and MMC intervention group members from late March through early April, 2015, and obtained data for the MMC comparison group in May 2015. CHCS provided data in April 2015.

[4] There is no viable contemporaneous comparison group that can be drawn from the assessment data provided by Felton or from data obtained from the county mental health departments. Felton reports that most eligible individuals in the intervention counties have been or will be enrolled in the intervention, thereby limiting our ability to identify a comparison group. As described in detail in Section VIII, we also were unable to identify a comparison group for ICSI.

anticipate being able to acquire from Vinfen the individual-level data needed to support rigorous quantitative analyses.

We report program effectiveness measures for five awardees. Table I.4 summarizes the basic design we used to estimate program impact for each awardee, including a brief description of the intervention population, comparison group, measures, and key limitations. Further details about the methods and data sources uses for the analyses are included in the awardee-specific chapters and technical appendices.

Table I.3. Key data limitations for awardees without program effectiveness measures

Awardee	Currently available data source for estimating program effectiveness	Estimate of size of analytic sample using currently available data (Medicaid population, FFS Medicare population)	Rationale for not estimating program effectiveness at this stage of evaluation
Feinstein	Alpha-MAX for NY through Q3 2013 Medicare files through 2014	13, 58 [a]	Pre-post analyses would be substantially underpowered because of small sample sizes for both Medicare and Medicaid populations; intervention participants are in multiple states.
HLN	Alpha-MAX for WA, MT, WY through Q3 2013 Medicare files through 2014	30, 100 [b]	Pre-post analyses would be substantially underpowered because of small sample sizes for both Medicare and Medicaid populations.
FPHNY	Alpha-MAX for NY through Q3 2013	34, 20 [c]	Pre-post analyses would be substantially underpowered because of small sample size.
Felton	Awardee assessment and encounter data and county mental health system data[d]	220, 0	No data were available to develop a comparison group. Further examination of the reliability and completeness of the data sources is necessary before conducting pre-post analyses of the intervention group.
Vinfen	Awardee assessment and encounter data	250, 100	No data were available to develop a comparison group. There were no individual-level claims data available on program effectiveness measures for intervention participants.[e]

[a] By April 1, 2015, Alpha-MAX data for NY were available for analysis for services through September 30, 2013. Allowing for three months of intervention exposure before the end of the analysis period, only 13 Medicaid participants were enrolled as of June 30, 2013, based on the Q4 program report that Feinstein submitted to the Lewin Group's reporting website. All 13 Medicaid participants were at the NY site. Data were available for Medicare services through September 30, 2014 for this report. However, based on the Q10 program report that Feinstein submitted to the Lewin Group, Feinstein had only 58 fee-for-service Medicare enrolled participants through September 30, 2014.

[b] By April 1, 2015, Alpha-MAX data for MT, WY, and WA were available for services provided through September 30, 2013. Allowing for three months of intervention exposure before the end of the analysis period, only 30 Medicaid participants in the three states were enrolled as of June 30, 2013 based on the Q4 program report that HLN submitted to the Lewin Group reporting website. Data were available for Medicare services through September 30, 2014 for this report. However, based on data provided by HLN, only 100 fee-for-service Medicare participants were enrolled in the three states as of September 30, 2014.

[c] By April 1, 2015 Alpha-MAX data for NY were available for analysis for services through September 30, 2013. Allowing for three months of intervention exposure before the end of the analysis period, only 34 Medicaid participants were enrolled as of June 30, 2013, based on data provided by FPHNY. The majority of the other participants enrolled at that time were uninsured. Data were available for Medicare services through September 30, 2014 for this report. However, based on analysis of data reported to the Lewin Group's site for Q10, we estimated only 20 fee-for-service Medicare enrollees had participated as of September 30, 2014.

[d] We plan to use Felton's assessment/encounter data and/or county mental health system data for future analysis due to limitations of the Alpha-MAX data available for CA. CA has high managed care penetration for general health services and inconsistent managed care encounter data reporting in Alpha-MAX. In addition, the county mental health system data include services provided to uninsured individuals who are not represented in Medicaid data.

[e] Individual-level claims data to report CMMI core program effectiveness measures are not available because consented participants did not give permission to access their claims.

Table I.4. Overview of program effectiveness evaluations for five awardees

Awardee	Intervention population	Comparison group	Data source(s) for outcomes [a,b]	Outcome measures	Key limitations
CHCS	Adults (19–64) who are homeless in Texas	Awardee randomly assigned treatment and control groups	Awardee-provided baseline and follow-up assessments	SF-6D, Brief Symptom Index, Adult Trait Hope Scale, University of Rhode Island Change Assessment	Data on expenditures, hospitalizations, and ED visits not available
ICSI	High-risk adult patients with Medicare or Medicaid coverage who have active depression plus uncontrolled diabetes or cardiovascular disease or both	None[c]	Medicare 2010–2014	ED visits, hospitalizations and hospital readmissions, and total expenditures	Medicare analysis only
KMHS	Adults with severe mental illness and children with SED who receive services at KMHS	Developed comparison group using claims data, matching at person level to MH facility client in Washington	Medicare 2010–2014	ED visits, hospitalizations and hospital readmissions, and total expenditures	Medicare analysis only
MMC	Adults with SMI living in southwest Brooklyn	Developed comparison group from claims data, matching at person level to SMI population in Philadelphia.	Medicare 2010–2014	ED visits, hospitalizations and hospital readmissions, and total expenditures	Medicare analysis only

Awardee	Intervention population	Comparison group	Data source(s) for outcomes [a,b]	Outcome measures	Key limitations
ValueOptions	Plan members with two or more detoxification admissions in Massachusetts	Awardee-selected comparison group of members meeting intervention participation criteria, but receiving care at treatment as usual sites.	Awardee-provided Medicaid claims and clinical assessment data	ED visits, short-term residential treatment stays, days of intensive day treatment, and total expenditures	Substantial part of sample missing clinical assessment information on baseline alcohol and drug use.

Note: For each awardee for which we have a comparison group, we conducted multivariate longitudinal analysis of intervention and comparison group outcomes or costs controlling for factors specific to that awardee using a "difference-in-differences" paradigm.

[a] Medicare Advantage participants are excluded from our 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 due to lack of available data.

[b] We did not conduct Medicaid analyses for ICSI, KMHS, and MMC. ICSI intervention group participants reside in multiple states; as a result, few participants are in any one state. In addition, many ICSI participants are enrolled in managed care and the consistency of reporting those managed care encounters in Alpha-MAX is unknown. KMHS Medicaid participants are in WA, which has a high rate of managed care enrollment. We will assess the usability of the WA state Medicaid managed care encounter data in the future. FPHNY has recently facilitated transmission of Medicaid data from NY state to support the evaluation of FPHNY and MMC; these data should have less lag time compared with Alpha-MAX to allow for more timely analysis. We expect New York to provide relevant data in early 2016.

[c] Intervention participation criteria cannot be sufficiently proxied in claims data to identify a comparison group. See Chapter VIII for further information about this issue.

ED=Emergency department; SED=Severe emotional disturbance; SMI=Severe mental illness.

E. Roadmap to the report

In the next chapter, we describe cross-awardee findings and themes for each of the evaluation questions. Later chapters (III–XII) contain findings for individual awardees, based on qualitative data and on the quantitative data on program effectiveness that were available for certain awardees. The chapters on each individual awardee include the following:

- A brief description of the goals and structure of the awardee's project and the data we used for this report
- Summaries of implementation effectiveness (including a description of the awardee's target population and enrollment progress), program effectiveness, and the contextual factors that influenced implementation
- A discussion of limitations of the analyses and overall conclusions

II. CROSS-AWARDEE FINDINGS

A. Introduction

In this chapter, we examine the progress the 10 awardees have made over the past year. We identify patterns in their implementation of the programs and highlight the similarities and differences in their experiences. The findings are based on the qualitative and quantitative data described in the previous chapter. The report's technical appendices and the chapters on each awardee contain more details about the specific data sources for each awardee. We organize this discussion by the evaluation's designated research domains and the questions they encompass:

- Implementation effectiveness
  1. What was the target population and was enrolled in the program?
  2. To what extent was implementation timely, conducted as planned, and responsive to site-level constraints?
  3. How did unexpected events support or conflict with successful implementation of the innovation?
- Program effectiveness
  1. To what extent did the program change Medicaid or Medicare expenditures on all care given to the target population?
  2. To what extent have rates of hospitalization changed?
  3. To what extent have rates of readmission changed?
  4. To what extent have levels of emergency department (ED) utilization changed?
  5. To what degree did these projects affect other outcomes?
  6. From the perspective of key informants, to what extent does the program improve desired health outcomes and quality of life for participants?
- Context
  1. To what extent did organizational features support or conflict with implementation?
  2. To what extent did the policy and political environment support or conflict with implementation?

B. Implementation effectiveness

In this section, we address three questions in the domain of implementation effectiveness:

- What was the target population and who was enrolled?
- To what extent was implementation timely, conducted as planned, and responsive to site-level constraints? (RQ 17) [5]

[5] RQ = Research question, as identified by CMMI.

- How did unexpected events support or conflict with successful implementation of the innovation? (RQ 8)

1. What was the target population and who was enrolled?

The target population, state setting, and number of implementation sites were different for each awardee (Table II.1). Some awardees worked with individuals who had specific mental health conditions, such as schizophrenia and depression, and others addressed a broader range of severe mental illness (SMI) and substance use disorders. Several awardees served individuals with both SMI and chronic physical health conditions, while a few focused on behavioral health services for specific subgroups, such as homeless individuals, residents of rural areas, and individuals being discharged from a detoxification facility.

Seven awardees implemented their innovation in a single state, with four limiting their program to individuals residing in specific geographic areas within those states. Almost all awardees worked with multiple sites to deliver their interventions. These sites were usually clinical providers such as primary care clinics, community mental health providers, and detoxification centers, but they also included respite centers, mobile care teams, and a homeless shelter.

Table II.1. Target population and setting, by awardee

Awardee	Overview of intervention	Target population	State	Number of sites
CHCS	Integrate primary care clinic into behavioral health service setting	Homeless adults with SMI who stayed in a San Antonio shelter	TX	1
Felton	Implement an integrated treatment model to improve intervention for psychosis	Youth (14–29) with schizophrenia, schizoaffective disorder, or schizophreniform disorder	CA	2
Feinstein	Improve intervention for schizophrenia through training, care management, and new technology	Patients with schizophrenia, recently discharged from hospital and receiving care at community treatment centers	FL, IN, MI, MO, NH, NM, NY, OR	10
FPHNY	Provide crisis intervention services to facilitate early engagement and continuity of care, combining community-based care with access to primary care	Individuals who have been diagnosed with psychosis	NY (Manhattan, Brooklyn, the Bronx, and Queens)	8
HLN	Provide behavioral care services via telehealth to individuals in rural areas lacking access to these services	Patients with behavioral health needs in areas with shortages of behavioral health clinicians	MT, WA, WY	100

Awardee	Overview of intervention	Target population	State	Number of sites
ICSI	Implement collaborative care management for patients with active depression and uncontrolled diabetes or cardiovascular disease	High-risk adult enrollees in Medicare and Medicaid who have active depression and uncontrolled diabetes or cardiovascular disease or both	CA, CO, FL, MA, MI, MN, PA, WA	171
KMHS	Integrate primary and mental health care	Adults with SMI and one comorbidity; children with severe emotional disturbance and one physical comorbidity	WA (Kitsap County)	1
MMC	Coordinate mental and physical health care through advanced health information technology	Adults with SMI	NY (Brooklyn)	12
ValueOptions	Provide care coordination	Adults with two or more detoxification admissions	MA	4
Vinfen	Integrate health care services into existing behavioral health outreach teams	Individuals with SMI who receive care from community-based behavioral health providers	MA (Boston)	4

Awardees also had different enrollment goals and levels of success in achieving them (Table II.2). The wide range in enrollment goals reflect the diversity in the participants, the care settings, and the number of implementing sites. By the end of the three-year award period, five awardees (CHCS, Felton, ICSI, MMC, and ValueOptions) met their enrollment targets. One awardee (HLN) achieved close to 90 percent of its enrollment goal, whereas three others (Feinstein, FPHNY, and Vinfen) attained less than 70 percent of their target enrollment.

Regardless of their results, most awardees faced challenges in recruiting and enrolling program participants. Although each awardee encountered its own challenges, establishing an effective referral process and identifying successful outreach and engagement strategies were the challenges cited most often. Awardees dealt with these challenges in a variety of ways, including broadening the criteria for eligibility, providing targeted training to staff, and redirecting their efforts to the most successful recruitment methods. We describe these and other strategies for improving engagement and outreach in the awardee-specific chapters of this report.

Table II.2. Awardee enrollment status at the end of year 3

Awardee	Enrollment goal	Number of direct participants[a] (percent of enrollment target met)	Major enrollment challenges
CHCS	260	261 (100)	• No notable challenges
Felton	140	188 (100)	• Establishing an effective referral system
Feinstein	770	506 (66)	• Delays in recruiting participants due to late notice of award
FPHNY	2208	1388 (63)	• Low enrollment due to restrictive participant eligibility criteria in one site
HLN	1534	1348 (88)	• Recruiting provider referral sites • Lower than expected volume of participant referrals at provider sites
ICSI	2700	2726 (100)	• Implementation delays in some sites • Identifying effective outreach and recruitment strategies for eligible population • Initially restrictive eligibility criteria
KMHS[c]	N.A.[c]	2033 (N.A.)	• Not applicable
MMC	500	635 (100)	• Outdated state-provided contact information for potential participants
ValueOptions	1492	1893 (100)	• Changes in state Medicaid program diminished the participant pool
Vinfen	400	216 (54)	• Some participants initially reluctant to use program services

[a] Figures are based on Q12 data provided by awardees to the Lewin Group's reporting website, unless otherwise specified. As of the drafting of this report, enrollment has concluded for all awardees.

[c] KMHS did not define a specific enrollment target for Race to Health! because KMHS staff intended the program to reach all patients who used KMHS' outpatient services. KMHS defined direct participants (2033 participants at the end of Q12) as patients who were referred by its primary care practice partner for behavioral health-related services, or patients who received direct services from either HCIA-funded medical assistants or the HCIA-funded healthy families' coordinator. KMHS defined indirect participants (3142 participants at the end of Q12) as patients on its Adult and Youth Care teams. KMHS provided Mathematica with an extract including all outpatients (7,891) served at the facility from January 2013 through June 2015, which we used for our analyses.

2. To what extent was implementation timely, conducted as planned, and responsive to site-level constraints? (RQ 17)

Although most awardees kept relatively close to their planned implementation schedule, they all confronted challenges in doing so. These challenges, and the strategies awardees used to address them, provide insights for policy makers, health care providers, and other stakeholders who may be considering similar innovations in behavioral health care. Below, we highlight common challenges and the strategies awardees used to overcome them.

Several awardees had difficulty hiring and retaining behavioral health specialists. Four awardees (CHCS, Felton, HLN, and Vinfen) faced obstacles in their efforts to recruit and retain qualified staff. For example, Felton had problems securing psychiatric nurse practitioners and psychiatrists because the pool of candidates was so limited in the rural counties that Felton served. HLN, which also focused on rural areas, did not face this particular challenge, because its

innovation provided services remotely. However, both of these awardees lost several staff to more lucrative opportunities. Vinfen also had trouble retaining nurse practitioners, in part due to the intense work involved in delivering community-based primary care to individuals with complex conditions. And CHCS struggled to find a primary care physician who combined an appreciation for the program's behavioral health-focused recovery model with experience treating homeless individuals. Organizations that are considering implementing similar programs should expect challenges in hiring and retaining qualified staff.

Establishing partnerships and collaboration between multiple sites and providers meant overcoming significant bureaucratic and legal hurdles. Two awardees, HLN and ICSI, implemented their programs in many different sites that were subject to a variety of requirements and procedures. The awardees could not readily obtain all the required approvals and partner agreements from the participating sites, and this hampered their ability to implement the program on time. ICSI—which implemented the COMPASS program in 197 sites—faced implementation delays due to the complexity of instituting cross-organizational legal agreements. To provide health care services in primary care sites, HLN's mental health specialists had to be credentialed to verify their professional experience and qualifications; obtaining the necessary credentials from each of the primary care clinics was a protracted and expensive process that caused delays in recruiting sites and enrolling participants. Any health care organization that seeks to implement a program at multiple sites with a variety of providers must plan for the time and resources necessary to establish the appropriate legal arrangements for effective collaboration.

Adopting technology, such as electronic health record (EHR) systems or telehealth platforms, proved to be more resource- and labor-intensive than some awardees expected. Three awardees (CHCS, ICSI, and KMHS) reported challenges with using EHRs or other forms of health IT. For example, ICSI partner sites reported their EHR systems were not compatible with the COMPASS registry, which meant that staff had to enter the same information into two and sometimes three different health IT systems. KMHS was unable to obtain electronic access to patients' health information to automatically populate its EHR system; as a result, program staff often had to manually enter the information. CHCS encountered delays in transitioning to a new EHR system because its IT resources were inadequate to the task, and the program struggled to find a system that would be useful to both primary care and behavioral health staff.

The practice of using specialized technology to monitor program participants presented other challenges. For example, during the award period, the technology company that developed one of Feinstein's planned program components (an ingestible sensor) was acquired by another technology company; as a result, Feinstein could not obtain the sensor in time to implement this component of the program. Vinfen found many clients were resistant to using the Health Buddy console, an in-home device for monitoring daily symptoms and medication adherence; however, Vinfen increased access to and use of the system with the launch of a web and mobile phone version. When adopting new IT systems, provider organizations must ensure the availability of sufficient resources and support for both the implementation effort and staff training. Clients' preferences for the use of new technology should also be carefully considered.

To address the differences between implementing sites, most awardees had to balance standardizing their programs and making allowances for sites that sought to customize the programs to their specific needs. Examples of such customization included allowing sites to adapt their recruitment procedures and operational processes to their own unique organizational contexts. For example, MMC allowed its partner organizations to adapt the program's staffing model to their own existing staffing structure and preferences. Each site determined which staff position was best equipped to take responsibility for participant outreach, intake, enrollment, and care coordination. Given the diversity of organizational structures, policies, and workflows in the implementing sites, awardees considered such flexibility to be both necessary and beneficial. Other providers considering multi-site integration should be mindful of establishing an appropriate balance between cross-site standardization and adaptation of an intervention to address the differences between sites.

Several awardees specifically designed their programs so they could continuously adapt them in response to challenges and lessons learned. For example, based on feedback from workforce members, Vinfen made mid-implementation changes to its service delivery structure and the protocol for administering its curriculum. KMHS hired more medical assistants and an additional healthy families coordinator in the final year of the program after the positions proved to be a positive addition. Awardees reported that the program modifications they made worked well and helped to keep the implementation on course or even improve it. In addition, they stressed that identifying needed program adjustments depended on effective internal feedback loops that involved staff directly responsible for administering program services. To ensure implementation effectiveness, health care and provider organizations adopting new programs should carefully consider methods to continuously gather and apply workforce input to improve program implementation.

3. How did unexpected events support or conflict with successful implementation of the innovation? (RQ 8)

As detailed in the individual awardee chapters, seven awardees reported unexpected events that affected their program implementation. The unexpected events were different for each awardee and included a natural disaster that delayed implementation, challenges in obtaining technology, loss of key program leaders, and state-level policy changes that impacted the potential pool of program participants. Although these unexpected events initially posed challenges to program implementation, the awardees were resourceful and creative in finding solutions that allowed them to stay on or close to the planned schedule.

C. Program effectiveness

In this section, we use quantitative data to address the following questions in the domain of program effectiveness:

- To what extent did the program change Medicaid or Medicare expenditures on all care for the target population? (CMMI core measure)
- To what extent have rates of hospitalization changed? (CMMI core measure)
- To what extent have rates of hospital readmissions changed? (CMMI core measure)

- To what extent have levels of ED utilization changed? (CMMI core measure)
- To what degree did these programs affect other outcomes? (awardee-specific measure)

Within the same domain, we use qualitative data to address the following question:

- From the perspective of key informants, to what extent does the intervention improve desired health outcomes for participants? (RQ 23 and RQ 29)

As described in Chapter I, we analyzed quantitative data for five of the awardees: CHCS, ICSI, KMHS, MMC, and ValueOptions. We explain why quantitative analyses were limited to these five awardees in Chapter I, in the awardee-specific chapters, and in the technical appendices of this report. We conducted impact analyses for CHCS, KMHS, MMC, and ValueOptions and pre-post analyses for ICSI. We plan to update these analyses and conduct additional analyses for our third annual report.

Table II.3 provides an overview of our approach to the quantitative analyses for the five awardees for which these analyses were possible. The analyses presented here include only the individuals enrolled in the programs through December 2013 (ICSI, KMHS, and ValueOptions) or March 2014 (MMC). We analyze outcomes through September 2014.

Table II.3. Overview of impact analysis approach, by awardee

Awardee	Number of unique intervention group members included in analyses	Period of analysis	Comparison group	Subset of population included in analysis
CHCS	208	March 2013–September 2014	Awardee randomly assigned intervention and control groups	All intervention participants
ICSI	263	February 2012–September 2014	N.A.[b]	Only FFS Medicare population[a] enrolled in intervention by December 2013
KMHS	874	January 2010–September 2014	Mathematica developed comparison group from claims data, matching at person level to MH facility client in Washington	Only FFS Medicare population[a] using services at KMHS through December 2013
MMC	236	March 2012–September 2014	Mathematica developed comparison group from claims data, matching at person level to SMI population in Philadelphia[c]	Only FFS Medicare population[a] enrolled in intervention by March 2014
ValueOptions	239	March 2012–September 2014	Awardee-selected comparison group of members meeting the program's participation criteria, but receiving care at treatment as usual sites[d]	All program participants (all are Medicaid enrolled) enrolled through December 2013

[a] The population included in the analyses is limited to individuals who have Medicare as their primary payer, are enrolled in Medicare Parts A and B, and are not enrolled in Medicare Advantage.

[b] Intervention participation criteria cannot be sufficiently proxied in claims data to identify a comparison group. See Chapter VIII for further information about this issue.

[c] We adjusted for geographic differences in Medicare payment rates in our expenditure analysis.

[d] Comparison sites were a convenience sample of non-participating MBHP sites.

The period covered by our analyses varied. For CHCS, we analyzed baseline and 12-month follow-up survey data. For ICSI, MMC, and ValueOptions, we analyzed data for one year pre- and post-intervention. For KMHS, we analyzed up to three pre-intervention and two post-intervention years.[6]

We employed a quasi-experimental design for KMHS, MMC, and ValueOptions. For CHCS, we used an experimental design because CHCS randomly assigned individuals to intervention or control groups. For ICSI, only pre-post analyses were possible. For CHCS and ValueOptions, we analyzed data only for participants who had enough exposure to the program for it to have an impact on the participant's outcomes.[7] For ICSI, KMHS, and MMC, we analyzed data only from Medicare fee-for-service (FFS) enrollees.

1. To what extent did the program change Medicaid or Medicare expenditures for all care in the target population?

Overall, there was little evidence that the interventions affected Medicaid or Medicare expenditures in the first intervention year. Currently available evidence indicates that the KMHS intervention resulted in a decrease in Medicare expenditures in the second intervention year. (KMHS is the only awardee for which we had data that covered the second post-intervention year.)

In Table II.4, we report the estimated impact of the intervention on Medicare or Medicaid expenditures per patient for the three awardees that had enough data available to support the analysis. We conducted pre-post analyses for one awardee.

- Starting out in the baseline period, the intervention group members at ICSI ($15,655) and KMHS ($15,869) had similar expenditure levels. Average expenditures for members of the ValueOptions ($21,867) and MMC ($28,469) intervention groups were somewhat higher than those for ICSI and KMHS populations in the baseline period. These baseline differences may be caused by a number of factors, including differences in the health needs of the eligible population for each intervention, variations in payment rates in the associated geographic areas and insurance programs,[8] and geographic differences in provider practice patterns.

[6] Data for ICSI, KMHS, MMC, and ValueOptions were analyzed in one-year periods. For individuals for whom less than one year of data were available in a given analysis period, estimates for the observed period were prorated. In regression models, the observations were weighted according to the proportion of the analysis period for which an individual was enrolled.

[7] For CHCS, we required individuals to have been enrolled for at least six months; that is, data from the six-month follow-up interview had to be available for us to include the individual in our analysis. For ValueOptions, we required individuals to have 9 months of continuous enrollment in the Massachusetts Behavioral Health Plan (MBHP) after they enrolled in the program.

[8] Note that ValueOptions expenditures are for Medicaid enrollees and the expenditures for the other three awardees assessed here are for Medicare FFS coverage.

- Between the baseline year and the first intervention year, average expenditures for both the participants in the intervention groups and the members of the comparison groups at ValueOptions increased. In contrast, at MMC expenditures declined for both groups between the baseline and the first year following enrollment in the program. For KMHS, average expenditures for the comparison group increased between the baseline and the first program year, but those for the intervention group declined. These differences in trends between the awardees may be the result of many factors, including differences in the health needs of participants, geographic variation in payment rates and practice patterns, and effects of the program.
- Estimates for ICSI indicate a significant increase (p-value <.01) in expenditures after one year of the intervention as a result of the program, suggesting that expenditures for program participants increased relative to the baseline period. The ICSI findings should be considered tentative because they were sensitive to model specification.
- For KMHS, we estimate the impact on expenditures during the second year of participation for participants who were enrolled in the program for two years. Our analyses indicate a significant decrease in expenditures during the second program year, suggesting that expenditures for intervention participants decreased more relative to than baseline than expenditures for those who did not receive the intervention did.

We note that these findings address cost and service use outcomes that are of particular interest to CMMI and that can be measured in similar ways across the awardees. For most awardees, our analyses do not address measures of health or mental health status because we do not have access to relevant data. However, we do report clinical outcomes for CHCS and ValueOptions in their respective chapters (Chapters III and IX). We found no impacts on these outcomes, possibly because of the limited data available to us; therefore, we did not examine potential links between changes in expenditures and changes in health outcomes—an analysis that we will conduct, if warranted, for our next annual report.

Table II.4. Summary of estimates for expenditures per patient, annual results[a]

	Regression-adjusted mean for intervention group	Regression-adjusted mean for comparison group	Estimated impact[b]		
			Value	Percent	P-value
ICSI					
Baseline	$15,655	N.A.	N.A.	N.A.	N.A.
IY1	$22,053[c]	N.A.	N.A.	N.A.	N.A.
KMHS					
Baseline	$15,869	$14,921	N.A.	N.A.	N.A.
IY1	$15,264	$15,803	-$1,487	-9	0.37
IY2	$8,416	$11,114	-$3,646***	-30	<0.01
MMC					
Baseline	$28,469	$28,313	N.A.	N.A.	N.A.
IY1	$25,280	$20,035	5,088	25	0.29
ValueOptions					
Baseline	$21,867	$19,984	N.A.	N.A.	N.A.
IY1	$25,513	$27,146	$-3,516	-12	0.14

Source: For ICSI, KMHS, and MMC, we analyzed FFS Medicare expenditures and utilization data for February 2012–September 2014, January 2010–September 2014, and March 2012–September 2014, respectively. For ValueOptions, we analyzed MBHP provided program enrollment data, Medicaid administrative data, and baseline assessment data covering the period from March 2012 through September 2014.

[a] Expenditure estimates cover expenditures for a 12-month period for an individual who is enrolled for all 12 months.

[b] We derived the impact estimates in Stata using the lincom command to compare the difference between the intervention and baseline period means for the treatment and comparison groups.

[c] As noted, we conducted a pre-post analysis for ICSI. The difference between baseline and IY1 periods is $6,339 (p<.01).

*** Impact estimates are statistically different at the .01 level, two-tailed test.

N.A. = Not applicable. IY1 = Intervention year 1; IY2 = Intervention year 2.

2. To what extent have rates of hospitalization changed?

KMHS is the only awardee for which the current evidence indicates that the program resulted in a decrease in hospitalizations in its first year.

In Table II.5, we report the estimated impact of the programs on Medicare hospitalizations for three awardees.

- In the baseline period, intervention group members at KMHS had 0.7 hospitalizations per patient, respectively. Intervention group members at MMC had a somewhat higher hospitalization rate of 1.4 per patient.

- Between the baseline and the first program year, the hospitalization rate decreased for both KMHS and MMC intervention and comparison group members. For MMC, we estimated program impact on hospitalizations for the first program year only. For MMC, we identified no significant impact from the program. For KMHS, we estimated a marginally significant decrease of 0.1 (p-value = 0.07) in hospitalizations for both the first and second program year as a result of the program. For ICSI, we identified a significant increase of 0.5 hospitalizations per patient (p-value = 0.03) in the first program year relative to baseline.

Table II.5. Summary of estimates for hospitalizations per patient, annual results[a]

	Regression-adjusted mean for intervention group	Regression-adjusted mean for comparison group	Estimated impact[b]		
			Value	Percent	P-value
ICSI					
Baseline	0.8	N.A.	N.A.	N.A.	N.A.
IY1	1.2[c]	N.A.	N.A.	N.A.	N.A.
KMHS					
Baseline	0.7	0.7	N.A.	N.A.	N.A.
IY1	0.4	0.5	-0.1*	-17	0.07
IY2	0.3	0.4	-0.1*	-23	0.07
MMC					
Baseline	1.4	1.6	N.A.	N.A.	N.A.
IY1	0.9	0.9	0.2	23	0.58

Source: For ICSI, KMHS, and MMC, we analyzed FFS Medicare expenditures and utilization data for February 2012–September 2014, January 2010–September 2014, and March 2012–September 2014, respectively.

[a] Hospitalization estimates represent a 12-month period for a full-time enrolled individual.

[b] We derived the impact estimates in Stata using the lincom command to compare the difference between the intervention and baseline period means for the treatment and comparison groups.

[c] As noted, we conducted a pre-post analysis for ICSI. The difference between baseline and IY1 periods is .05 (p=.03).

* Impact estimates are statistically different at the .10 level, two-tailed test.

N.A. = Not applicable. IY1 = Intervention year 1; IY2 = Intervention year 2.

3. To what extent have rates of readmission changed?

We were able to estimate rates of readmission only for KMHS. Overall, there was little evidence that the programs affected hospital readmission rates in the first intervention year.

Specifically, we estimate the program's impact on Medicare hospital readmissions per full-time enrolled patient per year for KMHS (Table II.6). Readmission rates were not analyzed for ICSI, MMC, and ValueOptions, because there were too few observations and too few events. The results show no statistically significant impact on readmissions in year 1 or year 2 of the program.

Table II.6. Summary of estimates for readmissions per patient, annual results[a]

	Regression-adjusted mean for intervention group	Regression-adjusted mean for comparison group	Estimated impact[b]		
			Value	Percent	P-value
KMHS					
Baseline	0.13	0.12	N.A.	N.A.	N.A.
IY1	0.08	0.11	-0.03	-26	0.34
IY2	0.03	0.05	-0.03	-44	0.21

Source: Mathematica analysis of FFS Medicare expenditures and utilization data for January 2010–September 2014.

[a] Hospital readmission estimates represent a 12-month period for a full-time enrolled individual.

[b] We derived the impact estimates in Stata using the lincom command to compare the difference between the intervention and baseline period means for the treatment and comparison groups.

N.A. = Not applicable. IY1 = Intervention year 1; IY2 = Intervention year 2.

4. To what extent has ED utilization changed?

Overall, ValueOptions is the only awardee for which the current evidence indicates that the program resulted in a decrease in ED visits in its first year.

In Table II.7, we report the estimated impact of the program on ED visits per patient for three awardees and conducted pre-post analyses for one:

- Average ED visits per patient in the baseline year were in a similar range for ICSI (1.7), KMHS (1.5) and MMC (1.7) intervention group members. Average ED visits were substantially higher for ValueOptions intervention group members in the baseline year (6.7). These differences may be attributable to differences in the target population, regional variation in payment rates, and varying practice patterns.
- We observed differences in average ED visits between the baseline period and the first program year across the awardees. For MMC, these visits declined for both intervention and comparison group members. For KMHS, they rose slightly for both intervention and comparison group members. For ValueOptions, they declined for intervention group members but rose for comparison group members.
- For the first year post-program enrollment, we estimated the impact of the program on ED visits for the awardees. The estimates for KMHS and MMC indicate no significant effect on ED visits. The estimate for ValueOptions indicates a significant decrease of 2.1 ED visits per patient (p-value = 0.02) during the first program year. We also estimated the impact on ED visits during the second program year for KMHS only, and found no significant impact. Pre-post analyses for ICSI indicated a slight increase in ED visits relative to baseline, an increase that was not statistically significant.

Table II.7. Summary of estimates for ED visits per patient, annual results[a]

	Regression-adjusted mean for intervention group	Regression-adjusted mean for comparison group	Estimated impact[b]		
			Value	Percent	P-value
ICSI					
Baseline	1.7	N.A	N.A.	N.A.	N.A.
IY1	1.9[c]	N.A	N.A.	N.A.	N.A.
KMHS					
Baseline	1.5	1.5	N.A.	N.A.	N.A.
IY1	1.7	1.7	0.0	0	0.95
IY2	1.5	1.6	-0.1	-8	0.34
MMC					
Baseline	1.7	1.9	N.A.	N.A.	N.A.
IY1	1.4	1.5	0.2	14	0.61
ValueOptions					
Baseline	6.7	5.7	N.A.	N.A.	N.A.
IY1	5.5	6.5	-2.1**	-27	0.02

Source: For ICSI, KMHS, and MMC, we analyzed FFS Medicare expenditures and utilization data for February 2012–September 2014, January 2010–September 2014, and March 2012–September 2014, respectively. For ValueOptions, we analyzed MBHP provided program enrollment data, Medicaid administrative data, and baseline assessment data covering the period from March 2012 through September 2014.

[a] ED visit estimates represent a 12-month period for an individual enrolled in the program full time.

[b] We derived the impact estimates in Stata using the lincom command to compare the difference between the intervention and baseline period means for the treatment and comparison groups.

[c] As noted, we conducted a pre-post analysis for ICSI. The difference between baseline and IY1 periods is not significant.

** Impact estimate is statistically different at the .05 level, two-tailed test.

N.A. = Not applicable. IY1 = Intervention year 1; IY2 = Intervention year 2.

5. To what degree did these projects affect other outcomes?

We estimated program impacts for other outcome measures for CHCS and ValueOptions. The CHCS program is intended to reduce psychological stress, improve the participant's capacity for self-management, increase feelings of hope, and increase the participant's capacity to change. Our analysis focuses on four metrics related to these outcomes (Chapter III). We found no statistically significant impacts on the four outcomes we analyzed. However, the preliminary quantitative results presented in this report reflect changes in outcomes for only the first six months of the year-long intervention period.

For ValueOptions, intensive day treatment and residential treatment were commonly used services in the pre- and post-program period and the program intended to affect these outcomes. We found no significant effect of the program on these outcomes in the current analyses (Chapter XI).

6. From the perspective of key informants, to what extent does the intervention improve desired health outcomes and quality of life for participants? (RQ 23 and RQ29)

Although the 10 awardees implemented wide-ranging programs, key informants universally agreed that the programs had a positive effect on their participants' mental and physical health,

service use, and quality of life. We highlight examples of common themes expressed by key informants in this section. We provide additional examples of program effects, including respondents' perceptions of the program components and characteristics that facilitated improvements in participant well-being, in the awardee-specific chapters.

- **Better physical and mental health.** Respondents from six awardees (CHCS, HLN, ICSI, KMHS, ValueOptions, and Vinfen) noted positive changes in clinical indicators. For example, a ValueOptions RSN described a participant whose health improved significantly during his time in the program: "I had a client who was almost dying when we first met him— his HIV was over the roof, liver function was very bad, but by the time we discharged him, he was fine and doing so well." Key respondents from CHCS, Feinstein, HLN, ICSI, and ValueOptions also noted improvements in participant mental health status over the course of the program. Respondents described reduction in symptoms of anxiety and depression among participants, and, in some cases, reductions in medication use and associated side effects as a result of medication management offered through the program. Respondents from CHCS and ValueOptions also reported decreased substance abuse among some participants, which contributed to improvements in their physical and mental well-being.
- **Fewer hospitalizations, readmissions, and ED visits.** Based on anecdotal evidence, staff from CHCS, FPHNY, KMHS, MMC, and Vinfen believed that the program reduced participants' use of the emergency department and decreased their number of hospitalizations. Some attributed these changes to the program, noting the improved availability of preventive care and increased awareness among participants about appropriate sources for care. For example, KMHS program staff said that being more aware of the reasons behind participants' ED visits helped them redirect participants to appropriate care as necessary: "I think [patients] are overall getting better care because I feel like we're catching a lot of stuff that we didn't catch before. We're finding out why they're not following through with medical appointments, what's causing them to have their diabetes out of control, and what we can do to help them rein it in." Staff from Feinstein also reported that, based on their own internal analysis, readmission rates decreased by half for participants in the health technology program when compared to the program's standard reference group over a six-month period.
- **Increased access to services.** Respondents from CHCS, HLN, KMHS, MMC, and Vinfen reported improved access to routine health services among participants. An outreach team member from Vinfen 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." Respondents from HLN also reported lower rates of no-shows because participants could access mental health care at the PCP office instead of driving to a bigger city or facing the possible stigma associated with visiting a psychiatrist's office. As one respondent explained, "It seems that there's more follow-through with the patients when they come to a familiar area."
- **More self-management of health conditions.** Respondents from four awardees (FPHNY, ICSI, KMHS, and Vinfen) said participants were more aware of the importance of nutrition and primary care check-ups, self-care routines, and medication adherence, and they understood their own physical and mental health conditions better. These changes helped

participants improve their health behaviors and lifestyle choices, and in some cases, influenced their decisions on when to seek care. For example, Vinfen respondents explained that participants had become more knowledgeable about when a health concern was a real emergency, rather than an ongoing symptom. Staff from FPHNY, ICSI, and KMHS also described how participants had become more empowered to manage their own care and to advocate for themselves when talking with health professionals. For example, a respondent from FPHNY noted that participants in the program learned to talk with their doctors and care team members about lowering medication dosages to decrease the number and degree of side effects.

- **Enhanced self-esteem and social competence.** Respondents from six awardees (CHCS, Felton, FPHNY, ICSI, ValueOptions, and Vinfen) described the improvements in participants' self-esteem and social competence over the course of the program. In some cases, the social supports such as food, transportation, and clothing that were provided by program staff helped participants regain their self-respect. Other participants were able to lead fuller lives because they were not as consumed with physical or mental health issues. A FPHNY clinician said, "The clients feel like they're being successful by working with us, working towards their goals and achieving their goals. Several of our clients, we see them less frequently now because they have jobs, they have their own apartments and things like that so I think it's positive and also a lot of clients that I've worked with … say… we regard you almost as a friend or family member, which is really nice." Respondents from FPHNY, ICSI, and Feinstein also said their participants' social functioning and relationships with family and friends improved as a result of the programs. For example, a Vinfen workforce member described how the intervention helped his participants feel more connected to the community and cared for by intervention staff: "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."

Although key respondents from all 10 awardees cited multiple examples of positive program effects, some respondents also emphasized that they expect to see greater results only in the longer term, given the relatively short course of the intervention period. For example, respondents from Vinfen stressed that many participants entered the program with chronic health problems; therefore, they did not expect to see immediate or significant changes in health indicators. ICSI respondents reiterated this point, noting that smaller improvements can lead to larger changes in other areas over time. For example, one care manager pointed out that minor improvements in depression can lead to improvements in social functioning and self-care, which can lead to improvements in physical health. Nevertheless, respondents agreed that the programs had positive effects on participant health and well-being, by improving their physical and mental health status, their use of health services, and their self-esteem and social functioning.

D. Context

In this section, we address two questions within the context domain:

- To what extent did organizational features support or conflict with implementation? (RQ 104)

- To what extent did the policy and political environment support or conflict with implementation? (RQ 106)

1. To what extent did organizational features support or conflict with implementation? (RQ 104)

Respondents from all the awardee programs identified several organizational features that either supported or impeded program implementation. Cross-awardee analysis of these facilitators and barriers revealed several lessons learned that can be applied to the broader behavioral health field or to innovation efforts beyond those presented in this report (Table II.8).

Table II.8. Organizational facilitators and barriers to program implementation

Facilitators	Barriers
• Mission and vision aligned with program goals • Experience implementing similar program models • Visionary program leaders with strong commitment to program goals • Enthusiastic mid-level or clinical staff leaders who encourage staff buy-in • Emphasis on collaboration and communication when implementing program across multiple sites or partner organizations • Engagement with community stakeholders through existing relationships or strategic outreach	• Difficulty changing current practice to adopt new models of care • Varying management and staffing structures across implementing sites, limiting awardee ability to standardize and monitor program implementation • External community stakeholders unengaged or resistant to program model

Below, we describe these features and provide examples of how they facilitated or hindered program implementation.

Culture and experience. Organizational culture and experience with program models played an important role in implementing nearly all the programs. Respondents described the mission and values of several awardee organizations (Felton, KMHS, MMC, and Vinfen) as key factors that supported implementation. Several awardees (Feinstein, ICSI, and KMHS) also mentioned the ways that their implementing sites' experience made the implementation easier. For example, Feinstein selected community mental health centers they had partnered with before on similar research projects. Likewise, many of ICSI's COMPASS sites had experience with other mental health integration, primary care redesign, and care coordination programs. Awardees noted these sites benefited from having some of the necessary staff and infrastructure already in place.

At times, organizational culture slowed some awardees' implementation progress (CHCS, FPHNY, ICSI, and MMC), particularly when programs required frontline staff to adjust how they approached patient care. For example, some clinical staff in the FPHNY and ICSI programs initially resisted these program models because they conflicted with staff's training in more traditional models of care. Awardees overcame these barriers by providing increased training to program staff and creating more opportunities for communication and feedback.

Leadership. Strong leadership was a key facilitator of program implementation for most awardees (Felton, Feinstein, ICSI, KMHS, MMC, and Vinfen), both at the administrative and

frontline levels. For example, as Feinstein and KMHS planned and implemented their programs, both benefited from leadership teams with expertise in their respective programs' target population or type of intervention. Staff at other programs, such as Felton and Vinfen, highlighted their leaders' strong support and accessibility as key elements in successful program implementation. In addition to strong administrative leaders, ICSI staff singled out the support they received from clinical leaders as important to implementation of the COMPASS program. In particular, respondents said that enthusiastic and influential physician champions and committed mid-level staff were critical to achieving buy-in from other practice staff and promoting good communication among team members.

Collaboration among partners. Implementing and standardizing programs across multiple provider sites was challenging, yet effective implementation often required the involvement of multiple stakeholders or multiple sites. Respondents said that strong coordination and collaboration supported implementation of some awardee programs (ICSI, MMC, ValueOptions). For example, ICSI involved a diverse group of care delivery organizations in the COMPASS consortium and worked to gain buy-in and facilitate learning across the program's collaborators. MMC also benefited from effective coordination with a large and diverse group of partners. Respondents highlighted MMC's inclusive governance structure as important to ensuring that feedback from care management partner organizations informed all stages of program implementation.

The need to coordinate across diverse partner organizations also posed a barrier for several awardees (FPHNY, ICSI, ValueOptions, and Vinfen). For example, although the strong relationship between MBHP and local providers made it easier to implement the ValueOptions program, MBHP lacked control over staffing at the provider sites, and the sites' varying organizational structures and management styles sometimes meant a struggle to standardize services for research purposes. Most awardees addressed this challenge by (1) allowing sites to adapt the programs to their unique organizational needs and (2) developing opportunities for partner organizations to provide input on program implementation. For example, MBHP sought feedback from staff at implementing sites when considering changes to the ValueOptions program procedures; MBHP also encouraged sites to develop their own protocols for some program operations, such as distributing incentive payments to participants.

Engaging community stakeholders. In addition to working effectively with implementing partners, successful programs also needed to engage with community stakeholders such as social service providers, county health staff, and local hospitals. For example, respondents noted that the success of Felton's PREP program relied heavily on Felton's establishing an effective referral system that included local school districts and county social service providers. Felton invested heavily in an outreach strategy to engage these stakeholders who were initially apprehensive about having an out-of-county organization provide services to their residents and clients. Similarly, some of ICSI's COMPASS sites leveraged existing partnerships—including relationships with local YMCAs, community colleges, and fire departments—to better implement COMPASS in their communities. At least one awardee (FPHNY) struggled to engage and raise awareness among external stakeholders such as hospitals and emergency departments, which made it difficult to generate appropriate referrals at some sites.

2. To what extent did the policy and political environment support or conflict with implementation? (RQ 106)

Awardees identified several aspects of the policy and political environment that supported or hindered implementation. In this section, we summarize the key policy and political factors awardees reported at the federal, state, and local levels.

Only a few awardees highlighted federal policy as a key influence on program implementation. Two awardees cited the effects the effects of their states' decisions about Medicaid expansion. CHCS respondents reported that Texas's decision not to expand Medicaid coverage posed a barrier to improved outcomes, because a large majority of participants remained uninsured during implementation, and required services beyond those available through the HCIA-funded primary care physician. In contrast, KMHS respondents indicated that Medicaid expansion in Washington State resulted in more demand for mental health services. In the absence of a parallel increase in the number of providers qualified to provide these services, KMHS and its partner organizations struggled to increase their staff capacity enough to meet this demand. KMHS respondents noted that the need to continuously hire and train new staff sometimes distracted from program implementation efforts. For HLN, the lack of standardized regulations and guidelines for hospital and clinic credentialing caused confusion among providers and significantly hampered implementation.

Awardees more commonly highlighted the role of state policies and programs in implementation. For example:

- Felton respondents stressed that California's Mental Health Services Act of 2004 (MHSA), which provides funding for prevention and early intervention services, created an incentive for counties to adopt or collaborate with programs like PREP. The two counties with HCIA-supported programs will use MHSA funding to sustain Felton's PREP program.
- FPHNY respondents noted that New York State's Medicaid redesign provides an opportunity for sustainability of the program. In 2015, the state changed its policy to provide reimbursement for Parachute NYC services, including peer support and CRCs. According to Parachute staff, this change was originally slated to take effect at the beginning of 2015 but was delayed until October 2015.
- HLN respondents cited several state policies supporting the use of telemedicine. In Washington State, a new provision states that EDs are not allowed to hold psychiatric patients overnight without a psychiatrist consult. Respondents believed this has resulted in increased demand for mental health coverage, including tele-psychiatry services like HLN's. In Minnesota, parity regulations require commercial payers to provide reimbursement of telemedicine services at a level equal to that provided for in-person services.
- MMC program leaders noted that its HCIA-funded program developed alongside New York State's health homes program, which will continue to support program services for most MMC participants after the award period.
- Vinfen's program aligns with Massachusetts' One Care program, which launched in October 2013, for individuals dually eligible for Medicare and Medicaid. Vinfen hopes to see some of the intervention components integrated into the One Care program.

Awardees also identified several local policy or political factors in program implementation. For example, the city of New York provided FPHNY with in-kind support for external evaluation of its Parachute NYC program. Respondents also identified some local environmental challenges to implementation—most commonly, a lack of available and affordable housing for the patient population (Feinstein, FPHNY, KMHS, and MMC). For example, respondents from FPHNY and MMC cited New York City's lack of affordable housing as a major barrier for frontline staff's effective implementation of these programs, because many program participants were homeless or lacked consistent, stable housing.

E. Limitations of our analysis

A number of limitations to our analyses are important to consider when interpreting our results:

- **Design limitations.** As noted, for KMHS, MMC, and ValueOptions we used a quasi-experimental design for the analysis. For ValueOptions, the comparison group included individuals who met all the criteria to be eligible for the program, but who received treatment as usual at a convenience sample of non-participating sites. For ICSI, we were able to conduct only pre-post analyses because intervention participation criteria cannot be sufficiently proxied in claims data to identify a comparison group. For KMHS, and MMC, we used matching techniques on Medicare administrative data to develop comparison groups. The matching techniques used for these three awardees are intended to reduce the differences between members of the intervention group and members of the comparison group on observable measures related to the outcomes of interest. However, estimates from quasi-experimental models may be biased if the selected comparison population differs from the intervention population on unobserved characteristics that are correlated with the outcome measures of interest. For example, the two groups may differ in their marital status—a variable for which we have no data and which may be correlated with health outcomes in this population. This unobserved difference may bias outcomes in ways that we cannot detect.
- **Only subsets of intervention groups and of costs were assessed.** For ICSI, KMHS, and MMC, the current analysis is limited to Medicare FFS enrollees and Medicare covered services. Our findings may not be generalizable to other program participants (for example, patients covered by Medicare Advantage, Medicaid only, or commercial insurance). Moreover, the services and costs reported for the Medicare population we analyzed may substantially underestimate service utilization and costs for this population, particularly for dual Medicare-Medicaid enrollees, because they do not include services covered by Medicaid or other health insurers, or services not covered through any insurance (for example, psychiatric hospitalizations that exceed allowable limits).
- **Small sample sizes.** In all of the analyses except the one for KMHS, the size of the intervention group is less than 300. The small size of the analysis population may make the results more sensitive to outliers and to the model specification. We may obtain more stable results in later analyses that are based on larger samples.
- **Initial phase of intervention only.** As noted, the analyses presented here include only the individuals enrolled in the programs through December 2013 (ICSI, KMHS, and ValueOptions) or March 2014 (MMC). In addition, available data allow us to analyze

outcomes only through September 2014. The programs, however, did not end until June 2015. Thus, we expect to incorporate data for more participants and time periods in future analyses. For several reasons, findings may change when we analyze additional data. For example, staff may refine the programs over time as they learn lessons from early experiences. The characteristics of later enrollees may be different from the characteristics of those who enrolled early. In addition, the interventions may have greater impacts on individuals who have been enrolled longer.

- **A focus on selected quantitative outcome measures.** In response to CMMI's request, we focused on assembling standard administrative and claims data from Medicaid and Medicare files that would allow us to assess four core outcome measures for both participant and comparison groups (when available). We also used data provided by the awardees if such data were available for both groups. Most of the awardees gathered additional data on participants, such as the number of services received, clinical indices of health or mental health status, and surveys of participant experiences with care. These awardee-specific sources of data can provide important information about the implementation and outcome of the programs. However, given CMMI's priorities, we did not routinely gather and analyze these data for this report. As a result, our evaluation does not address some important questions regarding the implementation and effects of these programs.

In light of these limitations and the preliminary nature of our findings to date, we urge readers to avoid drawing premature conclusions regarding the impact of the awardee programs.

F. Summary and conclusion

The 10 HCIA awardees in behavioral health 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. The cross-awardee analyses presented in this chapter can help inform CMS as the agency considers plans to support further development or scaling up of selected intervention approaches. Our qualitative and quantitative analyses to date reveal several key findings:

- Despite the range of programs that were implemented, the awardees encountered many of the same challenges, including difficulty hiring and retaining qualified behavioral health specialists, establishing partnerships and collaboration across multiple sites and providers, and adopting technology such as new EHR platforms or telehealth devices. Future endeavors can avoid similar pitfalls by learning from these awardees' implementation challenges and their approaches to addressing them.
- Flexibility and adaptability proved to be essential for successful implementation. For programs with multiple sites or providers, awardees found it necessary to balance program standardization with customization based on site-specific needs. In addition, several awardees specifically designed their programs to include continuous adaptation in response to unanticipated challenges and lessons learned.
- Although quantitative data analyses have not yet revealed a widespread program impact on core measure outcomes, respondents universally agreed that the programs have had a clear and positive effect on participants' mental and physical health, service use, and quality of life. Many respondents said they did not expect the programs to have a significant impact on

core measures until several years after their implementation. Future quantitative analysis may reveal increased program impact.

- According to respondents, organizational characteristics can shape program implementation positively or negatively. The positive characteristics include having a mission and vision that is aligned with program goals, experience implementing similar program models, visionary and committed leaders, enthusiastic mid-level or clinical staff who encourage staff buy-in, an emphasis on collaboration and communication, and engagement with community stakeholders. Organizational barriers include difficulties with adapting to new models of care when staff are entrenched in earlier ways of providing services; site-to-site differences in management and staffing structures, which can result in limitations in awardees' ability to standardize and monitor programs across sites; and a lack of community engagement in some organizations.
- Although federal policies impacted a few awardee programs, awardees more commonly reported that state-level policies and programs have considerable effects on interventions. Future endeavors should consider the effects of potential shifts in state policies and programs when planning to implement new programs.

In the coming year, we will analyze more quantitative and qualitative data. Additional Medicaid data will likely become available during the next six to nine months; consequently, we expect to conduct additional and more comprehensive analyses than were possible for this report. Furthermore, we will be analyzing qualitative data about awardees' intentions for sustaining key elements of their programs. In addition, we expect to integrate qualitative and survey data regarding frontline workers—a key component of the delivery system for individuals with SMI and substance use disorders. Finally, in our third annual report, we expect to synthesize information from these diverse analyses. We expect this synthesis to give CMS more evidence that can be used to improve care and services and reduce costs for this group of Medicare and Medicaid beneficiaries.

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

III. CENTER FOR HEALTH CARE SERVICES

A. Introduction

The Center for Health Care Services (CHCS), a San Antonio, Texas-based mental health care provider, is implementing Project HEALTH (Homeless Engagement Addressing Limitations to Healthcare) on the Haven for Hope campus, a planned campus for homeless individuals that includes a traditional shelter and a range of mental health and social services. Project HEALTH is designed to integrate behavioral and physical health care for homeless adults with serious mental illness—or serious mental illness combined with a substance use disorder—who have or are at risk of developing a chronic physical disease. The approach focuses first on stabilizing behavioral health conditions and then on actively engaging the participant in primary care, prevention, and wellness services. With the Project HEALTH model, CHCS aspires to lower costs and improve health care and health outcomes by reducing emergency department visits and psychiatric and medical hospitalizations.[9]

CHCS successfully reached its enrollment target of 520 homeless adults, all of whom lived at Prospects Courtyard, a safe outdoor sleeping area located on the Haven for Hope campus. CHCS randomly assigned 261 of these adults to an intervention group and 259 to a control group. Members of the intervention group received integrated behavioral and physical health services, as well as peer support to build and sustain their readiness for change, motivation, and treatment compliance. Members of the control group received more standard care, including assistance from community guest specialists who provided linkages and referrals to existing services and resources.

The findings in this report are based on quantitative and qualitative data collected or received by April 30, 2015, and June 1, 2015, respectively, and enrollment data reported throughout the award period. Data sources included:

- Enrollment data submitted by CHCS to the reporting website maintained by the Lewin Group for the HCIA initiative
- Baseline and follow-up assessment data on all participants, submitted by CHCS to Mathematica. Assessments measure a variety of health outcomes including physical function, psychological distress, feelings of hope, and motivation for change.
- Qualitative data, including telephone interviews and an in-person site visit in March 2015. Mathematica conducted in-depth interviews with key awardee staff, members of the workforce, and other stakeholders. In addition, Mathematica convened focus groups with members of the workforce and with members of the intervention and control groups (known as participants and non-participants). Data from the participant and non-participant focus groups, however, are not part of this report; instead, we provide an in-depth analysis of those focus groups in our seventh quarterly report.

[9] CHCS received a no-cost extension from CMMI to extend its program for 12 months, so we use the present tense to describe the program's activities. CHCS will maintain program services through October 2015 and continue its evaluation, analysis, and dissemination of data through June 2016.

In this chapter, we also present the results of quantitative analysis on demographic information and selected health outcomes. This analysis is based on data CHCS provided for both the intervention group and the randomly assigned control group. CHCS's data files included variables related to key program outcomes, such as psychological stress, capacity for self-management, feelings of hope, and capacity for life change; we provide impact estimates for these outcomes. This chapter does not include analysis of CMMI's four core measures, because few CHCS participants are enrolled in Medicaid or Medicare, and CHCS's data files do not include information corresponding to these measures.

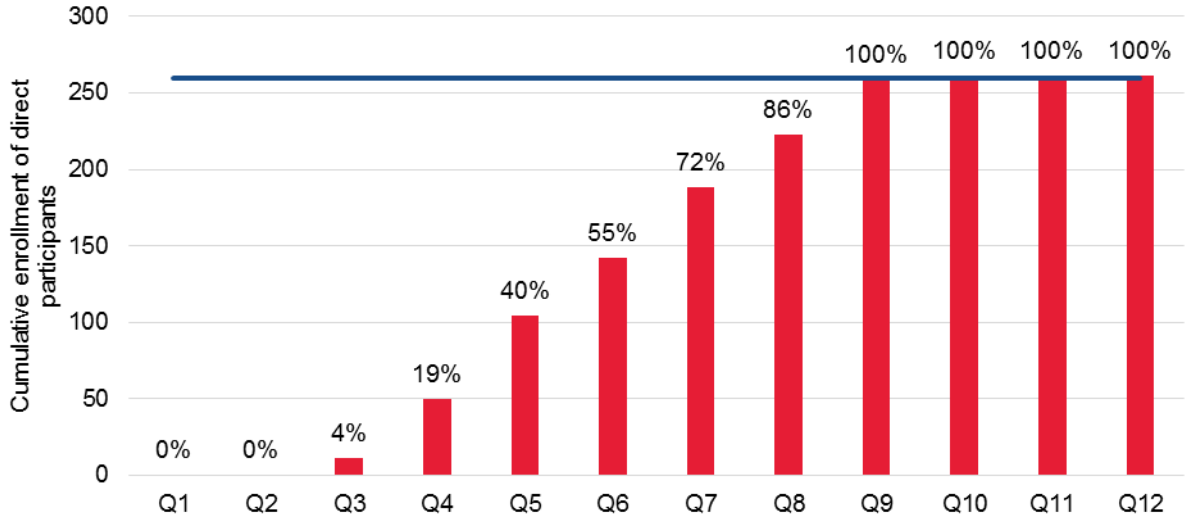
The research questions we address in this report were selected in consultation with our COR. Specifically, we examine questions related to implementation effectiveness in Section B, program effectiveness in Section C, and the program and policy context in Section D. We close with an overall summary and conclusion (Section E).[10]

B. Implementation effectiveness

1. Who was enrolled in the program?

By the end of the 12th quarter (June 30, 2015), CHCS had enrolled 261 direct participants in Project HEALTH, exceeding its original goal of 260 by one (Figure III.1). In the first two quarters, CHCS set up and piloted the project. Enrollment began in the third quarter with 11 direct participants and increased steadily between quarters 4 and 10, with CHCS reaching its target enrollment in quarter 9.

Figure III.1. Percent of target enrollment achieved by quarter, Q1–Q12—CHCS



Source: Quarterly reports submitted by CHCS to the Lewin Group's reporting website for the HCIA initiative.
Note: As displayed by the horizontal bar, CHCS's target enrollment was 260 unique participants.

[10] CHCS may conduct analyses of data that they have collected on program participants and comparison groups and publish reports about program outcomes. These analyses are likely to use somewhat different data, focus on different outcome measures, and address different specific questions compared with the data and measures we use in this report to address questions of direct interest to CMMI. As a result, CHCS may reach conclusions about program effects that differ substantially from the ones we present in this and future reports.

2. To what extent was implementation timely, conducted as planned, and responsive to site-level constraints? (RQ 17)

CHCS has largely succeeded in implementing the program in a timely manner. However, several challenges delayed or hindered its achievement of program goals:

- **Hiring primary care staff.** It took longer than expected for the program to hire a physician with both an appreciation of the recovery-oriented model and expertise in working with homeless individuals who have complex physical and mental health needs. However, in September 2013, the program hired a qualified full-time physician to oversee delivery of primary care to Project HEALTH participants.
- **Transitioning to a new electronic health record platform.** CHCS' original electronic health record (EHR) system—Anasazi—was designed for use in community mental health centers, and was not ideal for supporting the integrated care model used in this program. For this reason, CHCS began transitioning to Cerner Power Charts, a fully integrated and highly functional EHR system, in August 2014. The transition to the new system has been a slow and challenging one, due to inadequate technological resources and the time needed to train staff. In addition, the Cerner system does not allow for detailed documentation of information collected during mental health assessments. As a result, most mental health staff continue to use Anasazi. CHCS is currently working to improve the Cerner system and is optimistic that a full transition will be completed in the months ahead.
- **Changing the peer referral process.** In the third award year, CHCS modified its peer referral process to relax its requirements for how peers connect to and work with participants. Although health navigators officially assign participants to one peer for tracking purposes, participants are encouraged to meet and develop relationships with multiple peers during their time in the program. CHCS made this change based on the approach promoted by ViaHOPE, the training program through which Project HEALTH peers receive peer certification. Peer services are most effective and person-centered when participants are compatible with a peer and the peer is able to offer empathy on issues that are important to the participant. Although the health navigator attempts to make a good match during the initial assignment process, allowing participants to work with any peer helps compensate for the occasional poor match. In addition, this approach helps to leverage the full peer team's skills and availability.

3. How did unexpected events support or conflict with successful implementation of the innovation? (RQ 8)

In January 2015, CHCS began providing integrated physical and behavioral health care to select Prospects Courtyard residents through the state's 1115 Healthcare Transformation waiver.[11] CHCS initially offered these services to individuals with the greatest need for them, but plans to eventually offer integrated services to the broader Prospects Courtyard population. Although this development has the potential to affect members of the control group, who would not have otherwise had access to physical health services on the Haven for Hope campus, the expansion of integrated services took place late in the implementation process and was made

[11] Medicaid eligibility is not a requirement for use of the broader integrated physical and behavioral health services offered through Texas' 1115 Healthcare Transformation waiver.

available to only a limited number of residents during its pilot phase. Furthermore, CHCS is able to identify in the data those members of the control group who accessed medical services under the waiver, thus allowing Mathematica to control for the effects of those services in our analyses. Therefore, the effects of the expanded services should not affect our ability to detect differences in outcomes between the original intervention and control group.

CHCS also faced unexpected staffing challenges during the third year of the program when several members of the peer support team, who were in recovery from substance abuse or mental health conditions, experienced relapses. Their caseloads were absorbed by other peers and community guest specialists, some of whom struggled to maintain the increased workload. CHCS dealt with this challenge by reassigning caseloads among its workforce; in addition, one part-time employee moved to full-time status. As workforce members begin transitioning to different internal or external positions during the final months of the project, CHCS may also consider training peers and non-Project HEALTH staff to conduct follow-up assessments.

C. Program effectiveness

In this section, we present the results of our quantitative analysis of four intervention outcomes. The CHCS program is intended to reduce psychological stress, improve the individual's capacity for self-management, increase feelings of hope, and increase the individual's capacity for change. Our analysis focuses on four metrics related to these outcomes. As noted (Section III.A), we cannot analyze outcomes on CMMI's four core measures because we do not have data on the use of health care services or on expenditures. We begin the section by describing our quantitative methods. We then report the findings and limitations of these analyses. Next, we discuss qualitative findings about the perspectives of key informants on program effectiveness.

1. Overview of quantitative methods and population

CHCS provided us with data for members of the intervention and control groups; the data were from assessments administered at enrollment, at approximately 6 months post-enrollment, and at approximately 12 months post-enrollment. Our analysis included only the first two assessment points, because data from the 12-month post-enrollment assessment were only available for about half the members of the intervention and control groups.

We further limited our analysis to members of the intervention and control groups who completed both assessments; this sample includes about 82 percent of the intervention group and 80 percent of the control group. The excluded individuals could not be reached during the window of availability to complete a six-month follow-up assessment.[12]

The data included scores on the following four scales. See Appendix A for additional information on these assessments.

- **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 scale ranging from 0 (not at all) to 4 (extremely). Total scores, calculated by

[12] Based on conversations with CHCS.

summing the question ratings, range from 0 to 72, with higher scores indicating higher levels of global psychological distress.

- **University of Rhode Island Change Assessment (URICA).** The URICA assesses motivational readiness for change. Respondents 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). Responses within each stage are averaged to create a stage-level score.[13] These four stage scores are then combined to create a score indicating motivational readiness to change, which is equal to the sum of the contemplation, action, and maintenance scores, minus the pre-contemplation score. Ranging from 2 to 14, higher scores represent higher levels of motivational readiness to change.
- **Short Form 36 Health Survey, Version 1 (SF-36).** The SF-36 assesses quality of life across multiple domains, including physical functioning; physical and emotional role limitations; general health; pain; emotional well-being; social functioning; and energy or fatigue. Improvements on this measure may mean that CHCS is achieving its goal of improving the participant's capacity for self-management. Ratings of selected items within each domain[14] are combined into a composite measure (the SF-6D) ranging from 0 to 1, with scores closer to 1 indicating better quality of life.
- **Adult Hope Scale (AHS).** The AHS assesses the participant's feelings of hope on two subscales. The first subscale, pathways, measures the ability to plan routes to achieve desired goals. The second subscale, agency, measures the ability to initiate and sustain the use of those pathways. Participants rate 12 statements using a Likert scale ranging from 1 (definitely false) to 8 (definitely true). Item ratings from separate sets of four items are summed to derive subscale scores,[15] and the two subscale scores are summed to create a global score of hope, ranging from 8 to 64; higher global scores represent greater feelings of hope.

Core measures are not included in the analysis for this awardee. For both participant and control groups combined, a total of 47 individuals are enrolled in Medicare and 77 in Medicaid, representing less than one-quarter of the total number of individuals involved with this program. As a result, CMS administrative data are not an appropriate data source for this awardee. Further, CHCS's data files do not include information relevant to these core measures. We have not identified an alternative data source that fully encompasses the population and services needed to calculate the core measures.

[13] One of the eight questions in each of the four stages is omitted from the calculation of the stage-level score.

[14] Specifically, 11 items are used from the following domains: physical functioning (three questions), physical and emotional role limitations (two), pain (two), emotional well-being (two), social functioning (one), and energy or fatigue (one).

[15] The remaining four items are considered "distracter" questions and are not included in score calculations.

2. Quantitative findings on program effectiveness

None of the estimated impacts were statistically significant for any of the four assessment scores included in this analysis (Table III.1).

Table III.1. Impact estimates for CHCS, annual results

	Regression-adjusted mean for intervention group	Regression-adjusted mean for control group	Estimated impact[d]		
			Value	Percent	P-value
BSI-18 global distress[a]					
At enrollment	30.77	30.72	N.A.	N.A.	N.A.
6 months after enrollment	20.86	20.96	-0.14	-1	0.93
URICA readiness[a]					
At enrollment	9.85	9.75	N.A.	N.A.	N.A.
6 months after enrollment	9.11	9.35	-0.33	-3	0.11
SF-6D[b]					
At enrollment	0.56	0.57	N.A.	N.A.	N.A.
6 months after enrollment	0.61	0.62	-0.01	-2	0.46
AHS global hope[c]					
At enrollment	41.68	41.87	N.A.	N.A.	N.A.
6 months after enrollment	44.24	43.52	0.92	2	0.40

Source: Mathematica analysis of CHCS data for those enrolled between March 2013 and September 2014.

Note: To be included in the analysis, individuals in both participant and control groups must have had both baseline and follow-up scores. All regression models control for age, gender, employment status, insurance status, education status, living situation, and month of enrollment.

[a]n = 208 program participants, n = 203 control group members.

[b]n = 170 program participants, n = 159 control group members.

[c]n = 208 program participants, n = 202 control group members.

[d] We derived the impact estimates in Stata using the lincom command to compare the difference between the intervention and baseline period means for the treatment and comparison groups.

N.A. = Not applicable.

3. Limitations of the quantitative analysis

In light of the preliminary nature of the current findings, it is important to avoid drawing premature conclusions about the implementation or outcomes of CHCS' program. Limitations of the analysis include:

- **Restricted time frame of available data.** Findings reflect only the first six months of the year-long program. As data become available for the full period of the program, findings may change.
- **Small sample size.** The sample size for this analysis may be too small to detect small changes in outcomes. If new data become available for more participants and control group members for the 12-month follow-up assessment, the findings may change.

- **Lack of core measures.** Although we examined impacts on selected behavioral and physical health outcomes, we cannot assess the program's impact on core outcome measures.
- **Lack of a dosage measure.** We were unable to include a dosage measure to control for the varying degree of intervention services that different individuals received. Although data on the use of program services are available, we are working with CHCS to classify those services into meaningful categories. This metric will be incorporated into future impact analyses.

4. Qualitative findings on program effectiveness: From the perspective of key informants, to what extent does the intervention improve desired health outcomes and quality of life for participants? (RQ 23 and RQ29)

Overall, key respondents believe the program has had a positive impact on its participants' health and well-being. For example:

- Respondents said that many participants had better scores on clinical indicators such as blood pressure and hemoglobin A1c values.
- Key respondents reported that, because participants can access preventive care more easily, they may be using the emergency department less often.
- Although primary care staff struggled to link participants to affordable and available specialty care, they gave several examples of participants who successfully accessed much-needed medical services with the help of the Project HEATH team. Examples included a woman who had an ovarian tumor removed and a man who had surgery to remove long-standing cataracts.

Respondents reported that the mental health of some participants also improved significantly. For example:

- One participant who struggled with alcoholism is now living in the Haven for Hope campus dormitories and has greatly reduced his drinking. A respondent reported that "he is now living over here in the member's dorm, and is doing great, not drinking, attending all his groups. He is a success story. I'm proud of him."
- Another participant worked closely with a peer support specialist to access mental health services after confessing that he was suicidal: "Five months later, he has recovered from all drugs, and is now over here volunteering at Haven. You'll see him over there wearing a suit."

Key workforce staff emphasized that, for many participants, satisfying basic needs such as those for housing or food is a higher priority than improving their health behaviors. For this reason, linking participants to sustained housing is an important indicator of the program's success. Respondents reported that:

- Workforce members helped transition a number of participants into off-campus housing or wellness dorms on the Haven for Hope campus.

- The program helped many participants get back on their feet and regain their self-respect by connecting them with other basic needs such as food, transportation, or clothing.
- One respondent described how he was able to help a participant who needed a pair of black shoes to get a job: "And that's what's preventing him from getting his job, from moving on, getting out of here, and getting back into normal life. Simple things like that."

D. Context

1. To what extent did organizational features support or conflict with implementation? (RQ 104)

Key awardee staff generally viewed CHCS as supportive of the innovation and its goals. In particular, respondents reported that CHCS views Project HEALTH as an important first step toward the ultimate goal of integrating physical health care and the peer support role into existing programs and services. During the past year CHCS expanded the availability of integrated physical and behavioral health care through the state's 1115 Healthcare Transformation waiver. In that same period, CHCS also implemented an initiative to integrate peers into all four of its divisions and on the Haven for Hope campus.

Respondents believe the organizational culture of both CHCS and Haven for Hope has contributed to the successful implementation of the peer support program. Peers' tasks include connecting participants to the various programs and support services available on the Haven for Hope campus. The program's effectiveness relies on the campus and program staff recognizing the peers as respected colleagues. However, establishing this level of acceptance was a slow process; many staff were initially uncertain about the role of peers, and were hesitant to embrace colleagues who may have once resided in the campus shelter. To address this issue, senior staff at CHCS did outreach and training to clarify the peer role and highlight how integrating the peers would benefit other program services. This helped make the peers feel both recognized and respected. As one noted, "I feel like an asset, a somebody, over here."

As an organization primarily focused on providing mental health services, CHCS faced three notable challenges in establishing the necessary infrastructure to support a primary care clinic:

- First, CHCS bases its services on the recovery model, an approach that may conflict with the medical model used by many primary care physicians. As a result, CHCS encountered some difficulty in finding a PCP with an appreciation for the principles of a recovery model.
- Second, the organization lacked a strong provider network of medical supply vendors and specialty care services. As a result, the primary care staff had to advocate for supplies and take steps to develop a referral network of community health specialists.
- Third, CHCS' EHR system, as noted in section III.B, did not adequately support the documentation and tracking of primary care services. As a result, the organization transitioned to a new EHR system—one that serves the needs of both primary care and mental health service staff—midway through implementation. This process continues to be challenging because of insufficient resources and limited IT capacity.

2. To what extent did the policy and political environment support or conflict with implementation? (RQ 106)

The state of Texas did not expand Medicaid under the Affordable Care Act, and a large majority of program participants remained uninsured during Project HEALTH's implementation. This has had implications for the program's implementation. In particular, because the majority of participants are poor and uninsured, and because the primary care physician is not equipped to provide specialty care, many physical health needs have remained unmet. Despite significant community outreach by primary care staff, few local providers have volunteered to provide free specialty care to participants. The participants' poverty and their lack of insurance coverage have also prevented the primary care physician from prescribing adequate medications. Instead, she relies on cheaper drugs, which, in some cases, are less effective.

E. Summary and conclusions

Despite minor changes to its original protocol, including migration to a new EHR platform and modifications to the peer referral process, CHCS's implementation of Project HEALTH has been generally timely and conducted according to plan. Qualitative findings indicate the program has had positive effects on indicators of participant well-being and health, and a number of participants have transitioned successfully into off-campus housing or wellness dorms in Prospects Courtyard. CHCS views Project HEALTH as a first step in its efforts to expand the availability of integrated care and peer support services to other programs at CHCS. However, because the project represents one of the organization's first efforts to integrate primary care services, the Project HEALTH team encountered several significant challenges, including recruitment of primary care staff with a behavioral health background, development of a network of medical supply vendors and medical specialists, and adoption of an EHR system that successfully integrates documentation of physical and mental health services. In addition, the state's political environment has affected program implementation. Texas did not expand Medicaid under the Affordable Care Act, and the majority of participants remain uninsured; because they are also poor, they cannot afford to purchase specialty care for physical health needs that exceed those addressed by the primary care services offered through the program.

Despite positive reports from key informants about the program's effectiveness, we found no statistically significant impacts on the four outcomes we analyzed. However, the preliminary quantitative results presented here reflect changes in outcomes for only the first six months of the year-long program period, and the sample included in these analyses may be too small to detect small impacts. Furthermore, the analyses reported here included only a limited set of outcomes measures, which do not fully reflect the anticipated effects of the program. In addition, these analyses do not include a dosage measure of service utilization to control for the degree to which participants received program services.

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

IV. FELTON INSTITUTE

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, Felton began to operate the Prevention and Recovery in Early Psychosis (PREP) Program in San Francisco in 2007. An outpatient intervention program, PREP offered an integrated suite of evidence-based medication and psychosocial interventions to stabilize and promote the remission of schizophrenia in individuals ages 14 to 29. A transformative workforce of therapists, psychiatrists, psychiatric nurses, case aides, and employment and education specialists administered the interventions.

The PREP program included the following components: (1) community education to promote early identification of psychosis and entry into intervention, (2) early diagnosis, (3) algorithm-guided medication management that would promote long-term compliance with medication regimens as well as stable remission, and (4) 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.

Felton replicated the San Francisco PREP program in Alameda and San Mateo counties, both of which sustained the program with funding generated by the California Mental Health Services Act of 2004. Felton used HCIA funding to expand and launch the program in two additional counties: San Joaquin and Monterey.

The PREP program targeted individuals ages 14 to 29 who had either prodromal symptoms of schizophrenia or early onset schizophrenia. In the two HCIA-funded sites, the PREP team sought to administer the Structured Clinical Interview for the DSM-IV (SCID) to at least 140 individuals, most of whom were expected to be eligible for PREP services.

The findings in this report are based on qualitative data that Mathematica collected or received by June 1, 2015, as well as enrollment data reported by Fulton throughout the award period. Data sources included:

- **Enrollment data** submitted by Felton to the reporting website maintained by the Lewin Group for HCIA, Round 1.
- **Qualitative data** collected by Mathematica during a site visit and telephone interviews in May 2015. 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. Results from the participants' focus group are not included in this report, but our seventh quarterly report does include an in-depth analysis of the focus group data.

At this time, we cannot conduct an impact analysis because data limitations have prevented us from identifying an appropriate comparison group. These limitations are discussed in detail in Section C.

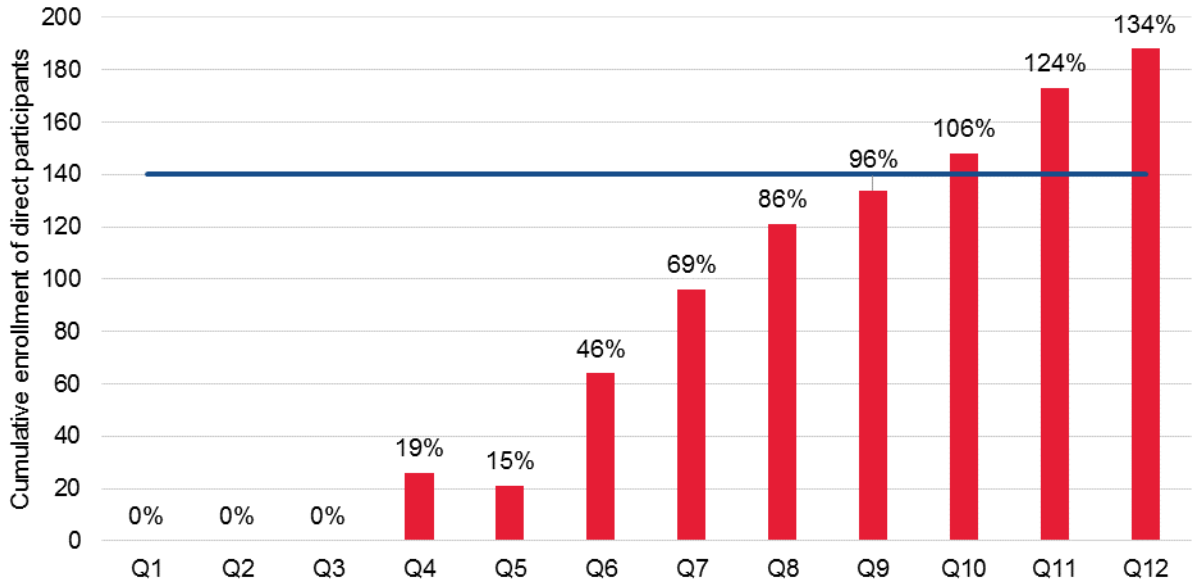
The questions that we address in this report were selected in consultation with our COR. Questions related to implementation effectiveness are discussed in Section B; to program effectiveness, in Section C; and to the program and policy context, in Section D.[16]

B. Implementation effectiveness

1. Who was enrolled in the program?

By the end of the program's 12th quarter (June 30, 2015), Felton had administered the SCID to 188 "unique" participants in the two HCIA-funded sites (Figure IV.1), surpassing its total goal of 140 participants. Felton defined a unique participant as someone who was referred to the PREP program and who completed the SCID assessment. Not all individuals who completed the SCID, however, were found to be eligible for PREP services. Enrollment began in the fourth quarter, rose steadily through quarter 9, and surpassed the enrollment goal of 140 in quarter 10.

Figure IV.1. Percent of target enrollment achieved by quarter, Q1–Q12—Felton



Source: Awardee's self-reported enrollment data, Lewin quarterly reports.

Note: As displayed by the vertical bar, Felton's target enrollment was140 unique participants.

2. To what extent was implementation timely, conducted as planned, and responsive to site-level constraints? (RQ 17)

Although Felton successfully implemented the PREP program in the two HCIA counties, four noteworthy challenges limited the reach of the program.

[16] Felton may conduct analyses of data that they have collected on program participants and publish their own reports about program outcomes. These analyses are likely to use different data, focus on different outcome measures, and address different specific questions compared with the data and measures we use in this report to address questions of direct interest to CMMI. As a result, Felton may reach conclusions about program effects that differ substantially from the ones we present in this and future reports.

- **Establishing an effective referral system.** Despite an effective, efficient referral system, several referral-related challenges emerged as the program was being implemented. For example:
  - Although referrals came from a range of organizations, the majority originated in the county mental health department. It took a significant amount of time for county staff to be fully trained in the tenets of the PREP program and in understanding the referral criteria.
  - In the early months of implementation in particular, a large number of people referred to the PREP program were ineligible for services. As one PREP team member explained, "County staff will get mad, they'll say, 'Oh I referred this person to you and you didn't take him,' and we'll say, 'Well, he actually has bipolar not schizophrenia'." So we get a lot of pressure, but . . . it's a friendly pressure because they really know what we do works."
  - Inappropriate referrals resulted in the misuse of resources and in disappointment among the potential participants and their families. PREP and the counties therefore invested significant resources in training staff, and as a result, the proportion of inappropriate referrals declined over time.
- **Engaging the broader community.** The size and rural nature of the HCIA counties required a high level of community outreach and engagement. Moreover, some county organizations—particularly school districts—were reluctant to work with the PREP program at all. As one PREP staff member explained, "Collaborating with all the other agencies is sometimes challenging because we're a new thing, and we're changing the way that people think about and approach psychosis and this population." Establishing these necessary collaborative relationships required significant outreach on the part of PREP staff.
- **Recruiting and retaining staff.** Hiring and retaining qualified staff in the two HCIA-funded programs—which are in larger, more rural counties (San Joaquin and Monterey) than the original three—was a challenge for Felton. For example:
  - Felton preferred candidates with roots in the local communities and an understanding of local resources; it was difficult to achieve this, however, because of the relatively small hiring pool in San Joaquin and Monterey counties. For instance, the psychiatrist for the program in Monterey County commutes once a week from San Francisco.
  - Once hired, PREP staff took part in an intensive training regimen on the various evidence-based practices they would be using in their new roles. This specialized training increased the staff members' knowledge and skill sets, making them more attractive candidates for other open positions in their field. As a nonprofit organization, Felton could not provide the most competitive salary or benefits package relative to those offered by other local organizations. Consequently, Felton struggled to retain staff because they were enticed by more lucrative employment opportunities.
- **Identifying transportation options.** The counties that were home to the two HCIA-funded programs are large, rural, and lacking in public transportation systems. PREP staff therefore had to travel more extensively, and many participants could not access the PREP office.

- Compared with staff in the original three counties, which were smaller and more urban, PREP staff in San Joaquin and Monterey counties spent a significant amount of time traveling to meet participants in their homes or communities. As one PREP team member noted, "We need to drive far away to provide the services, to connect with communities. For people to come into the program is another challenge because transportation is not like living in San Francisco. They have transportation everywhere, and here it's not that. So for clients or for us, it's kind of difficult."
- Although the PREP teams found ways to structure their schedules to accommodate field visits, the lack of transportation prevented many families from participating in the multifamily group sessions, a critical component of the PREP program. According to one team member, "It's late at night, and family members might want to come, but they're afraid to come by themselves in the dark on the bus. And we also don't have a car or a van to go pick people up. We don't have our own thing. We can't drive them; we don't have that option. We finally have [the multifamily group] going, but it took a long time to really get it going and we only have three or four families that come. So that's a challenge for us and I think it has to do with where we are regionally."

3. How did unexpected events support or conflict with successful implementation of the innovation? (RQ 8)

No unexpected events affected the successful implementation of the innovation.

C. Program effectiveness

As noted in Section A, we cannot use quantitative data to measure program effectiveness. The following specific limitations in data availability make an impact analysis infeasible at this time:

- **We do not have data for a comparison group.** To develop a comparison group, we may be able to use data provided by mental health departments in the participating counties. These data could be used to identify individuals who meet PREP referral criteria but who have not been referred to the program. To date, however, we have received data files from only two counties. Early assessments of the data in these files suggest that (a) they may have serious gaps in comprehensiveness and comparability, and (b) we may not be able to identify which individuals in these files are members of the intervention group because at least one county mental health department did not include data that could be linked to Felton's assessment and encounter data system. Using the county data to create a comparison group is further complicated by the fact that most individuals who met the eligibility criteria were referred to PREP, thus limiting the size the pool from which potential comparison group members can be drawn. In addition, because it is difficult to diagnose schizophrenia, the number of individuals with this disease is likely to be underreported in the county data.
- **Felton's assessment and encounter data system contains information about the intervention group only.** As a result, we cannot use these files to identify a comparison group. The files also include data on hospitalization and emergency department use that are entered manually by staff on five county teams. We have not been able to confirm the accuracy or the consistency in this data entry process. Moreover, hospitalization and emergency department data are not uniformly available for all participants. For example, the

data were available only for the period in which a participant received the intervention, which ranged from 6 to 24 months. Felton is in the process of making these data equally available for a segment of participants.

If developing a comparison group becomes possible, we will need to consider strategies for mitigating the risk of an underpowered impact analysis. Felton's sample is small, and as a result, we may want to consider adjusting the usual level of significance and conducting only certain sensitivity tests, depending on the availability of data. The sensitivity tests would provide evidence about the likelihood that specific program effects had actually occurred. We will continue to work with Felton to obtain appropriate data and, depending on the results of our efforts, we may be able to conduct impact analyses and include the results in the next annual report.

1. From the perspective of key informants, to what extent does the intervention improve desired health outcomes and quality of life for participants? (RQ 23 and RQ29)

All the stakeholders we interviewed—including members of the workforce, program managers, and county representatives—strongly believe that, for the majority of participants who received PREP services, the quality of both their health and their life improved as a result of the program. This is particularly true when we look at the limited options that were available in the county before the PREP program was implemented. According to workforce staff, it is the collection of therapies and evidence-based practices, not any one component that makes PREP effective. As one staff member recounted:

> "It's all the factors. Everything. Resuming the therapy from the clinicians, receiving services from the voc-ed care advocate, assisting the families, helping them understand what's happening, and also being consistent with taking the medication. Some of our clients will refuse to take medication when they begin, and you know, we're still able to help them without the medication, but then we eventually guide them to that extra support."

Although the staff indicated that all PREP components were important to the program's success, several highlighted the value of the CBTp, which taught participants critical coping skills. Within a few months of receiving PREP services, most participants began to show an improvement in this area. As one staff member stated:

> "I think some of my clients, they're becoming, they haven't fully yet, but they're becoming goal-oriented, you know, such as being able to be out in the community, or beginning to take a class, or even just having the thought of 'I can take a class' is a lot for some of my clients. But they're on that spectrum, whether they're becoming goal-oriented, they're able to see what their capacities are, as well as how to get to the goal. That's definitely something that I see in my clients."

The ultimate objective of PREP was to help participants become stable, goal-oriented, and independent within two years of when they entered the program. Accordingly to PREP staff, most participants reached this goal—or "graduated"—within one year of when they entered the

program. Within a relatively short period of time, therefore, staff observed significant changes in participants' behaviors. In the words of one staff member:

> "I was attending some graduations and the expression that [the participants] had . . . 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. Context

1. To what extent did organizational features support or conflict with implementation? (RQ 104)

Three organizational features in particular affected Felton's implementation of PREP: Felton's reputation and relationship with external stakeholders, the commitment of Felton's senior leaders to PREP, and the remote location of field staff working far from Felton's headquarters.

First, Felton's positive reputation made it easier for the organization to work with external stakeholders. Before PREP was implemented, some county staff, local providers, and community stakeholders were apprehensive about an out-of-county organization providing services to residents and their clients. These concerns were alleviated somewhat by Felton's reputation for providing social services to vulnerable populations. As explained by a representative of a county mental health department:

> "We [the county] try and be most effective; we try and bring in evidence-based practices. But I think what was attractive about Felton Institute was they really had a good model. They've been working in other counties, so they had clout and a history of working well in other counties where they could say, 'We have found that if you do X, Y, and Z, you'll improve outcomes and reduce costs.' So who wouldn't want that?"

Second, members of the workforce highlighted the accessibility of Felton's senior leaders and their commitment to PREP as critical to the successful implementation of the program. As one team member said:

> "At Felton, Bob [the CEO] cares about PREP. He is invested in this program; he can speak passionately about it. When something comes up, he is like, 'How are we going to make this work?' He is there for this program and has been a huge asset."]

In addition, many team members mentioned that they were drawn to PREP not only because of its leaders' commitment to evidence-based practice but also because of the opportunity to develop their own knowledge and skills in this area. Felton's investment in developing its capacity to provide field staff with intensive and ongoing training, coaching, and support contributed to workforce satisfaction. According to one member of the workforce, "Working for a company that really values evidence-based science is really invigorating for me."]

Finally, PREP field staff worked far from Felton's San Francisco headquarters; this organizational characteristic posed a challenge to program implementation. But Felton was able to mitigate this challenge by investing in videoconferencing technology, which allowed the mostly San Francisco-based clinical support team to provide remote coaching to the PREP field staff. Several members of the field staff recognized and appreciated Felton's communication efforts and its commitment to them. However, some field staff said that they felt isolated and frustrated when Felton made decisions about the PREP model without much consultation regarding the unique county context in which the individual programs operated. As one workforce member noted, "Things keep changing, the rules keep changing, and sometimes things aren't communicated to us."

2. To what extent did the policy and political environment support or conflict with implementation? (RQ 106)

The policy and political environment in San Joaquin and Monterey counties supported the implementation of the PREP program. From the perspective of county stakeholders, PREP was appealing because it offered a proven solution to a problem facing the county mental health system: treating a complex and expensive disease. According to one county stakeholder,

> "Monterey County is very committed to this whole concept of 'recovery happens,' so we want to move away from the idea that once an individual has a psychotic episode that forever they will be ill, they will be doomed to a life in chronic mental health care systems and end up living in boarding care. We've really been trying to say, okay, we need to do things differently; people can recover from having a psychotic illness. So that thinking already was here in the county."

In addition, county stakeholders expected the collaboration with Felton's PREP program to improve the county's capacity to work with the broader population of individuals with schizophrenia. In this context, Felton's introduction of evidence-based practices to Monterey County produced a positive spillover effect. Not only did the presence of the program increase awareness of the disease and its treatment options, but Felton is also exploring options for training county staff to administer some of the treatment techniques to individuals who are not eligible for PREP services, such as those who do not meet the age criteria.

Finally, California's Mental Health Services Act of 2004, which provides funding for prevention and early intervention services, created an incentive for counties to adopt or to collaborate with programs like PREP. As this report was being written, both counties were intending to use MHSA funding to sustain the PREP program.

E. Summary and conclusions

In implementing the PREP program in San Joaquin and Monterey counties, Felton benefitted from the lessons it learned when it implemented the program in San Francisco, Alameda, and San Mateo counties. Despite this experience, Felton encountered some county-specific challenges, which were related to establishing an effective referral system, hiring and retaining a local workforce, and managing a program in two geographically large and mostly rural counties. The organization did, however, navigate these challenges, and when this report

was being written, both counties were on course for maintaining the program beyond HCIA funding, an indicator of implementation success.

Information on whether PREP had an impact on health, service, and cost outcomes is not yet available, as explained in Section C. We hope to be able to conduct impact analyses in the future.

V. FEINSTEIN INSTITUTE FOR MEDICAL RESEARCH

A. Introduction

Staff at the Feinstein Institute for Medical Research (Feinstein), the research division of the North Shore-Long Island Jewish Health System in New York, used HCIA funding to implement the Improving Care-Reducing Costs (ICRC) program. This program aimed to lower the costs of care and improve the mental health status of 770 individuals with schizophrenia, schizoaffective disorder, or psychotic disorder not otherwise specified, who were at risk for hospital readmissions. In addition to relapse prevention counseling, 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 serving program participants in 8 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. These were mental health/health technology (MH/HT) case managers; they provided relapse prevention counseling, training, and ongoing support to participants during the six-month program period,[17] and facilitated the participants' use of complementary technologies and services.

To help assemble the package of complementary technologies and services for participants, Feinstein partnered with technology developers and researchers at six other institutions: Dartmouth College, Boston University, the University of Minnesota, the University of Pittsburgh, Proteus Digital Health, and the Nathan S. Klein Institute for Psychiatric Research.

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

- Between the ages of 18 and 60
- Clinical diagnosis of schizophrenia, schizoaffective disorder, or psychotic disorder not otherwise specified
- Hospitalized or within 30 days post-psychiatric hospitalization at enrollment
- Enrolled in or eligible for a Medicaid program, privately insured, or uninsured

Participants had to be able to participate in research assessments conducted in English and be able to give fully informed consent. Before recruiting ICRC participants, each participating clinic also enrolled about 10 individuals to serve as a standard reference group. Standard reference group members met the same criteria as the target population and received the same assessments that ICRC participants did, but they did not participate in the ICRC program. These members of the standard reference group were excluded from the analyses in this report.

[17] Program participants received program services for six months.

The findings in this report are based on qualitative data Mathematica collected or received by June 1, 2015, as well as enrollment data reported throughout the award period. Data sources included:

- Enrollment data submitted by Feinstein to the reporting website maintained by the Lewin Group for the HCIA initiative
- Qualitative data, including telephone interviews and an in-person site visit in May 2015. Mathematica conducted in-depth interviews with key awardee staff, members of the workforce, and other stakeholders. In addition, Mathematica convened focus groups with members of the workforce and with ICRC program participants. Data from the latter are not part of this report; instead, we provide an in-depth analysis of those focus groups in our seventh quarterly report.

This chapter does not include analyses of CMMI's four core measures, because we do not yet have sufficient quantitative data to evaluate the program's impacts. We expect to conduct outcome and impact analyses for Feinstein in the upcoming year, as data on more participants become available. We will report the results of those analyses in our third annual report.

The research questions we address in this report were selected in consultation with our COR. Specifically, we examine questions related to implementation effectiveness in Section B, program effectiveness in Section C, and the program and policy context in Section D. We close with an overall summary and conclusion (Section E).[18]

B. Implementation effectiveness

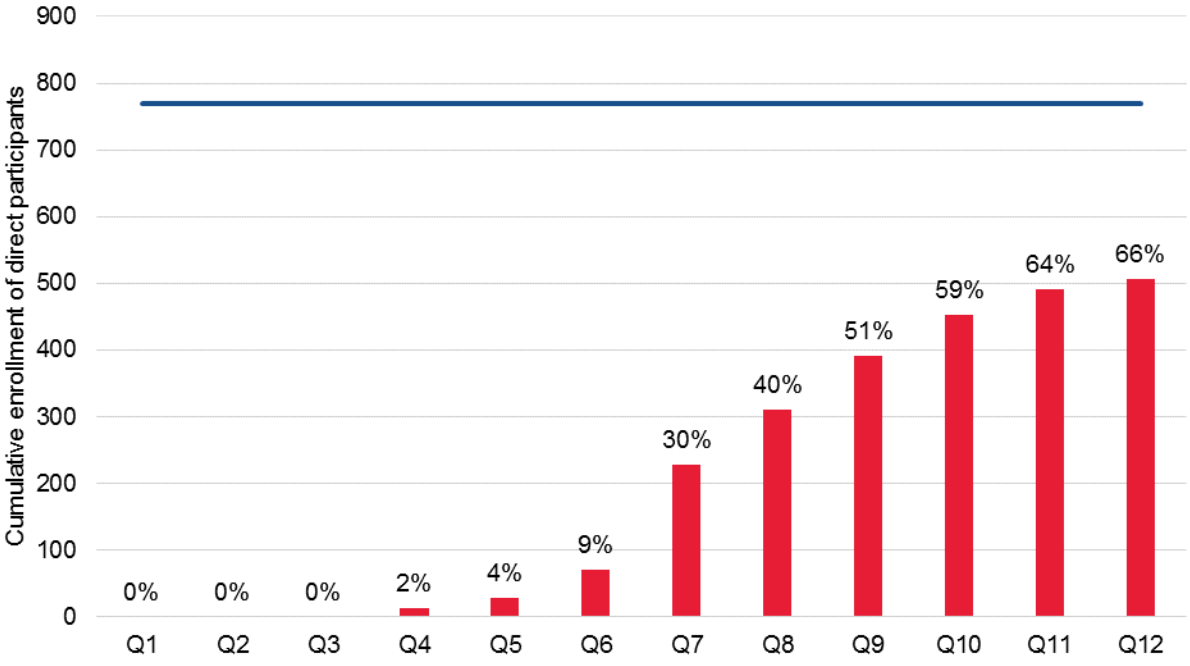
1. Who was enrolled in the program?

By the end of the 12th quarter (June 30, 2015), Feinstein and its partners had enrolled 506 unique participants, about 66 percent of the total enrollment target of 770 participants (Figure V.1). Recruitment for the program was initially delayed because Feinstein staff were notified of the award later than they expected to be.[19] ICRC program leaders worked quickly with each site's staff to develop recruitment plans and outreach materials after they learned about the award, and they began enrolling participants in Q4. Enrollment increased steadily over the course of the award, and Feinstein reported that its overall recruitment rate increased in the later quarters of the project. However, Feinstein ultimately did not reach its enrollment target of 770 participants.

[18] Feinstein may conduct analyses of data that they have collected on program participants and publish their own reports about program outcomes. These analyses are likely to use different data, focus on different outcome measures, and address different specific questions compared with the data and measures we use in this report to address questions of direct interest to CMMI. As a result, Feinstein may reach conclusions about program effects that differ substantially from the ones we present in this and future reports.

[19] Feinstein received formal notification of the award on November 1, 2012.

Figure V.1. Percent of target enrollment achieved by quarter, Q1–Q12—Feinstein



Source: Quarterly reports submitted to Lewin Group's reporting website for HCIA initiative.
Note: As displayed by the horizontal bar, Feinstein's target enrollment was 770 unique participants.

2. To what extent was implementation timely, conducted as planned, and responsive to site-level constraints? (RQ 17)

Feinstein was largely successful in implementing the program in a timely manner. However, respondents identified some challenges to achieving the program's goals:

- **Participant recruitment.** As described in the first annual report, the late notice of award caused delays in the program's initial efforts to recruit and enroll participants. To accelerate enrollment, program leaders worked with staff at individual sites to tailor specific recruitment and outreach plans to each site. To encourage more referrals, ICRC program staff continued to reach out to their referral sources over the course of the award to give them information and education about the ICRC program and share success stories about the participants.
- **Case manager role.** Program leaders noted that one of the largest initial challenges was orienting the MH/HT case managers to the idea that the ICRC program required outreach and follow-up in the community, with more intensive outreach in the beginning to engage participants who were newly released from the hospital and might otherwise be lost to care.
  - One MH/HT case manager described how the role of the ICRC case manager departed from standard practice in mental health treatment, noting that in standard practice, if a patient missed three appointments with a therapist, "usually your case would be closed within the clinic, but instead we show up at your house and are knocking on your door, trying to engage you."

  - A program leader commented that it was important and challenging to communicate to case managers that, in contrast with other interventions that require participants to be relatively stable, "in order to start delivering interventions and engaging deep in the [program's health] technology, they didn't have to wait until people were 'well enough' to participate."
  - Program leaders noted that individual case managers' interest in and willingness to adapt to this new role were critical to success of the program.

- Ongoing virtual training for dispersed site staff. Program leaders mentioned some challenges associated with providing ongoing training and technical support to program staff located in sites around the country because "some people aren't great at interacting on the telephone." Respondents at the leadership and frontline levels reported that initial, in-person training was important to establish a foundational understanding of the program and its components and to help program leaders develop a rapport with program staff.
- Continued involvement of technology developers. Initially, program leaders expected to "get the technologies up and running and then they'd just kind of take care of themselves." In reality, program leaders found that the entire ICRC program required ongoing attention from technology developers. For example, smartphone applications often needed changes when platforms were updated, requiring "constant minor tweaking." Technology developers were available throughout the award period to address technology upgrades and provide technical assistance.
- Ingestible sensor and personal monitor. Feinstein did not implement one of its program components, the MIND1 ingestible sensor and personal monitoring patch (formerly Proteus), as originally planned. We describe this divergence from the initial plan next.

3. How did unexpected events support or conflict with successful implementation of the innovation? (RQ 8)

In addition to its early challenges associated with the late notification of award, Feinstein had intended to deliver the ingestible sensor component of the ICRC's health technology program (HTP) to 100 participants at the Zucker Hillside Hospital site; however, it was unable to acquire the technology in time to provide it during the HCIA award period as planned. 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. Feinstein intends to provide this technology to individuals who meet the same criteria as the ICRC participants as part of a separate, future research study.

C. Program effectiveness

We do not present analyses of the four core measures because we do not yet have enough quantitative data to evaluate program impacts. We expect to conduct outcome and impact analyses for Feinstein in the upcoming year, as data lags diminish and data on more participants become available, and report our results in the third annual report. We do, however, report on perceptions of the program's effectiveness from the perspectives of key awardee staff, members of the workforce, and other stakeholders.

1. From the perspective of key informants, to what extent does the program improve desired health outcomes and quality of life for participants? (RQ 23 and RQ29)

Program leadership believed the program improved desired health outcomes for participants. For example:

- Leadership reported that hospital readmission rates decreased by half for participants in the HTP compared with the rates in the program's standard reference group over a six-month period.[20]
- The leadership emphasized that if participants completed the six-month program, "we made an impact in the sense that we've gotten people over the highest risk period" for hospital readmission.
- Leaders also noted that although some participants returned to the hospital after the six-month program ended, participants did well while enrolled in the program. As one respondent commented, "Unfortunately, once they finished, sometimes they were readmitted without that extra support. But I think while they're in the program, it's an accomplishment."

Staff working in the implementing sites provided anecdotal examples of the program's impact. For example, one respondent described a participant who, at the beginning of the program, "couldn't get through a conversation without crying, and she was so paranoid that she wanted to live in the hospital." After participating in the program and working with the MH/HT case manager, "she no longer sees the hospital as the place that she's going to be the most safe. That's not where she goes anymore. She smiles and she's like a totally different person."

Many program leaders and frontline staff at implementing sites also remarked on the benefits of the HTP technologies for ICRC participants. For example:

- Respondents described how the technologies could align with participants' personal preferences. As one respondent stated, "Patients can get treatment whenever they wish. If they want to do CBT [cognitive behavioral therapy] at three in the morning, they can do CBT at three in the morning."
- Some of the frontline staff and leaders commented that the technologies helped participants understand that other people experience similar symptoms or challenges. For instance:
  - Participants recognized their own experiences in examples provided in the online CBT program and were "amazed to learn that other people had the same experiences."
  - Participants used the daily support website to connect with and learn from other people with similar experiences.

[20] Leaders' conclusions were based on program assessment data analyzed by the Feinstein Institute. Mathematica cannot confirm the accuracy of these statements.

Several program leaders and prescribers noted that the program's prescriber decision assistant (PDA) allowed prescribers to be more aware of medication side effects and to address them, which in turn helped participants adhere to their medication regimens. For example:

- Respondents noted that because of the PDA's structured questionnaire and associated prompts, prescribers learned about the sexual side effects experienced by participants and were able to adjust their medications accordingly. One program leader commented that "prescribers had a real shock and they found out the actual prevalence of these [sexual] side effects among their patients."
- Some said that participants appreciated the PDA because it led to "more thorough" appointments with prescribers.
    - One respondent noted that participants did not appear to mind answering the same questions each time, because it helped them remember information to tell the doctor and "explain their situation or how they're feeling."
    - Another respondent said, "I would ask, 'Have you noticed any shaking of your hands?' And [the participant] would realize, oh, yeah, I do have some of that going on."

Many respondents who worked in the implementing sites described the positive effects of giving smartphones to the participants. The advantage that was most commonly cited was being able to reach participants easily—many of these participants had unreliable access to telephones before they entered the program. In addition:

- One MH/HT case manager noted that simply "having access to the cell phone was paramount to people's stability. Not only for their ability to communicate with providers, but also with family members. And for folks that are looking for housing, employment, or school, they have a phone number that people can call."
- Another MH/HT case manager said she tried to "incorporate the phone as much as possible with [the participant's] other goals. So if listening to music was helpful for coping, we'd download Pandora, or if socializing was a goal, we would do Facebook. A lot of people had weight loss goals, so we'd download healthy living apps and exercise apps."

Nearly all respondents commented on the benefit of having the MH/HT case manager available to participants throughout the program, and cited several reasons for this:

- One program leader described the MH/HT case manager's role as "the spoke in the wheel" of the ICRC program.
- Several respondents remarked that it was important to have someone in the case manager role to help ease the participants into using technology.
- In addition, respondents noted that the MH/HT case managers' monitoring of participants' use of the technology alerted the staff to changes and prompted them to follow up with participants when needed. For example, if participants stopped responding to the FOCUS app as usual, the MH/HT case manager could see that and reach out to the participant.

Respondents also said the relationships that participants developed with their MH/HT case managers and the more intensive follow-up support these staff sometimes provided had positive effects on the participants. Several respondents provided examples of how this increased engagement appeared to benefit participants. For instance:

- One respondent commented that after six months in the program, many participants were "in a much more stable situation socially in terms of housing and employment or income."
- Another respondent said that by working with case managers, participants knew "there's a person that they can turn to who will help get them to where they need to be."
- One MH/HT case manager described the positive effects she's seen in her work with participants, noting that it was rewarding to see them "have some insight about their illness and be linked to services and participating in their treatment, and eventually finding a job or reconnecting with their family members and their friends."

D. Context

1. To what extent did organizational features support or conflict with implementation? (RQ 104)

Respondents described several unique characteristics of Feinstein and its implementing sites that facilitated implementation. For example:

- The Feinstein leadership team had significant experience conducting research on schizophrenia and relapse prevention in partnership with mental health sites across the country. Leaders were able to bring this experience to bear as they implemented the ICRC program.
- Feinstein created a strong central administrative support structure to facilitate ICRC program implementation at participating mental health centers around the country. For example, MH/HT case managers had access to a research coordinator at Feinstein who provided "real-time tech support" to troubleshoot problems with the program's health technologies. MH/HT case managers also received ongoing group telephone consultation from the researchers who developed the program's components.
- Feinstein staff selected sites they had worked with on earlier research projects, and noted common implementation facilitators across sites:
    - Respondents from one site noted that their experience participating in research was an asset and they already had the infrastructure in place to carry out this program's research.
    - Respondents from the same site also said that "an inpatient unit and the continuum of care for behavioral health within the organization certainly helped." Program leaders echoed this sentiment, reporting that implementing sites affiliated with larger hospitals or health systems tended to have greater access to resources and 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.

- ICRC program leaders reported leadership at these sites knew they were likely going to "lose money on the deal but they were more about trying to improve care" and were willing to absorb the cost.

2. To what extent did the policy and political environment support or conflict with implementation? (RQ 106)

Disparity in states' reimbursement practices for case management led to some differences between the sites in terms of the MH/HT case manager's responsibilities. In particular, New York State does not reimburse for case management services. As a result, the award paid for 100 percent of the cost of MH/HT case managers in the New York site; these case managers provided traditional case management services in addition to supporting participants in the use of the program-specific technology and offered the program's relapse prevention counseling. The amount and type of case management varied at the other sites, and many case managers reported working with existing, traditional case management staff at their sites to coordinate care for shared participants. These different reimbursement practices required flexibility in the program's structure so it could accommodate the various financing systems.

Staff at one implementing site noted they faced some recruitment challenges because this program was similar to existing programs that offer care coordination to this population. A respondent noted challenges with the local Medicaid agency, saying that Medicaid "felt like it was mutually exclusive to be in this program and then also to be on the ACT team, so we weren't able to recruit from that [ACT team participants] group." However, site staff did say that staff at the Medicaid agency, and at the state health agency, were otherwise supportive of the program.

E. Summary and conclusions

Once Feinstein was notified of its award, it was able to ramp up recruitment and enrollment. Ultimately, however, Feinstein was unable to reach its goal of 770 program participants. Although Feinstein fell short of its enrollment target, its staff worked throughout the award period to address challenges and adapt the program's implementation as needed. For instance, technology developers stayed engaged throughout implementation to provide any necessary updates and technical assistance, although program leaders did not initially anticipate this ongoing role. Moreover, nearly all interview respondents believed the on-demand support from smartphone applications and online CBT, and the extra help from case managers, had considerably improved participants' health outcomes and their quality of life. Certain characteristics of the Feinstein Institute and its participating mental health centers, such as their involvement in previous studies and their affiliation with inpatient units, aided in the program's implementation.

Although we do not yet have the quantitative data we need to assess program impacts, we plan to conduct outcome and impact analyses for Feinstein in the future and report those results in our third annual report.

VI. FUND FOR PUBLIC HEALTH IN NEW YORK

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, is partnering with the Division of Mental Hygiene (DOHMH), in New York City's Department of Health and Mental Hygiene, to implement Parachute NYC. This project aims to improve care and lower costs for adults in New York City who experience a mental health crisis and a first episode of psychosis or severe mental illness. Parachute NYC's goal is to remake the usual crisis model of care and focus on patient-centered care; long-term, community-integrated treatment; and increased access to primary care services.[21]

As part of its need-adapted treatment model (NATM), Parachute NYC uses intentional peer support (IPS), need-adapted mobile crisis teams (NA-MCTs), and a crisis respite center (CRC) in each of four boroughs (Brooklyn, the Bronx, Manhattan, and Queens), as well as a citywide support line. Parachute NYC is the first large-scale implementation of the NATM in the United States. This model is similar to the open dialogue model, which had positive outcomes in Finland when used to treat psychosis.

Although the CRCs were newly introduced with the Parachute NYC program, the mobile crisis teams were already in place before the program began. However, these teams were trained on the new IPS and NA-MCT treatment modalities and asked to incorporate peers into their teams and their treatment practices.

FPHNY targets New York City residents ages 18 to 65 who experience a mental health crisis and a first episode of psychosis or severe mental illness. In Brooklyn, FPHNY expanded the age criteria to include youth ages 16 and older.

The findings in this report are based on qualitative data collected or received by June 1, 2015, and enrollment data reported throughout the award period. Data sources included:

- Enrollment data submitted by FPHNY to the reporting website maintained by the Lewin Group for the HCIA initiative
- Qualitative data, including phone interviews and an in-person site visit in April–May 2015. Mathematica conducted in-depth interviews with key awardee staff, members of the workforce, and other stakeholders. In addition, Mathematica convened focus groups with members of the workforce and with program participants. Data from the latter are not part of this report; instead, we provide an in-depth analysis of those focus groups in our seventh quarterly report.

This chapter does not include findings on CMMI's four core measures, because very few FPHNY participants are enrolled in Medicaid or Medicare.

[21] FPHNY received a no-cost extension to continue program and evaluation activities through 12/31/2015.

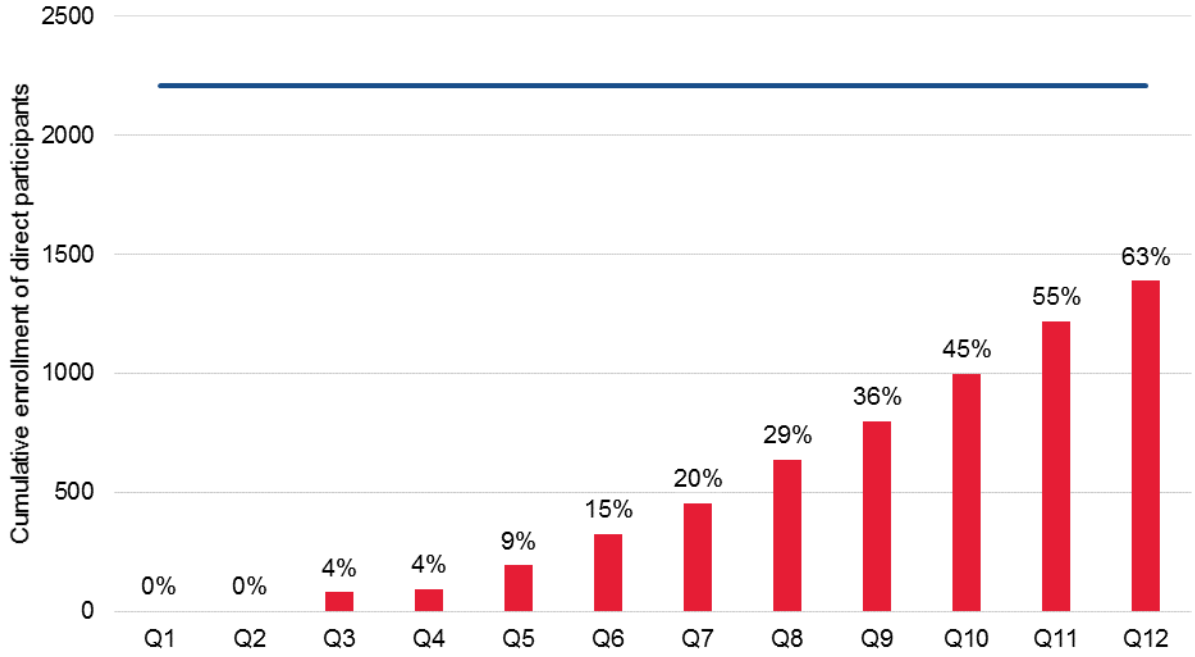
The research questions we address in this report were selected in consultation with our COR. Specifically, we examine questions related to implementation effectiveness in Section B, program effectiveness in Section C, and the program and policy context in Section D. We close with an overall summary and conclusion (Section E).[22]

B. Implementation effectiveness

1. Who was enrolled in the program?

By the end of the 12th quarter (June 30, 2015), FPHNY had enrolled 1,388 program participants, 63 percent of its enrollment target (Figure VI.1). In the first two quarters, the FPHNY project staff were establishing the program. Enrollment began in the third quarter with 79 participants and grew steadily each month.

Figure VI.1. Percent of target enrollment achieved by quarter, Q1-Q12—FPHNY



Source: Quarterly reports submitted to the Lewin Group's reporting website for the HCIA initiative.
Note: As displayed by the horizontal bar, FPHNY's target enrollment was 2,208 unique participants.

[22] FPHNY may conduct analyses of data that they have collected on program participants and publish their own reports about program outcomes. These analyses are likely to use different data, focus on different outcome measures, and address different specific questions compared with the data and measures we use in this report to address questions of direct interest to CMMI. As a result, FPHNY may reach conclusions about program effects that differ substantially from the ones we present in this and future reports.

2. To what extent was implementation timely, conducted as planned, and responsive to site-level constraints? (RQ 17)

FPHNY implemented all components of Parachute NYC on a schedule that generally ran according to plan in all four of the targeted boroughs. In addition, the project staff were responsive to site-level constraints. For example:

- In Brooklyn, consistently low enrollment prompted an expansion of the eligibility criteria. The Brooklyn CRC was initially limited to participants between the ages of 18 and 30 who were experiencing a first episode of psychosis. Because the services were underutilized, this requirement was expanded to include all individuals 16 and older who were experiencing any serious mental illness. The result was a significant increase in the volume of participants served at the Brooklyn CRC and a corresponding reduction in vacancy rates.
- Based on feedback from staff, Parachute NYC developed a formal protocol to use with hospital emergency departments (EDs); it detailed the referral process (for example, who pays for the taxi from the ED to the CRC) to make it easier for ED staff to refer a person to Parachute NYC.

Several challenges delayed or hindered the achievement of Parachute NYC's goals. These included:

- **Integrating peers.** Initially, many staff members reported challenges in integrating the supporting peers into the existing care teams. For example:
  - Several members of the mobile crisis teams, both peers and clinicians, reported they had trouble working together. Peers sensed a lack of respect for their opinions and did not feel truly accepted in the teams. 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 peers would relapse.
  - CRC staff also reported difficulties with integrating the peers into the teams in the beginning. 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.
  - Both Parachute NYC staff and DOHMH leaders agreed that it got easier to integrate the peers over time, but there was still significant room for improvement.
    - CRC leaders reported that they developed a deeper knowledge of the characteristics of successful peers as the implementation moved forward. These characteristics included resourcefulness, sustained success with recovery, and an ability to balance professionalism with empathy and approachability.
    - Parachute NYC staff also said that regular debriefings were useful because they ensured open communication and dialogue within the teams. Several respondents identified a need to define and clarify the roles and responsibilities of peer staff.

  - One staff member spoke of the early challenges in integrating peers in the mobile teams: "When the peers came on, we just kind of brought them to the meetings and then that was it. And then the peers would be quiet because they didn't really understand their role. But once we had a better understanding of their role, in getting—giving—providing the peer background knowledge about the case, that helps them to inform what they say sometimes at the meetings [themselves]. So that was helpful, but in the beginning, they kind of just came along for the ride and we didn't know where to fit them in."
  - Another staff member said, "The integration of—especially—more traditional medical model clinical staff with peer staff is going to take some time, and just be aware of that … quite frankly addressing that upfront would probably be best."

- **Conversion of mobile teams to the Parachute NYC model of care.** Mobile crisis teams existed before Parachute NYC and many staff members on these teams struggled to balance the Parachute NYC training and implementation with their existing work.
  - Leaders within several mobile teams discussed difficulties with planning the conversion to Parachute NYC, as well as the steep learning curve involved in the program's implementation. As one staff member described, "It took two years to get to the point where we were fully functioning … it was very difficult just mechanically, operationally, getting things started and getting the team as they were trained, as they joined and we had … some turnover just to kind of get things moving along."
  - The Parachute NYC model involves longer-term interactions than the traditional episodic care model that had been used by the mobile crisis teams. Some respondents noted that not all individuals will benefit from this longer-term care, and establishing appropriate referrals for those best suited to the NA-MCTs was a challenge in transitioning many respondents to the Parachute model. As one mobile team member stated: "I think for me at times, it's just difficult to get the appropriate referrals or the best referrals, and so that can be tough."
  - Implementing Parachute NYC was more complicated for CRCs managed by large provider organizations, due to their more complicated hiring and bureaucratic processes. Senior leaders reported working with some organizations that took more than six months to hire new staff members for Parachute NYC.

3. How did unexpected events support or conflict with successful implementation of the innovation? (RQ 8)

In October 2012, Hurricane Sandy caused disruptions in training schedules and resulted in a six-week delay in opening the citywide support line. Travel schedules, training locations, CRC site startup activities, and other logistics were significantly impacted by this event and the resulting citywide emergency response.

In addition, the program relied on foreign experts to provide initial training to staff members on the Parachute NYC model. Obtaining work visas and arranging travel for these trainers caused minor delays in the implementation of Parachute NYC.

C. Program effectiveness

At this time, we cannot provide findings on the four core measures, because very few FPHNY participants are enrolled in Medicaid or Medicare, and alternative data sources that would cover a large number of the participants are not available. We do, however, report on perceptions of the program's effectiveness from the perspectives of key awardee staff, members of the workforce, and other stakeholders.

1. From the perspective of key informants, to what extent does the program improve desired health outcomes and quality of life for participants? (RQ 23 and RQ29)

Parachute NYC staff in a variety of roles (including frontline staff, peers, and senior leaders) gave numerous examples of positive developments in the participants' ability to function in society. Some staff members discussed Parachute NYC participants who went back to work or school, and many participants were less isolated socially and improved their relationships with family and friends. One clinician noted:

> "The clients feel like they're being successful by working with us, working towards their goals and achieving their goals. Several of our clients, we see them less frequently now because they have jobs, they have their own apartments and things like that, so I think it's positive; and also a lot of clients that I've worked with … say [they] regard [me] almost as a friend or family member, which is really nice."

CRC and support line staff also saw improvements in the social skills of program participants. For example, one peer specialist discussed an experience with a support line caller:

> "A few months ago she was talking on the phone and … [staff] roles came up and she [said] you're my friend and I [responded], I'm not really your friend. I work on the support line, you call the support line; and we kind of talked about that … that was hard to hear but we talked about it and then after that she had this realization … I actually don't have friends but I could have friends."

Several peers and clinicians shared that participants often seemed more independent and "vibrant" after Parachute NYC. Staff members believed Parachute NYC gave its participants hope and empowered them to change their lives. For example:

- A peer noted, "We see people come in that don't think that it's possible anymore to have direction in their lives, and from interacting with us, they see that they can and they then actually want to."
- A 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."

In addition, several staff members described improvements in the participants' physical health. For example:

- Many staff members reported that they saw decreases in ED visits and hospitalization among program participants.
- 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.
- Several Parachute NYC staff also reported that participants were more capable of advocating for themselves about their medication. Parachute participants learned to talk to their doctors and care teams about lowering dosages to decrease the number and degree of side effects.

D. Context

1. To what extent did organizational features support or conflict with implementation? (RQ 104)

Because FPHNY is the fiscal agent of the Parachute NYC project, subcontracted providers delivered all the components of the program, including the crisis respite center and the mobile team services. This administrative arrangement resulted in a large and complex organizational structure that posed some challenges to implementation. For example:

- Program leaders reported that some larger provider organizations were often hard to work with because of the amount of time and number of requirements it took to hire staff.
- Some subcontractors were unfamiliar with deliverable-based contracting. As a result, contract negotiations were longer and more difficult.
- Various subcontractors also faced budget constraints; some initiated hiring freezes. This prevented some organizations from hiring the appropriate clinical staff for the mobile crisis team.

Additionally, staff reported that partnering with hospitals for referrals and ED diversions was more challenging than expected. For example, Parachute NYC staff reported that many hospitals believed they were the most appropriate providers to serve potential Parachute NYC participants. Staff also shared that some hospital staff thought CRCs offered a lower level of care than hospitals did.

Staff also reported that cultural changes in attitudes about the treatment of psychosis took time. Some of these staff specifically mentioned challenges faced by clinicians or supervisors on the mobile crisis team who were accustomed to more traditional approaches to treating mental health conditions and were asked to do things they regarded as the opposite of standard practice. For example, providers in traditional inpatient or outpatient facilities, in weekly appointments with their patients, monitored and insisted on compliance with medications. In contrast, the Parachute NYC model was based on the participant's preferences. The participant set the meeting schedule and decided what, if any, medications he or she wished to take.

Project leaders found that increased supervision from trainers and others staff—and providing a forum for clinicians to express their personal difficulties with the Parachute NYC model—helped clinicians deal with the dissonance between past practice and the NATM approach. In some instances, turnover also helped with this culture change. Some staff members said that one traditional psychiatrist left the program and was replaced with a psychiatrist whose beliefs were more aligned with and supportive of the NATM approach.

2. To what extent did the policy and political environment support or conflict with implementation? (RQ 106)

FPHNY leaders thought the city and state governments were very supportive of the Parachute NYC project. These key informants reported that government leaders were crucial partners who assisted them in removing many barriers to successful implementation. For example, the city of New York provided in-kind money to fund an external evaluation of Parachute NYC that was outside the award funding. Additionally, staff said the peer movement is strong in New York and the peer community is quite supportive of the project.

New York City's lack of affordable housing created barriers for frontline staff. Both peers and staff on the mobile crisis team reported many Parachute NYC participants were homeless or unstably housed, which often impeded recovery. Senior leaders at DOHMH echoed this finding. One staff member stated, "The biggest challenge, and probably the biggest complaint, I don't think has anything to do with the service, it has to do with the housing crisis in New York City [because we hear:] I don't like where I'm living or I don't have a super stable place to live and can't your Crisis Respite Center mobile team magically find me something that I can afford in two weeks?"

New York State's Medicaid redesign offers an opportunity to sustain the program. In 2015, the state changed its policy to provide reimbursement for Parachute NYC services, including peer support and CRCs. According to Parachute NYC staff, this change was originally slated to take effect at the beginning of 2015, but was delayed until October 2015.

E. Summary and conclusions

FPHNY was largely successful in implementing the Parachute NYC program, effectively launching the CRCs and mobile teams in four New York boroughs to reach the target population. Respondents observed that Parachute NYC is a completely new model in mental health services. Most of the major challenges faced by the program were related to its innovations and the corresponding lack of staff experience and systems of care to support the program. Challenges included integrating peer specialists into the program, transforming existing mobile crisis teams to the Parachute NYC model of care, partnering with appropriate referral sites, and managing the various subcontractors and administrative bureaucracies involved in the large-scale project. Over time, these challenges were largely addressed by the awardee staff and leaders.

Although we could not calculate the four core measures for this report because the necessary data were unavailable for most participants, the respondents we talked to believed the program had a positive impact on the treatment of serious mental illness, including reducing psychiatric hospitalization and ED visits.

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

VII. HEALTHLINKNOW, INC.

A. Introduction

HealthLinkNow, Inc. (HLN), a provider network organization, used HCIA funding to integrate telemedicine with a patient-centered medical home model. The goal of the project—to link mental health specialists, such as psychiatrists, therapists and counselors, more closely with primary care providers (PCPs) and thus improve participants' access to mental health services—advanced HLN's overall mission: to use a health IT platform to improve access to mental health services for enrollees of Medicare, Medicaid, and the Children's Health Insurance Program who live in rural areas. Within a single web-based system, the integrated health IT platform supported participant and provider communication, telemedicine, e-prescribing, practice management, scheduling, billing, and electronic health records. The project aimed to enroll 1,534 participants, who received services in 85 primary care clinics and hospitals in Montana, Wyoming, and Washington.

The program was designed to address a wide range of mental health and substance use disorders in children and adults. Participants were referred to the program by primary care physicians in Montana, Wyoming, and Washington. Each participant was assigned a care navigator. The care navigators interacted with the participant, the clinicians, and the primary care staff. They focused primarily on helping patients overcome barriers and ensuring that all those involved with an individual patient's care (including the PCP and therapist) were connected and communicating. All participants in Montana and Wyoming lived in frontier or rural areas that offered limited mental health services. In Washington, most participants also lived in rural areas.

The findings in this report are based on qualitative data collected by June 1, 2015, and enrollment data reported throughout the award period. Data sources included:

- Enrollment data submitted by HLN to the reporting website maintained by the Lewin Group for the HCIA initiative
- Qualitative data, including phone interviews and an in-person site visit in May 2015. Mathematica conducted in-depth interviews with key awardee staff, members of the workforce, and other stakeholders. In addition, Mathematica convened focus groups with site coordinators, primary care providers, and program participants. Data from the latter are not part of this report; instead, we provide an in-depth analysis of those focus groups in the seventh quarterly report.

We are not conducting impact analyses for HLN at this time due to small sample sizes and limitations in the availability of Medicaid data. HLN's program was designed for individuals in three separate states who were enrolled in Medicaid or Medicare. The program also targeted people of different ages with many different health conditions. Therefore, sample sizes are relatively small at the subpopulation level (such as enrollees in each state, elderly enrollees, or children). In addition, incomplete encounter data prevent us from analyzing Medicaid enrollees who are enrolled in managed care in Washington State. Finally, due to significant time lags in the availability of Medicaid data for all states, at the time when files were required to be available to allow us to conduct the analyses for this report (April 2015) Medicaid data for Montana, Washington, and Wyoming were only available through September 2013.

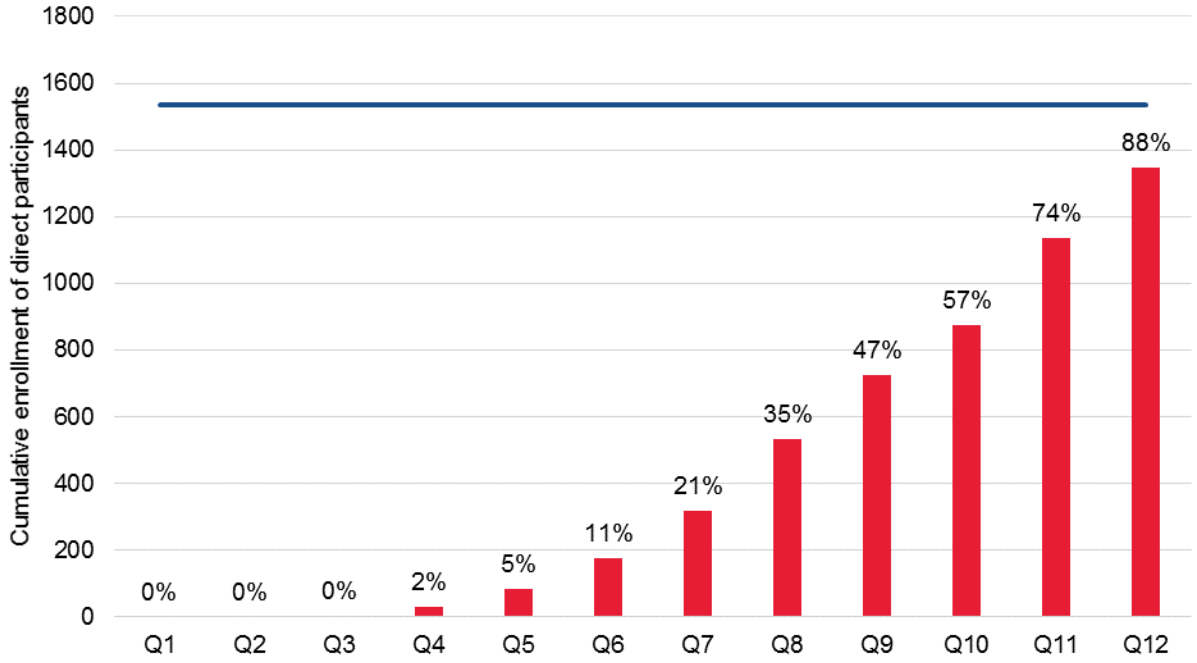
The research questions we address in this report were selected in consultation with our COR. Specifically, we examine questions related to implementation effectiveness in Section B, program effectiveness in Section C, and the program and policy context in Section D. We close with an overall summary and conclusion (Section E).[23]

B. Implementation effectiveness

1. Who was enrolled in the program?

By the end of the 12th quarter (June 30, 2015), the HLN program had provided services to 1,348 participants, achieving 88 percent of its enrollment target (Figure VII.1). In the first three quarters, HLN recruited providers and set up the project. HLN started enrolling participants in the third quarter, and enrollment levels rose steadily over the following months.

Figure VII.1. Percent of target enrollment achieved by quarter, Q1–Q12—HLN



Source: Quarterly reports submitted to the Lewin Group's reporting website for the HCIA initiative.

Note: As displayed by the horizontal bar, HLN's target enrollment was 1,534 unique participants.

[23] HLN may conduct analyses of data that they have collected on program participants and publish their own reports about program outcomes. These analyses are likely to use different data, focus on different outcome measures, and address different specific questions compared with the data and measures we use in this report to address questions of direct interest to CMMI. As a result, HLN may reach conclusions about program effects that differ substantially from the ones we present in this and future reports.

2. To what extent was implementation timely, conducted as planned, and responsive to site-level constraints? (RQ 17)

HLN was largely successful in implementing the program in a timely manner. However, several challenges delayed or hindered achievement of the program's goals:

- **Provider credentialing.** Providers had to be credentialed in order to provide HLN's health care services at a given facility. Most of the health care organizations HLN worked with credentialed providers themselves. This process varied based on the facility conducting it, but often included validating the provider's medical licenses, education, and insurance coverage. Providers were also required to submit a significant amount of paperwork. Respondents reported that obtaining credentials for HLN mental health specialists from PCP clinics was frequently a long, expensive, and complicated process, leading to delays in site recruitment and participant enrollment and, in some cases, affecting provider retention. For example:
  - Credentialing delays exceeded six months at many sites, and the provider sites could not enroll participants until their organization's HLN mental health specialist was credentialed.
  - Some respondents also linked these delays in credentialing to problems with retaining providers. One staff member said, "We also lose psychiatrists because you bring them on, train them and make all that investment, and then if the clinics do not process the paperwork, they can't start working. So they seek employment somewhere else."

  Delays in provider credentialing were a persistent issue for the following reasons:

  - Many of the clinics required the approval of their boards to work with HLN, and board meetings took place infrequently.
  - Credentialing practices and requirements varied significantly from one hospital to the next. Each facility required HLN providers to be credentialed according to its unique bylaws. As one respondent commented,

    > "The biggest challenge, I think, is dealing with the different hospitals and insurance companies, as there's not a set standard on things that are required, so it is kind of hit-or-miss. Some places will give me a list of things they need, other places don't … I send the application and then they'll come back with 50 emails saying you need x, y, and z."

  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, issues with the credentialing process were cited as the most significant barrier to implementation.

- **Recruitment of PCP and hospital sites.** The program's recruitment of PCP and hospital sites proceeded slowly and yielded fewer referrals than expected. Initially, HLN's enrollment strategy involved approaching clinics and hospitals in rural areas on an individual basis. Although this approach was appropriate for the target communities, in

which face-to-face contact was key to building business relationships, it was also time-consuming and not consistently successful. HLN staff often had to schedule multiple in-person meetings to develop the site staff's trust in the program and to recruit potential sites. This effort, coupled with the size and topography of the states (especially in Montana and Wyoming), produced delays in site recruitment.

- **Referrals.** Respondents also reported that the sites recruited by HLN did not initially refer as many participants as they were expected to. In response, HLN expanded the program's reach to Washington State and worked to recruit larger health systems. HLN also began to receive more word-of-mouth referrals as its services became better known. HLN enrolled more participants as a result of these efforts, but several respondents said it would have been a good idea to dedicate more resources to business development and marketing.
- **Billing.** HLN staff emphasized that two major challenges with billing were the lack of insurance coverage for tele-psychiatry and the lack of a dedicated staff member to handle billing issues. For example:
  - Budget constraints prevented the project from hiring a project manager, so billing was primarily handled by care navigators, who were also responsible for supporting participants, acting as liaisons between HLN providers and PCPs, and credentialing.
  - These care navigators described spending significant amounts of time in contract negotiations with insurance companies, asking the companies to cover participants' services.
  - Care navigators also reported that they sometimes found themselves educating payers about tele-psychiatry and why it was important to cover these services for their beneficiaries.
  - While respondents noted that insurance companies were becoming more aware of and knowledgeable about tele-psychiatry, some of the companies still did not reimburse HLN for these services in all cases. In these situations, PCP sites sometimes reimbursed HLN for the tele-psychiatry sessions and sometimes billed the participant directly. In rare cases, HLN was able to obtain coverage for an individual's tele-psychiatry by instituting a single case agreement with the insurance company.
  - HLN tried to minimize these billing issues by maintaining a tracking spreadsheet of health plans that did offer telemedicine as a covered benefit. However, in some cases, exceptions in these plans prevented HLN from being paid for the service. For instance, one health plan declined to reimburse out-of-state providers, even when the participant lived in state. To receive payment from that health plan, HLN had to contract with the plan in the HLN provider's state of residence.
- **Technical issues.** Many respondents reported that technical issues, including network and Internet problems, interfered with providing the HLN services. The underlying reasons for these issues were unclear, although several respondents speculated they may have been due to deficiencies in the PCP sites' technical infrastructure. Respondents indicated that HLN care navigators and IT staff were responsive in helping the sites quickly resolve any technical issues that arose. Some reported that the technology issues initially occurred frequently, but lessened over time as technical staff fine-tuned the systems. A few

respondents, however, said they were still experiencing problems, which were sometimes severe enough to require rescheduling sessions with participants.

- **Staffing challenges.** Respondents identified several key challenges with hiring and retaining the program's mental health specialists. For example:
  - Abundant demand for mental health providers nationwide caused some HLN clinicians to take higher-paying positions with other organizations. This, in addition to the credentialing challenges, was a cause of clinician turnover.
  - HLN staff recommended that, if affordable, the program should recruit a full-time, on-site psychiatrist to formally supervise the other mental health specialists on staff, who are all participating remotely.
  - Turnover in the program's mental health specialists affected PCP staff and participants. Several PCP staff reported that it would have been helpful to have the same mental health specialist see participants over time:
    - One respondent said, "We've been through quite a few [specialists]. And that's been a little bit frustrating, but we live in the real world that you kind of form a … relationship … from the clinical part, we'd go in with a certain amount of knowledge about what they wanted in the previous meetings, about what information [was needed and] how well they know the patient."
    - Another respondent agreed, saying that it can be challenging for participants and staff at the primary care site to "always have to start at Square One" when their psychiatrist changed.
    - Several respondents at provider sites also articulated the need for more flexible and numerous appointment times.

3. How did unexpected events support or conflict with successful implementation of the innovation? (RQ 8)

HLN staff and partners reported no unexpected events.

C. Program effectiveness

We do not provide findings on CMMI's four core measures, due to limitations in data availability for HLN. We do, however, report on perceptions of the program's effectiveness from the perspectives of key awardee staff, members of the workforce, and other stakeholders.

1. From the perspective of key informants, to what extent does the program improve desired health outcomes and quality of life for participants? (RQ 23 and RQ29)

Respondents believed the program had a positive impact on its participants' health and well-being, in large part because it targeted areas where access to mental health services was low. As one PCP staff member explained, "Because we are in a very rural area, [there are] no psychiatric services 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, 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 participants would have had to go without mental health services if they had not been able to enroll in HLN.

Respondents reported that participants responded positively to HLN's services, and their sense of mental well-being and quality of life were enhanced. PCP staff described seeing improvements in participants' health, including fewer symptoms of anxiety and depression and more positive behavior changes. Participants also 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.

Respondents shared stories of the specific improvements in participants' physical health. For example, one HLN staff member talked about a participant who lost a significant amount of weight and started exercising. Another staff member described a young participant who was having dietary issues, and who began to eat and gain weight after starting therapy.

In addition, several respondents said no-show rates were lower and adherence to treatment was greater in the HLN program compared with traditional mental health services. They believed this difference was due to the added convenience of locating the HLN appointments at the PCP's office. One respondent said, "It seems that there's more follow-through with the patients when they come to a familiar area." Speaking to the accessibility of HLN services, one PCP staff member said,

> "They come to our clinic. They come to the same room. They see the same nurses. They don't have to drive to the bigger city areas. They don't have to find parking … it's not a new situation. And so that's comforting to them. I think it does remove stigma because everyone knows you have to go to your primary care office … it takes away the stigma of walking into a psychiatrist's office [on the] first day."

Most respondents also praised the high quality of the HLN providers, saying the participants easily developed rapport and trust with their providers. Many respondents were initially skeptical about the effectiveness of delivering remote psychiatric services, but observed that participants responded positively to HLN's format and were comfortable using the technology to interact with their provider.

Respondents considered the care navigator instrumental in the improvements they observed in participants' health outcomes and quality of life. Care navigators focused on helping participants overcome barriers to using the service, and assisted with such tasks as scheduling appointments and ensuring that everyone involved with an individual participant's care (including the PCP and therapist) was connected and communicating. Feedback from some respondents indicated the care navigators might also have contributed to participant well-being by reinforcing positive health behavior and treatment adherence. One respondent said, "The care navigators have a huge influence on keeping people on the treatment plan and getting the

medication." Another respondent spoke about a care navigator who talked to a suicidal participant on the phone until help could arrive.

D. Context

1. To what extent did organizational features support or conflict with implementation? (RQ 104)

Respondents cited several organizational features that contributed to a successful implementation of HLN. For example:

- PCP sites without the capacity to provide mental health services themselves seemed particularly motivated to implement HLN and to appreciate the value of the service.
- HLN staff also reported that having the support of both clinical and administrative leaders was important. 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 came up. One respondent said, "There's some cases where even once [HLN] is 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."
- Although setting up the HLN software did not require substantial technical ability, organizations with complex IT systems, including numerous layers of security handled by different staff, found the process challenging. Respondents reported that such organizations needed a significant amount of time to locate the staff responsible for each component and to obtain the necessary permissions to install the software.
- HLN was recently accredited by the American Telemedicine Association to provide online patient consultation, and they believe this will help their recruitment and expansion efforts and provide more visibility for their services.

2. To what extent did the policy and political environment support or conflict with implementation? (RQ 106)

As pioneers in the delivery of tele-psychiatry services in rural areas, HLN leaders and staff made significant efforts to keep up to date on local and national legislation and regulations about telemedicine. Respondents identified several aspects of the policy and political environment that affected implementation. For example:

- The lack of standardized regulations and guidelines for hospital and clinic credentialing significantly hampered implementation. One respondent said, "If there [were] some sort of guideline or standard on what is expected, then it [would help] us be better prepared. We can advise our providers up front on exactly what we need and there's not that back and forth." Another respondent suggested that having a national credentialing body would address this issue.
- HLN staff cited CMS's proxy credentialing procedure, which can significantly streamline the credentialing process for hospitals and clinics that are part of a network, as favorable to

program implementation. However, respondents added that staff at some facilities were unaware of this policy and believed they still had to go through their own credentialing processes. This caused delays in implementation.

- Recently, the three states in which HLN implemented its program passed legislation in support of telemedicine. For example:
  - In Washington, a new provision states that emergency departments are not allowed to hold psychiatric patients overnight without a consult from a psychiatrist. Respondents believed this has resulted in increased demand for mental health coverage, including tele-psychiatry services like HLN's, in the state.
  - Parity regulations in Montana require commercial payers to cover telemedicine and to provide reimbursement for these services at a level equal to that provided for in-person services.
  - Respondents noted that even though some specific state regulations were supportive of tele-psychiatry and favorable to the growth of the HLN model, there were often gray areas and exceptions that kept HLN from being reimbursed for its services.

E. Summary and conclusions

HLN implemented its program relatively smoothly, working effectively to incorporate the technology at PCP sites and provide the innovation components to participants. Yet the awardee did experience several challenges that delayed the progress of implementation, including provider credentialing issues, PCP recruitment rates that were lower than expected, reimbursement issues, and staffing challenges. HLN worked throughout the award period to address these challenges and adapt its implementation strategies as needed; for example, modifying its recruitment approach to increase PCP referral rates and adapting its internal policies and staff training procedures to streamline the billing and credentialing process. Although we could not calculate the four core measures due to limitations in data availability, we learned that key respondents perceived HLN was largely successful in its goal to improve access to psychiatry services in rural areas. These key staff, stakeholders, and workforce members emphasized the program's positive impact on participants' well-being and quality of life.

VIII. INSTITUTE FOR CLINICAL SYSTEMS IMPROVEMENT

A. Introduction

The Institute for Clinical Systems Improvement (ICSI), a regional health improvement collaborative in Minnesota, implemented a collaborative care management model called Care of Mental, Physical and Substance Use Syndromes (COMPASS) with seven partners in California, Colorado, Florida, Massachusetts, Michigan, Pennsylvania, and Washington. ICSI and its partners, known collectively as the COMPASS consortium, tested national dissemination of the model. This project aimed to improve care and lower costs for 8,000 adults (some of them direct participants and some indirect[24]) enrolled in Medicare or Medicaid who, in addition to active depression, had uncontrolled diabetes, uncontrolled cardiovascular disease, or both. Each participant had a personalized care plan and a care team that incorporated care managers and consulting specialists and regularly reviewed the participant's progress and outcomes (systematic case reviews). In this model, care managers not only function as a liaison between the care team and the participant, but also work to reduce barriers to care—addressing social or physical needs and providing health education.

ICSI's partners included health plans, independent physician practice groups, large integrated health care systems, a federally qualified clinic, and regional health care collaboratives. These partners varied substantially in their location and size, the characteristics of their patient populations, and their experience with COMPASS. Most clinical partners implemented the program at multiple sites. In addition, as a provider network, ICSI itself had four clinical settings participating in the consortium. Overall, in its final report to CMMI, ICSI noted that 171 clinical sites participated in the program. ICSI also had several technical partners who assisted in implementing and evaluating the intervention. These were the Advancing Innovative Mental Health Center (AIMS Center) and the HealthPartners Institute for Education and Research (HPIER).

Each COMPASS program site targeted Medicare and Medicaid enrollees who had active depression with uncontrolled diabetes, cardiovascular disease, or both. An enrollee with a score greater than 9 on the widely used Patient Health Questionnaire-9 (PHQ-9) was considered to have active depression. ICSI's target for enrolling direct participants was 2,700 adults who had active depression combined with uncontrolled cardiovascular disease or diabetes. By the end of the project (quarter 12), ICSI reached this target.

The findings in this report are based on quantitative data received by December 19, 2014 and qualitative data collected before June 1, 2015, as well as enrollment data reported throughout the award period. Data sources include:

- Enrollment data submitted by ICSI to the reporting website maintained by the Lewin Group for the HCIA initiative.

[24] Direct participants received care from a clinic or care manager/care coordinator funded by HCIA dollars. Indirect participants received COMPASS care even though the clinic or care manager/care coordinator was not funded by HCIA dollars.

- Medicare and Medicaid IDs submitted to Mathematica by COMPASS consortium members through December 19, 2014. By that date, most of ICSI's partners had submitted files with participant Medicare and/or Medicaid identification numbers, which we used to extract data on Medicare program enrollment, utilization, and expenditures. In the future, we may also analyze Medicaid enrollment and claims data, as well as electronic health records data (EHR) that provide information on program services and participant health outcomes.
- Qualitative data, including telephone interviews and an in-person site visit in April–May 2015. Mathematica conducted in-depth interviews with key awardee staff, members of the workforce, and other stakeholders.

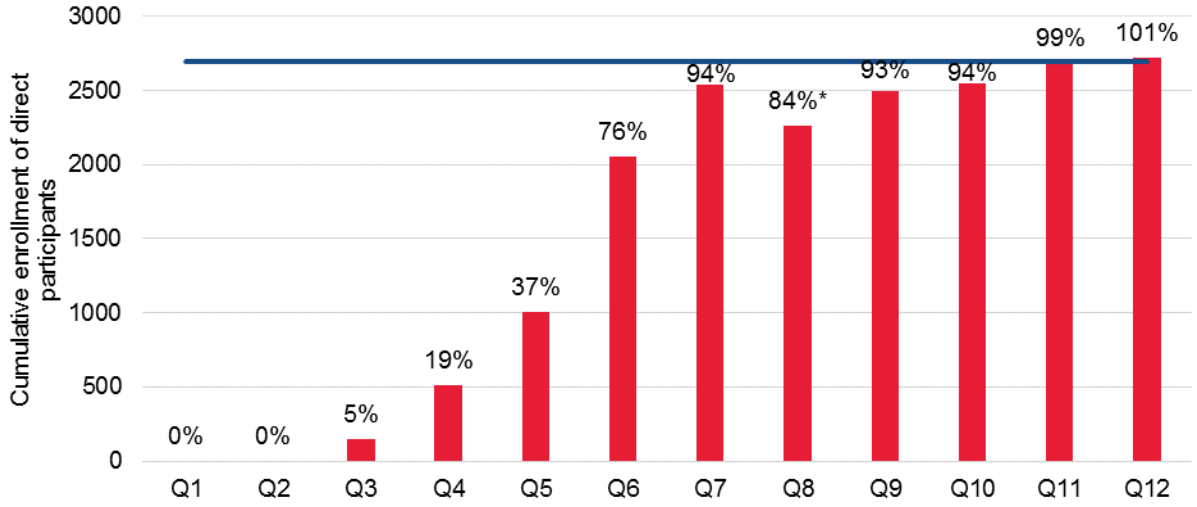
The questions we address in this report were selected in consultation with our COR. Specifically, we examine questions related to implementation effectiveness in Section B, program effectiveness in Section C, and the program and policy context in Section D. We close with an overall summary and conclusion (Section E).

B. Implementation effectiveness

1. Who was enrolled in the program?

By the end of the 12th quarter (June 30, 2015), ICSI had enrolled 2,726 direct participants, exceeding its original enrollment target of 2,700 (Figure VIII.1). In the first two quarters, ICSI was initiating and piloting the project. Enrollment began in the third quarter with 148 direct participants, increasing 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 VIII.1. Percent of target enrollment achieved by quarter, Q1–Q12—ICSI



Source: Quarterly reports submitted to the Lewin Group's reporting website for the HCIA initiative.
Note: As displayed by the horizontal bar, ICSI's target enrollment was 2,700 unique participants.
*In Q8, some participants left ICSI's COMPASS model.

2. To what extent was implementation timely, conducted as planned, and responsive to site-level constraints? (RQ 17)

ICSI was moderately successful in implementing COMPASS in a timely manner across its partner sites. However, several challenges affected ICSI's implementation of the program:

- **Complex legal agreements:** A significant barrier to implementation was the complexity of cross-organizational legal agreements—such as partner and medical group subcontracts, business associates' agreements for common use of the COMPASS registry's EHR data for participant information, and licensing of the registry at the participating sites—and the time required to obtain Institutional Review Board (IRB) approvals from multiple organizations. Fifty-five separate legal agreements and 12 IRB applications were required. In some cases, the team needed four months to obtain IRB approvals.
- **Slow enrollment:** Delays in implementation as well as challenges with outreach and recruitment strategies contributed to slow enrollment rates for many sites. To improve enrollment rates, the COMPASS consortium developed various strategies, such as asking partners to add clinical sites, increasing marketing efforts, awarding grants to sites for innovation in patient enrollment, extending enrollment incentives for ICSI partners and their sites, and broadening eligibility criteria to include participants without Medicaid or Medicare coverage, as well as participants over 65 years of age with uncontrolled hypertension and depression. The consortium found that broadening the program's eligibility criteria was the most effective way to increase enrollment.
- **Longer startup times:** Some partners reported that program startup took longer than anticipated because of the program's effect on clinical workflows and staff roles. These partners expected COMPASS to be up and running efficiently within several months, but staff reported they quickly learned COMPASS implementation would affect the entire clinical site. For example:
  - As one site coordinator described it, COMPASS affected "nursing administration and the IT people and the front desk and … the pharmacy and anyone you can think of."
  - Another coordinator reported staff took a long time to incorporate the program's screening procedures and learn its eligibility rules.
  - Even partners that had implemented similar programs before took more time then expected to fully launch COMPASS. These partners said that at least six months were needed for site staff to adapt to the change in staff roles and workflow and put all the necessary resources in place to fully implement the program.
- **Leadership and physician buy-in:** Strong buy-in from leaders and physicians was necessary for successful implementation of the COMPASS program because its innovative approach to providing services called for a cultural shift at the implementing sites. Some partners, however, did not have the support of their organization's senior leaders or physicians. One member of the ICSI leadership team said, "It's a real cultural change that has to occur and so I guess that does come down to leadership. [You've] got to … make sure that people are really on board." Some staff at partner sites shared their perception that their organization's leaders or physicians viewed COMPASS as a threat to their revenue stream. In particular, staff of one independent practice association said their leaders thought

participating in COMPASS could result in fewer patient visits to the practices, meaning less revenue for individual practices. ICSI leaders addressed this challenge by having ongoing conversations with leaders at partner organizations and continuously supporting the staff implementing COMPASS at its partner sites. In some instances, staff turnover resolved some of the issues with leadership or physician buy-in because new leaders who supported COMPASS came on board when less supportive leaders left.

- **Participant engagement:** Care managers struggled to keep participants engaged in the program. Many reported the biggest barrier to implementing COMPASS was retaining participants after they enrolled. Several care managers noted that most participants had problems securing stable housing, accessing transportation, and keeping their finances in order; these issues often took priority over their health care. Other care managers said some participants had no access to a reliable phone line, which led to problems when care managers tried to reach the participants. One care manager said,

  > "We've struggled with it. That the health system has never had social worker support out in the clinic [setting] and in primary care … if we are not able to address some of those social issues, we can have all the great programs put into place, but that patient doesn't have any way of achieving [their goals]. If they can't afford their insulin and if they're eating … from the food shelf and if they don't have transportation to the clinic, they can't follow the medical plan."

  Care managers thought COMPASS gave them the leverage to help address these concerns about social services. However, some respondents also thought participants did not have enough access to appropriate community resources.

- **Registry transition:** Some staff reported challenges integrating their EHR data with the COMPASS registry. These staff members described how the COMPASS registry was not compatible with their EHR and did not allow for information exchange between the two systems. This led to inefficiencies of double entry and, in some instances, triple entry in the organizations' various health IT systems.

3. How did unexpected events support or conflict with successful implementation of the innovation? (RQ 8)

ICSI staff reported that no unexpected events affected the program's implementation.

C. Program effectiveness

In this section, we present the results of a pre-post analysis on three core outcome measures for Medicare FFS beneficiaries participating in the COMPASS program: total expenditures, hospitalizations, and emergency department (ED) visits. (We did not conduct an analysis of hospital readmissions because there were too few readmissions in the data). First, we briefly review our analytic methods. We then report the findings and limitations of the analysis. Next, we include qualitative findings that address specific research questions on the program's effectiveness.

1. Overview of quantitative methods and population

Our pre-post analysis was based on data for program participants who met the following criteria in the baseline period—that is, the year before they enrolled in COMPASS: (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[25]). The residence criterion was based on the partners we selected for inclusion in this analysis; specifically, we selected four ICSI partners (with a total of 18 different clinic sites) based on the relatively large number of Medicare FFS participants served by these partners. Our initial analytic sample included those participants enrolled during the first 11 months of the program (February 2013 through December 2013). The baseline period (that is, the "pre-period") included data for 12 months prior to enrollment. The follow-up period (that is, the "post-period") included data for participants beginning at their enrollment through September 2014, thus allowing at least nine months of post-data for each identified participant.

We conducted a regression analysis between the baseline and one year follow-up to obtain regression-adjusted means for the treatment group. For each outcome, we estimated one model with a full set of control variables, including age (linear), gender, race (white, black, and other), original reason for entitlement (aged vs non-aged), dual eligibility status, indicator for whether 12 months of baseline data were available, and HCC score. This model also included indicators for the presence of diagnoses of various chronic conditions (depression, acute myocardial infarction, atrial fibrillation, congestive heart failure, hypertension, ischemic heart disease, hyperlipidemia, stroke, diabetes, hypothyroid, Alzheimer's disease, anemia, asthma, hyperplasia, cataract, chronic kidney disease, chronic obstructive pulmonary disease, glaucoma, hip/pelvis fracture, osteoporosis, arthritis, breast cancer, colon cancer, prostate cancer, lung cancer, and endometrial cancer).[26]

2. Quantitative findings on program effectiveness

In COMPASS's first year at the selected sites included in these analyses, we found statistical trends suggesting that participants had higher Medicare expenditures and higher hospitalization rates in the follow-up period compared with the baseline period (Table VIII.1). Specifically, for the baseline period, regression adjusted mean expenditures were $15,655 compared with $22,053 for the follow-up period. Hospitalizations increased from 0.8 to 1.2 per patient. Findings for both outcomes are statistically significant ($p<.05$). Rates of ED visits did not change from one period to the other. We consider these results very preliminary and subject to change as we conduct analyses over longer time frames and as additional participants reach the 12-month follow-up point and are added to the analyses.

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

[26] We also explored a more parsimonious model with fewer control variables. This model yielded somewhat different results. We report only the results for the full model because additional tests suggest that it is preferable to the model with fewer control variables. However, readers should consider our findings preliminary. We will further examine this issue of model stability when we have additional data next year.

Table VIII.1. Pre- and Post-Intervention Means for ICSI, annual results

	Regression-adjusted mean for intervention group	Differences		
		Value	Percent	P-value
Total expenditures per patient				
Baseline	$15,655	N.A.	N.A.	N.A.
IY1	$22,053	$6,339	40	<0.01
Hospitalizations per patient				
Baseline	0.8	N.A.	N.A.	N.A.
IY1	1.2	0.5	62	0.03
Readmissions per patient[a]				
Baseline	N.A.	N.A.	N.A.	N.A.
IY1	N.A.	N.A.	N.A.	N.A.
ED visits per patient				
Baseline	1.7	N.A.	N.A.	N.A.
IY1	1.9	0.2	9	0.49
Number of observations	526	N.A.	N.A.	N.A.
Number of unique patients	263	N.A.	N.A.	N.A.

Source: Mathematica analysis of FFS Medicare claims data for the period from February 2012–September 2014.

Notes: To be included in the analysis, individuals had to be enrolled in Medicare FFS Parts A and B and have Medicare as the primary payer for at least six months in the baseline period; have a diagnosis of either depression, cardiovascular disease, or diabetes; and live in one of the four states operating the ICSI program during the baseline period.

[a] Readmission rates were not analyzed due to small sample size and too few events.

[b] We derived the impact estimates in Stata using the lincom command.

N.A. = Not applicable.

3. Limitations of the quantitative analysis

One of the most important limitations of our analysis 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 not enrolled in the COMPASS program. For example, costs might have increased even further for these individuals in the absence of the COMPASS program.

We made considerable effort to construct a comparison group that would provide adequate counterfactual evidence. However, many of ICSI's enrollment criteria are based on data not observable in Medicare claims and 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 PHQ-9 and many of the clinical indices used to enroll program participants. These measures of depression and clinical status are not included in the Medicare data files. 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.

As part of our efforts to address this problem, we used diagnostic information reported in claims (for example, the ICD-9 codes) as a proxy for the PHQ-9 and clinical indices used to select ICSI participants. However, this approach had serious flaws. In particular, the diagnostic information on depression was often absent. For example, we knew that all ICSI participants had severe depression at the time of enrollment, as determined by their PHQ-9 scores. In theory, all of them should have had a diagnosis of depression in their claim records. However, we found that only 65 percent of participants had such a diagnosis reported in the claims files in the 12 months prior to enrollment.

Thus, we were faced with the task of finding individuals for the comparison group who were depressed but had no diagnosis of depression in their claims records. We examined various strategies for finding such individuals—such as looking for patterns of service use that would signal severe depression—but none were satisfactory. That is, none led us to a comparison group that would provide reliable and valid counterfactual evidence. Although our intervention and comparison groups might have looked well-matched on the variables that were in the records, they actually could have been quite different on other important variables (such as the extent of severe depression) that we could not observe but that are strongly related to costs and outcomes.

Moreover, we could not determine the direction and magnitude of the potential bias resulting from these potential differences. For example, if the participant group was more depressed or ill than the comparison group, our results might have been biased toward finding a negative effect of the intervention. That is, we might have found that, relative to the participant group, the comparison group did better. The cause of this difference might be because the members of the comparison group were healthier to begin with, not because the intervention didn't work. Alternatively, we might have found that the participant group did better; however, someone could argue that the improvement was simply the result of a regression to the mean that did not affect the comparison group, given that the comparison group was not as severely ill. Eventually, and in conjunction with our COR, we determined that developing a comparison group was not feasible. Consequently, we decided to conduct a pre- and post-intervention analysis with the treatment group only.

We note three additional limitations:

- Our results reflect the first year of services provided after an individual enrolls in the program. It is possible that the program may initially increase Medicare utilization and spending as providers and care managers ensure that participants obtain needed services, and that reductions in utilization and cost savings may appear in later periods. Consequently, we may find different results in the second post-program year, and those will be presented in the third annual report.
- Our current analysis is limited to Medicare FFS enrollees and Medicare-covered services at four ICSI partners (Table VIII.2). These four partners are not representative of all the states and partners implementing the COMPASS program. Our findings may not be generalizable to other ICSI participants (for example, those covered by Medicare Advantage, Medicaid only, or commercial insurance) at these sites or to Medicare FFS patients treated by other ICSI partners.

- The services and costs reported may substantially underestimate actual service utilization and costs for this population, particularly for any participants who were dual Medicare-Medicaid beneficiaries, because they do not include services paid for by Medicaid or other health insurers, or services not covered through any insurance (for example, psychiatric hospitalizations that exceed allowable limits).

In light of these limitations and the preliminary nature of our findings to date, it is important to avoid drawing premature conclusions about the impact of ICSI's program.

Table VIII.2. ICSI partners and clinics included in analysis

ICSI partner	Location	Clinics
Mayo Clinic	Minnesota	Austin, Albert Lea, Owatonna/Faribault, Mankato, Red Wing, Rochester, and Jacksonville
Michigan Center for Clinical Systems Improvement (MiCCSI)	Michigan	Advantage Health, Lakeshore, Spectrum
Pittsburg Regional Health Initiative (PRHI)	Pennsylvania	Excela, Premier, St. Vincent's Medical Group
Institute for Clinical Systems Improvement (ICSI)	Minnesota	Entira, Essentia, Lakeview/Stillwater, North Memorial

Source: Information provided by ICSI.

4. Qualitative findings on program effectiveness: From the perspective of key informants, to what extent does the program improve desired health outcomes for participants and quality of life for participants? (RQ 23 and RQ29)

Program staff, particularly care managers, mentioned several ways COMPASS may have improved its participants' health outcomes:[27]

- Investment in care and care competence. Most staff members reported that participants were more interested and invested in their own care and gained knowledge about their health. For example, staff reported COMPASS participants:
  - Improved their compliance with care plans
  - Increased self-care activities, including improving their diet and working to get enough sleep
  - Increased their self-monitoring of blood pressure and glucose 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 provider], although most of the engagement has been with the care managers, and the care managers let the primary [care provider] know what's going on, but [participants have] become more engaged and compliant with testing and laboratory work [and] with medication adherence."

[27] In their final report to CMMI, ICSI provides program outcome information based on their own evaluation.

- **Social outcomes.** Many staff members noted that participants' social functioning improved. For many participants, this included starting or returning to work. Some physicians and care managers reported participants were less socially isolated, spending more time with family and friends.
- **Physical health.** Anecdotally, care managers and physicians noted some changes in the physical health of participants. These included decreases in A1C levels (a marker of diabetes), improved blood pressure, and lowered cholesterol. A few staff members noted a reduction in some participants' weight and body fat. 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 their depression improved. One staff member said one of her participant's PHQ-9 scores decreased markedly while she was involved in COMPASS. This respondent stated, "[The patient's] PHQ-9 score …dropped about 10 points over that six- to eight-month period."

Staff were careful to note that not all COMPASS participants experienced these improvements, adding that improvements can take a lot of 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 the "roller coaster" metaphor of care, in which 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 improvement in some of the domains. One care manager described the roller coaster in the following way:

> "And we have several [COMPASS] patients … 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 patient 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 patient is going to do okay because they're finally getting it."

D. Context

1. To what extent did organizational features support or conflict with implementation? (RQ 104)

Several organizational features supported COMPASS implementation. These included:

- **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 COMPASS 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 have social workers in their clinical sites; the transition to COMPASS 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 their existing partnerships—including relationships with local YMCAs, community colleges, and fire departments—to better implement COMPASS in their communities. For example, one site included a social work intern from a local university in its team to help the RN care coordinators meet participants' needs for social services.
- **Collaborative approach.** The structure of the COMPASS project—a lead coordinating organization (ICSI) and several implementation sites— posed administrative and legal challenges, as described. However, ICSI worked to ensure that partner sites viewed COMPASS as a collaborative effort rather than the work of one organization. This approach was especially important because of the diverse organizations involved in COMPASS; staff reported that the collaborative approach fostered partner buy-in, transparency, accountability, and the sharing of experiences and lessons learned throughout the consortium.
- **Supportive leaders.** Staff reported that the support of both administrative and clinical leaders was important to the successful implementation of COMPASS. In particular, enthusiastic and influential physician champions and committed mid-level staff were critical. As one 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 work flows 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 parties 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."

- **Balanced composition of the Systematic Case Review (SCR) team.** Respondents also emphasized the importance of having the right mix of individuals on the SCR team, selecting appropriate physician consultants, and including 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 SCR—one who could champion the project—was useful for gaining and maintaining buy-in within the site.
- **Effective communication strategies.** Organizations that developed effective ways of communicating SCR recommendations 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. To what extent did the policy and political environment support or conflict with implementation? (RQ 106)

Staff at ICSI and its partners frequently brought up the FFS payment model as the biggest barrier to implementation. Many of them noted that the traditional FFS model did not have payment systems for COMPASS services, particularly for care management time, and 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. As 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 planned to train the partner organizations in the total cost of care model in the beginning of June 2015.

E. Summary and conclusions

ICSI met its enrollment target and was largely successful in implementing the COMPASS program throughout its partner organizations. Many staff members believed COMPASS was effective for its participants. Partner organizations faced some difficulty in adapting their systems to implement COMPASS, and this presented challenges to implementation. ICSI leaders addressed these challenges by actively supporting their partners through the implementation process and responding effectively to their issues and concerns.

Our impact analyses provide very preliminary indications that the COMPASS program increased total expenditures and hospitalizations. To provide more definitive results, we will continue our analyses with additional data in the upcoming year.

This page is left blank for double-sided copying.

IX. KITSAP MENTAL HEALTH SERVICES

A. Introduction

Staff at Kitsap Mental Health Services (KMHS), 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 for adults and children receiving outpatient services at KMHS, thereby reducing the cost of their care. Specifically, the program aimed to (1) integrate treatment of mental health and substance use disorders with physical health monitoring, prevention, and intervention; (2) improve the connection between individuals with serious mental illness or severe emotional disturbance and their primary care providers (PCPs); and (3) support appropriate services to address mental health and substance use disorders in physical health care settings.

KMHS' program was agency-wide and designed to improve the integration and coordination of care for *all* who use KMHS outpatient services. For our evaluation, and to comport with KMHS' implementation procedures, we define the target population as all KMHS patients served after January 1, 2013.[28]

Although all KMHS patients were included in the target population, KMHS periodically identified subgroups of patients with more severe health conditions.[29] These were known as cohorts. Staff made a special effort to ensure that electronic health records (EHRs) for members of the cohorts contained key data about their health status and use of health services. The staff obtained this information by reaching out to patients' PCPs and tapping into the state's emergency department (ED) data system. KMHS used these more comprehensive EHR data to better understand the full range of patients' needs for health services and to improve coordination with PCPs.

The findings in this report are based on quantitative and qualitative data collected or received by February 1, 2015, and June 1, 2015, respectively, and enrollment data reported throughout the award period. Data sources include:

- Enrollment data submitted by KMHS to the reporting website maintained by the Lewin Group for the HCIA initiative.

[28] KMHS received its award in July 2012 and used the first six months of the award period to prepare and train staff, then gradually implement components of the program. KMHS fully implemented the program by January 1, 2013.

[29] 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.

- Medicare administrative data covering the period from January 2009 through September 2014.[30]
- Qualitative data, including telephone interviews and an in-person site visit in April 2015. Mathematica conducted in-depth interviews with key awardee staff, members of the workforce, and other stakeholders. In addition, Mathematica convened focus groups with members of the workforce and with individuals who receive outpatient services from KMHS. Data from the latter are not a part of this report; instead, we provide an in-depth analysis of the focus groups with individuals who received KMHS services in our seventh quarterly report.

In this chapter, we also present the results of our quantitative analyses on four core outcomes. The analysis of program impacts includes KMHS patients enrolled in fee-for-service (FFS) Medicare—a group that represents about 15 percent of KMHS patients. Trends in outcomes that are not evident now may become apparent as we conduct analyses of Medicaid data in the future. In light of the preliminary nature of our findings, it is important to avoid drawing conclusions about the effectiveness of KMHS' program based on these early analyses.

The questions we address in this report were selected in consultation with our COR. Specifically, we examine questions related to implementation effectiveness in Section B, program effectiveness in Section C, and the program and policy context in Section D. We close with an overall summary and conclusion (Section E).[31]

B. Implementation effectiveness

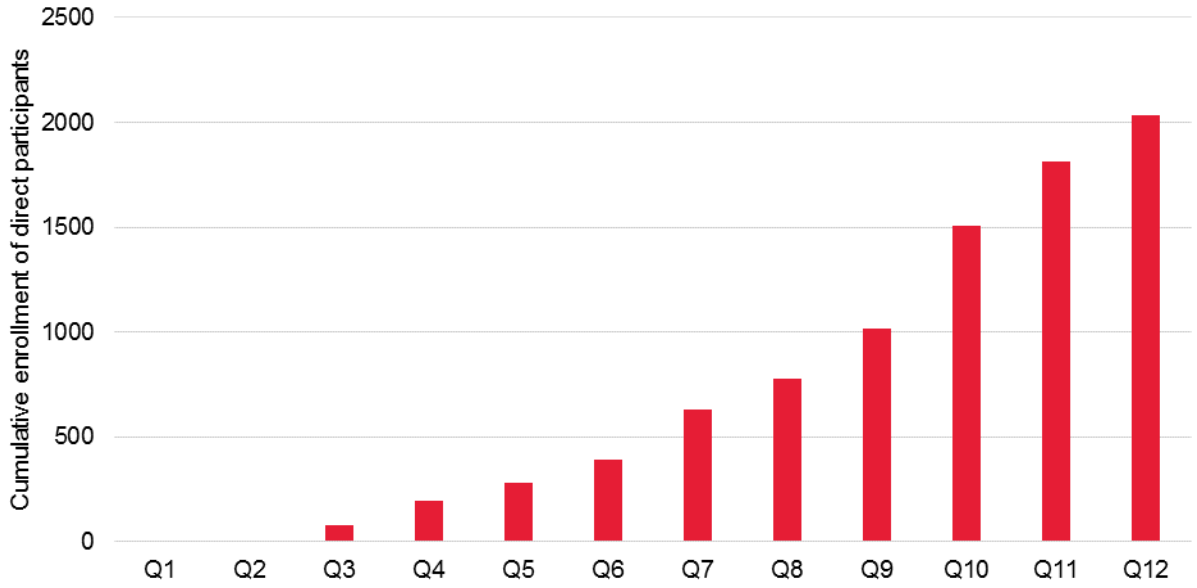
1. Who was enrolled in the program?

KMHS did not define a specific enrollment target for *Race to Health!* because KMHS staff intended the program to reach everyone who used KMHS' outpatient services. In its quarterly reporting to CMS, KMHS defines "direct participants" more narrowly as those whom its PCP partner refers for behavioral health consultation, or those who receive direct services from either the HCIA-funded medical assistants or the HCIA-funded Healthy Families Coordinator. KMHS began providing services to direct participants in Q3, and the number of participants it served steadily increased over the course of the award. By the end of the project's 12th quarter (June 30, 2015), KMHS reported that it had served 2,033 direct participants.

[30] Data from January 2010 through December 2013 were used to select potential members of treatment and comparison groups. The data from calendar year 2009 were only used to calculate HCC scores and the presence of substance use disorders for the year before treatment began for individuals who started treatment in calendar year 2010. Data for January through September 2014 were added to the analysis file for members of the treatment and comparison group after matching was complete. Individuals new to treatment in January through September 2014 were not included in the analysis.

[31] KMHS may conduct analyses of data that they have collected on program participants and publish their own reports about program outcomes. These analyses are likely to use different data, focus on different outcome measures, and address different specific questions compared with the data and measures we use in this report to address questions of direct interest to CMMI. As a result, KMHS may reach conclusions about program effects that differ substantially from the ones we present in this and future reports.

Figure IX.1. Cumulative enrollment by quarter, Q1–Q12—KMHS



Source: Quarterly reports submitted to the Lewin Group's reporting website for the HCIA initiative.

Note: KMHS did not define a specific enrollment target for Race to Health! because KMHS staff intended the program to reach all patients who used KMHS' outpatient services.

2. To what extent was implementation timely, conducted as planned, and responsive to site-level constraints? (RQ 17)

By design, KMHS continuously altered or improved the program components based on program leaders' observations and staff feedback. Consequently, KMHS' plans for the program evolved over the three-year award period. KMHS was largely successful in implementing its program according to schedule; however, key informants identified several challenges that slowed their progress toward program goals. These informants also described the ways the program adapted in the following areas:

- **Enhancing health information technology (IT) for data collection.** Early on, KMHS faced barriers in accessing health information from external data sources because a local hospital became affiliated with a regional hospital system. Before this affiliation, KMHS planned to join the hospital's health information exchange (HIE), which would have facilitated its access to electronic health data for more than half of the agency's patients. However, as part of the affiliation, the hospital transitioned to a different EHR system and consequently terminated the HIE.
  - This unexpected lack of electronic access to patients' health information created inefficiencies in patient data collection. For example, program staff did not have access to automated electronic alerts from the state emergency department information exchange (EDIE) within KMHS' system; medical assistants entered this information manually instead.

  - KMHS made progress in addressing this challenge in the third year of the program. For example, in October 2014, KMHS established an electronic EDIE interface that automatically populated KMHS' EHR system, Profiler, with information on patients' ED visits. KMHS launched a similar interface for patients' laboratory data in April 2015. However, program staff continued to manually enter other important health data, such as records of visits to PCPs.
  - Program leaders noted that without a shared EHR system, medical assistants had trouble deciding when to reach out to PCPs for updated medical information. At the time of the site visit (April 2015), program leaders reported that KMHS was beginning to work with a local federally qualified health center (Peninsula Community Health Services) as a partner to facilitate the agency's access to a state HIE, which they expected to "make things a lot easier."

- **Use of patient data.** Program leaders and frontline staff also remarked that it was a challenge to understand and effectively use patient health data once these data became available. One program leader described the agency's early experiences with patient 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 extract useful data and promote its use by care teams in ways that allowed "data to become information that can inform the treatment plan." For example, clinical program leaders developed guidelines for medical assistants' use of EDIE data, instructing medical assistants to run reports of ED discharge diagnoses only for patients who met certain thresholds (two ED visits in one week, or five ED visits in 90 days).
  - Several program leaders and frontline staff noted that KMHS program staff used data more consistently and expansively in the third year of the program's implementation. One frontline staff member, for example, described how KMHS staff would give hard copies of blood pressure readings obtained during mental health visits to patients with hypertension so they could share them with their PCP at PCP appointments. This gave PCPs opportunity to review readings that encompassed more points in time. Several staff members also said staff were increasingly using ED and other data in collaborative care conferences with internal and external partners about patients' care.
  - However, it was still a challenge to get staff to use some of the program's data resources. For example, one staff member commented that "the care management report was put together and available a year ago, and we're still just barely seeing teams using it."

- **Programming targeted to chronic disease management and wellness promotion; staff education.** Program leaders and staff noted Race to Health! was initially slow in implementing the components of the program that were designed to educate participants about chronic disease and wellness. However, program leaders commented that "in the last year we have really picked up the pace," implementing several groups focused on improving patients' understanding of wellness and on ways to self-manage a chronic disease.
  - Leaders said this aspect of the program took longer to establish because they first had to build an agency-wide commitment to whole health as a foundation of the work.

  - One program leader described Race to Health!'s health IT, staff education, and wellness programming components as "pistons in an engine," all working together to make the wellness programming run during the third year. For example, more tailored use of the health analytics capabilities made it easier to identify patients who could use wellness programming, while training and educating staff in physical health enabled staff to talk to patients about their whole health in a way they had not done previously.

- Evolving role of the medical assistant. As described in our 2014 annual report, KMHS created a new position of medical assistant; this position supported the care teams in collecting and using patients' physical health data. KMHS also sought more ways to engage the medical assistants in patients' care.
  - In the program's third year, both frontline staff and program leaders reported increasing clarity about the medical assistant role. One clinical leader commented that "the other people within the [care] team have come to understand [the medical assistant role] better and the role looks more similar across teams."
  - Program leaders suggested this role will continue to evolve in response to medical assistants' interest in increasing their interaction with patients. One program leader stated, "They have all indicated that they really have a desire to talk more with clients about their physical health, including referring them to groups or doing some coaching."
  - KMHS appeared to be moving toward more involvement with patients' physical health; at the time of the site visit, one medical assistant was in training to co-facilitate a diabetes education group for KMHS patients. In contrast with the 2014 site visit, this year medical assistants viewed their role as having "only minor differences from what you would do in a PCP office." To continue defining this role, KMHS program leaders engaged a consultant during the program's third year to clarify an appropriate scope of work for medical assistants and enable these staff to work at the top of the functions allowed under their license.

- **Redefined role of the nurse.** Throughout its implementation of the program, KMHS has found it challenging to redefine the role of its nursing staff within care teams.
  - One program leader said, "I don't think we have been nearly as successful as we want to be in moving our nurses away from that psychiatric role to a more whole-health role." Several respondents echoed this and added that some of the nursing staff have been somewhat reluctant to relinquish traditional psychiatric nursing responsibilities and focus more broadly on physical and behavioral health.
  - In addition, some respondents reported that nurses were concerned that the expanding responsibilities of the medical assistants would overlap with and detract from the nursing role. For example, psychiatric nurses suggested to a clinical leader that nursing staff should be the ones to call in prescription refills for patients; however, the program's consultant on the medical assistant's role said this responsibility would fit reasonably within the medical assistant scope of work.

  - Program leaders reflected that as they refined the responsibilities of the medical assistants, "it is helping us think about ways we can free up the nurses to do more of what we really want them to do," which is to serve as authorities on physical health for the entire care team, in addition to monitoring psychiatric medications. Program leaders reported some success with this in the realm of tobacco cessation, noting that nurses will receive referrals from other care team members when one of the care team's patients wants to quit smoking.
  - KMHS expected the ongoing challenge to remain "trying to figure out how we can give [nurses] new responsibilities without overloading them."

- **Expanded staffing.** KMHS increased its use of peer support specialists in the third year of the program.
  - Previously, KMHS assigned peers to some care teams on a part-time basis. The new chief clinical officer pushed for an increased peer presence. As of April 2015, all care teams had the equivalent of one full-time peer staff member.
  - KMHS did not use HCIA funding to support the peer position; however, program leaders described the peer role as instrumental in the program's health and wellness work. For example, KMHS had four peer staff trained in a peer support group training model (Whole Health Action Management) for use with patients as part of Race to Health!'s wellness programming.
  - KMHS also added other new staff in roles not funded by HCIA, but critical to Race to Health! implementation. As the agency added more care teams in response to increased caseloads (Section D.2), it also hired new medical assistants to support these care teams and another Healthy Families Coordinator to meet the demands of the child and family care teams.

- **Adoption of program protocols.** Program administrators reported that in the third year of the award period, the care teams' adoption of and conformity to protocols continued to vary.
  - One medical assistant noted, "I think each team does their emergency department alerts a little bit differently, but we all check them regularly."
  - Respondents attributed the lack of consistency to the rapid growth in care teams as a result of the expansion of Medicaid under the Affordable Care Act. This required program leaders to develop the capacity to take on the expanded Medicaid population at the same time it was managing the rollout of Race to Health! As one program leader put it, "It's simply slowed down the normalizing of all of the elements of Race to Health!."

- **Increased community engagement.** Overall, KMHS implemented its community-based program components as planned and on schedule.
  - Throughout the award period, the behavioral health provider increased her engagement with providers at three PCPs and received referrals for patients with a broader spectrum of behavioral health issues, including substance use issues. She continued to provide on-site consultation and brief interventions to patients at these PCPs. The psychiatric consultant also continued to provide training and consultation to other PCPs in the community.

- Respondents noted one continuing barrier to this work. Some providers remain unaware of their patients' misuse of prescription medication and are reluctant to change their prescribing habits. As one KMHS staff member commented, "we're really at stage one of systemic change." The behavioral health provider and psychiatric consultant engaged KMHS' internal consultant on co-occurring disorders to bring that person's expertise on substance use issues into the PCPs.

3. How did unexpected events support or conflict with successful implementation of the innovation? (RQ 8)

In addition to the project's unexpected loss of its HIE access, the agency's chief medical officer and its internal consultant on co-occurring disorders left their positions early in the program's implementation. Program leaders noted these roles were critical to shaping the direction of Race to Health! and the departures represented a considerable loss for the agency's substance use integration work. Over the course of the award period, the agency built up its clinical leadership by hiring a new chief medical officer, a chief clinical officer, and internal consultants who specialized in co-occurring disorders. Each of these individuals said they came to KMHS because of its leadership and innovation in integrated care.

C. Program effectiveness

In this section, we present the results of an impact analysis on four core outcome measures: total expenditures, hospitalizations, hospital readmissions, and ED visits. These measures include all Medicare covered services provided to the population included in the analysis. First, we provide an overview of our analytic methods. We then report the findings and limitations of the analysis. Next, we include qualitative findings addressing specific research questions related to program effectiveness.

1. Overview of quantitative methods and population

As noted in section A, we were only able to conduct a rigorous impact analysis for about 15percent of the overall population served by KMHS served because of limitations in data availability. Specifically, we could not include (1) non-dual Medicaid enrollees; we did not have data for managed care enrollees and there were substantial lags in the availability of Medicaid enrollment and claims data, or (2) Medicare enrollees in Medicare Advantage managed care plans, because data on these individual's use of services are unavailable. Thus, our analysis includes only patients who have Medicare as their primary payer, are enrolled in Medicare Parts A and B, and are not enrolled in Medicare Advantage. We obtained the Medicare administrative data for the analysis from the CMS Virtual Research Data Center.

We defined the members of the intervention group as individuals who had an outpatient mental health visit at KMHS between January 1, 2010, and December 31, 2013, as indicated by the existence of a Medicare claim for services (see Appendix A for more details). We then used propensity score matching methods to select a comparison group. Specifically, we took the following steps to identify the comparison group:

- Using the Substance Abuse and Mental Health Services Administration's Treatment Finder, we identified facilities in Washington State with characteristics similar to those of KMHS (see Appendix A for more details)

- In parallel with the approach used to identify the intervention group at KMHS, we identified all individuals who had an outpatient mental health visit between January 1, 2010, and December 31, 2013, at these 16 comparison facilities as the potential pool from which to construct a comparison group.
- Using this pool, we matched up to five of its individual members to each intervention group member. With the matching algorithm, we attempted to identify comparison group members who were similar to the members of the intervention group on several key characteristics that are predictive of future Medicare service use and expenditures, such as demographics, disability status, Hierarchical Condition Categories (HCC) score, [32]dual Medicare/Medicaid enrollee status, and, to the extent feasible, the most common mental health diagnosis listed on the person's claims.

Outcomes were analyzed from January 1, 2010 through September 30, 2014.[33] Detailed information on the matching methodology is available in Appendix A.

2. Quantitative findings on program effectiveness

Overall, the results suggest that the program reduced overall Medicare expenditures and hospitalizations for its participants relative to the experience of the comparison group, but the program did not appear to impact readmission rates or ED visits (Table IX.1). The analyses also have important limitations (described below) and should be interpreted with caution. Notable findings are as follows:

- In the program's first year, it had no significant impact on total Medicare expenditures. However, in the second year, we found that the total expenditures per patient decreased for both the intervention and comparison groups, but the decrease in the intervention group was greater. The estimated impact of the program was a $3,646 per-person decrease in total expenditures (a 30 percent decrease in expenditures). This difference was statistically significant at the .01 level (p-value < 0.01).
- The impact estimates indicated a decrease in hospitalizations for the intervention group in year 1 (17 percent) and year 2 (23 percent) of the program; however, both estimates were only marginally statistically significant at the .10 level (p-values = 0.07 and 0.07, respectively).

[32] HCC scores are calculated based on a risk adjustment model developed and used by CMS to adjust payments to Medicare Advantage for the relative risk of claims by enrollees. The score measures the predicted health expenditures for a given beneficiary relative to those for the average Medicare beneficiary. Each beneficiary's HCC score is estimated based on diagnoses in the past 12 months of Medicare claims data, as well as demographic characteristics. We calculated HCC scores following CMS methodology. They were calculated based on the conditions reported in Medicare claims data in the 12-month period before the month the enrollee first visited KMHS or a comparison facility. For individuals who had a mental health visit in January 2010, the HCC score was calculated based on Medicare data for January through December 2009.

[33] The pre-intervention or baseline period includes observable Medicare claims data for the analysis population from January 2010 through the date of enrollment in the program. For individuals who began treatment at KMHS or a comparison mental health facility in January 2013 or earlier, the intervention period begins in January 2013, the month that KMHS fully implemented Race to Health! The intervention period began in the month of the first visit at KMHS or a comparison facility for individuals whose first visit was after January 2013.

- Compared with the baseline year, readmissions decreased for both the intervention and comparison group during years 1 and 2 of the program, but the estimated impact of the program was not statistically significant in either year (p-values = 0.34 and 0.21, respectively).
- Compared with the baseline year, ED visits increased for both the intervention and comparison group in year 1, but there was no notable impact associated with the program. Between year 1 and year 2 of the program, ED visits declined for both the intervention and comparison group, but the decline was greater for the intervention group. However, the estimated impact of the program (an 8-percent decline) was not statistically significant in year 2 (p-value = 0.34).

Table IX.1. Impact estimates for KMHS, annual results

	Regression-adjusted mean for treatment group	Regression-adjusted mean for comparison group	Estimated impact[b]		
			Value	Percent	P-value
Total expenditures per patient[a]					
Baseline	$15,869	$14,921	N.A.	N.A.	N.A.
IY1	$15,264	$15,803	-$1,487	-9	0.37
IY2	$8,416	$11,114	-$3,646***	-30	<0.01
Hospitalizations per patient					
Baseline	0.71	0.70	N.A.	N.A.	N.A.
IY1	0.43	0.50	-0.09*	-17	0.07
IY2	0.29	0.36	-0.09*	-23	0.07
Readmissions per patient					
Baseline	0.13	0.12	N.A.	N.A.	N.A.
IY1	0.08	0.11	-0.03	-26	0.34
IY2	0.03	0.05	-0.03	-44	0.21
ED visits per patient					
Baseline	1.49	1.48	N.A.	N.A.	N.A.
IY1	1.73	1.71	0.01	0	0.95
IY2	1.48	1.60	-0.12	-8	0.34
Number of observations	3,610	13,407	N.A.	N.A.	N.A.
Number of unique patients	874	3,090	N.A.	N.A.	N.A.

Source: Mathematica analysis of FFS Medicare expenditures and utilization for the period January 2010–September 2014. Individuals treated at KMHS between January 2010 and December 2013 were included in the analysis. Data for calendar year 2009 were used to develop indicators of baseline health status.

Note: To be included in either the intervention group or the comparison group, individuals had to be enrolled in Medicare FFS, Parts A and B, and have Medicare as their primary payer for the entire analysis period. All regression models control for age (linear and logged), race, ethnicity, gender, dual Medicare/Medicaid enrollee status, time in mental health treatment (linear and logged), whether 12 months of baseline data were available, SMI diagnoses, alcohol and drug use disorder diagnoses, disability status, and HCC score.

[a] Results are based on a GLM model to parallel the observed distribution of expenditures in the analysis population and reduce sensitivity to outliers. Sensitivity testing indicates an OLS model would find the same direction of the effect and significance level.

[b] We derived the impact estimates in Stata using the lincom command to compare the difference between the intervention and baseline period means for the treatment and comparison groups.

*/**/*** Impact estimates are statistically different at the .10/.05/.01 level, two-tailed test.

N.A. = Not applicable.

IY1/IY2 = IY stands for intervention year.

3. Limitations of the quantitative analysis

The results presented in this section are preliminary and have important limitations:

- They reflect the intervention period (IY1) from January 2013 (the program start date) to September 2014. HCIA funding for the program continues through June 2015. Consequently, findings may change as data become available for additional participants and for the full period of the intervention.
- As explained above, the current analysis is limited to Medicare FFS enrollees and Medicare covered services. Therefore, our findings are not generalizable to all KMHS patients and services.
- In addition to providing services to all KMHS patients, a specific component of Race to Health! targeted a subpopulation of high-risk patients who were given more intensive monitoring and wellness programming. This group may benefit more from the intervention compared with all KMHS patients. We plan to analyze impacts for these high-risk patients separately in the future.

In light of the preliminary nature of our findings to date, it is important to avoid drawing premature conclusions about the impact of KMHS's program.

4. Qualitative findings on program effectiveness: From the perspective of key informants, to what extent does the intervention improve desired health outcomes and quality of life for participants? (RQ 23 and RQ29)

Many respondents described ways in which the program improved patients' physical health by connecting them to medical treatment or preventive services. Respondents noted that KMHS staff's improved awareness of physical health conditions enhanced their ability to discuss these issues with patients clients and help connect them to necessary medical care.

- For example, one respondent noted, "there are clients who've been identified with high blood pressure issues that would [otherwise] not have been because they don't go anywhere else for health care."
- Another respondent reported that patients appeared to appreciate this support: "I have heard several clients say things like, 'Wow, I'm surprised that you care about me in this way. I didn't know you cared about my health. Thank you for helping me to take care of it.'" Respondents suggested that increased awareness of whole health and advocacy on behalf of the patient were key facilitators of better health outcomes.

Respondents also described how the agency's use of ED data and other data on physical health helped staff in their efforts to improve participants' health. Some program staff attributed this to increased awareness of patients' ED visits, noting that this information helped them redirect participants to appropriate care.

- For example, one program staff member commented, "I think [patients] are overall getting better care because I feel like we're catching a lot of stuff that we didn't catch before. We're finding out why they're not following through with medical appointments, what's causing them to have their diabetes out of control and what we can do to help them rein it in."

- Another respondent noted that prescribers of psychiatric medication were gaining a better understanding of the medications prescribed to their patients because of the information medical assistants gathered from PCPs. The respondent thought this information "helps them to make the decision on what medication they're going to prescribe, which I think will also probably drive down emergency room visits as far as medication issues go."

In addition, respondents talked about how the program's wellness groups benefited patients. For example, one program staff member described a patient who decided that she wanted to quit smoking, joined the smoking cessation group, started walking for exercise, and now says "I walk every day and I feel great." In addition, respondents described how wellness programing empowered patients to manage their own care, noting that the Living Well group encourages patients to build relationships with their physicians and take more ownership of their health and care.

Staff at KMHS, Harrison Health Partners (HHP), and other community partner organizations commented that Race to Health! improved collaboration among the stakeholders involved in the patients' care and this should positively affect patient outcomes in the longer term. As part of Race to Health!, KMHS started conducting "collaborative care conferences," which brought together providers and other community stakeholders (EMTs, the police department, health plans) to coordinate a specific patient's care.

- One HHP provider noted: "I think getting everybody together on the same [page] and [taking] a concerted approach is going to pay dividends."
- Another KMHS frontline staff member noted the care conferences helped streamline treatment plans and engender partnerships with EDs, which will help patients get the best possible care.
- One KMHS respondent noted "It's much easier when you're working together rather than blaming the other person for not getting the patient better."

Respondents also described how the behavioral health professional helped patients who received primary care from providers at the three HHP practices.

- One physician commented that the behavioral health care provider supported patients on obtaining housing and addressing other social factors that influence health—factors the practice was otherwise not likely to address. The behavioral health care provider was able to identify appropriate resources in the community and quickly connect patients to those resources.
- In addition, respondents extolled the benefits of having the behavioral health care provider physically located at the practice. One respondent commented, "It's so much easier to deal with a patient when they're here rather than having them go home, waiting for somebody to call, and then playing phone tag back and forth," noting "there's none of this delay in trying to reach her."

D. Context

1. To what extent did organizational features support or conflict with implementation? (RQ 104)

As described in the first annual report, Race to Health! was the result of KMHS's long-standing intention to move toward integrated care. As one program leader expressed, "We didn't put anything in [the HCIA proposal] that we didn't already think we should be doing." Another program leader noted, "It was a model we wanted to do, whether we got [HCIA funding] or not." To prepare for this program and the associated cultural shift, the agency restructured its existing staff into care teams and physically moved staff within the buildings to facilitate collaboration.

In conceptualizing and launching its program, KMHS benefited from having a number of senior staff who are recognized throughout the state for their expertise. For example, several existing KMHS staff had strengths in co-occurring disorders and led the push for integration. These included the former chief medical officer, who specialized in addictions work, and the former internal consultant in co-occurring substance use disorders, lauded by program leaders as "truly a leader statewide."

Program leaders also suggested the agency's existing relationships with other providers and local and state officials were critical to the program's success. For example, KMHS had a strong existing relationship with HHP, and one of HHP's primary care providers was stationed weekly at KMHS. As one program leader stated, "we are not John Wayne; we don't do things by ourselves." Rather, the agency involved "40+ other strategic partners" in its work in the community. Respondents also noted that KMHS's unique position as the "sole provider" of public mental health services in Kitsap County fostered those relationships and helped PCPs understand the need for and embrace consultation services offered by the behavioral health professional and the psychiatric consultant.

Race to Health! also benefited from its ability to fill new program staff positions with internal candidates. One program leader noted, "I don't think we could have done it without hiring from within our [existing] staff." Clinical leaders echoed this sentiment, commenting that "the staff came with a strong knowledge base of their own work that was really integral. They needed that maturity" to implement the changes under the program.

Several financing mechanisms within the agency (described in the first annual report on this project) also proved important to restructuring and integrating substance use disorder services into care teams. The agency maintained a waiver of selected substance use disorder treatment regulations from the state legislature, which allowed KMHS to operate its integrated substance use and mental health treatment model under the state's mental health licensure requirements, rather than having to meet additional regulations on the treatment of substance use disorders. KMHS also worked with the county substance abuse treatment coordinator (with whom the agency has long collaborated) to obtain flexible state funding for activities not otherwise covered by Medicaid, such as training in co-occurring disorders, drop-in groups, and consultation for mental health clinicians. These arrangements, in addition to its sub-capitated, at-risk financing arrangement for mental health services, bolstered the agency's ability to pursue an integration strategy as part of the new program.

2. To what extent did the policy and political environment support or conflict with implementation? (RQ 106)

Respondents noted that Washington's expansion of Medicaid under the Affordable Care Act likely contributed to an influx of new participants, which affected implementation of Race to Health! Program leaders estimated a 40-percent increase in adults seeking outpatient services at KMHS.

- As a result, leaders had to shift some attention away from Race to Health! to build treatment capacity, and program staff initially experienced significantly increased caseloads. Program administrators mentioned that the influx of patients and increased caseloads posed somewhat of a barrier to implementation. Staff were asked to take on new ways of thinking and working under Race to Health! in the midst of an increasing, sometimes overwhelming, workload.
- To mitigate this challenge, KMHS expanded the number of outpatient care teams and the number of staff in each care team. One program leader described the magnitude of this change: "Had we not had the expansion, we would have better controls over all our variables, but you are talking about an infusion of something that's totally rocked our ship."

On a related note, several staff indicated that the Medicaid expansion posed a staffing challenge: the increased demand for mental health services resulted in an increased demand for qualified health care workers at KMHS and elsewhere. One respondent noted that this change affected not only KMHS, but also HHP, which was short-staffed. However, respondents believed staff were attracted to KMHS because of its innovative work, including Race to Health! One program leader described this challenge as "the impact of Medicaid expansion on the ability to not only find the staff, but then to quickly imbue them with this still newly unfolding, integrated care because they may or may not have been exposed to that previously. So we're in a sort of a continuous training state."

Several respondents also noted challenges posed by 42CFR Part 2, the federal regulation about 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 among providers. As one respondent noted, "We really won't have integration without finding some way to both protect people under 42CFR but allow us to share the chemical dependency information."

E. Summary and conclusions

KMHS implemented Race to Health! as scheduled, while expecting and allowing for the program to evolve and adjust over the course of the award period. The agency continued to refine its use of patient data, program staff roles, and program protocols in year 3, and built on its efforts during early implementation to introduce additional programming and incorporate new staff. Most respondents noted substantial improvements in patient health outcomes and quality of life. Several characteristics of KMHS may have aided program implementation, including the agency's funding arrangement, long-standing relationships with external partners, and the particular strengths of its staff.

Overall, the preliminary quantitative findings suggest that although Race to Health! had a limited impact on acute service use and expenditures for Medicare-enrolled participants in the initial year, the benefits of the program may have increased with exposure. The only significant decline observed in the initial year was a 17-percent decline in hospitalizations (p-value = 0.07); however the program achieved a substantial (30-percent) decline in total Medicare expenditures (p-value = <0.01) as well as a 23-percent decline in hospitalizations (p-value = 0.07) in year 2.

The quantitative results presented in this chapter are preliminary and only cover the first 21 months of the 30-month program. Moreover, they include only Medicare enrollees, who represent about 15 percent of KMHS patients. Therefore, it is important to avoid drawing premature conclusions about the implementation or outcomes of KMHS's program based on these preliminary findings.

X. MAIMONIDES MEDICAL CENTER

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 7,500 participants with serious mental illness (SMI). The program used an electronic care coordination platform and engaged a care management workforce to give participants a virtual medical and mental health home.

To implement the program, MMC collaborated with members of the Brooklyn Care Coordination Consortium, a group of more than 20 social service agencies and medical institutions. The program engaged participants' existing medical, mental health, and community service providers, who were supported by HCIA-funded care management staff, to create multidisciplinary care teams. Members of the care team shared information through the care coordination platform, thus giving participants a virtually co-located, coordinated medical home.

MMC and its partners targeted adults (age 18 and older) who lived or received care in certain zip codes in southwest Brooklyn. Individuals with the following diagnoses were eligible for the program:

- Certain depressive, bipolar, and other mood disorders
- Schizophrenia and certain other psychotic disorders

Before the HCIA funding, New York State granted MMC status as a Medicaid health home, and MMC ultimately[34] expected to provide health home services to roughly 7,000 Medicaid enrollees. In turn, the HCIA award gave MMC (1) the capacity to focus specifically on providing care management to 500 additional individuals with Medicare (including dual eligibles), commercial insurance, or no insurance, and (2) the startup funding for the training and technology infrastructure necessary to provide health homes to the larger target population of 7,500 participants with SMI—all of whom may have benefited from the infrastructure development and staff training funded through the HCIA.

For this evaluation, Medicare-only enrollees, commercially insured, or uninsured individuals are "direct participants." Enrollees with Medicaid, including dual eligibles, are "indirect participants" who also benefit from the improved infrastructure paid for by the HCIA. Medicare-only participants are considered "direct participants" because they benefit from HCIA funding for care management and the startup funding for better infrastructure, whereas participants with Medicaid benefits (including dual eligibles) are considered "indirect participants" who benefit from only the HCIA funding used for better infrastructure.

[34] As noted in section B.3., midway through the award period, MMC altered its original enrollment targets from 2,000 direct participants and 5,500 indirect participants to 500 direct participants and 7,000 indirect participants.

The findings in this report are based on quantitative and qualitative data collected or received through June 1, 2015, and enrollment data reported throughout the award period. Data sources included:

- Enrollment data submitted by MMC to the reporting website maintained by the Lewin Group for the HCIA initiative
- Medicare claims and enrollment data for program participants, which we extracted using participant identifiers provided by MMC
- Qualitative data, including telephone interviews and an in-person site visit in May 2015. Mathematica conducted in-depth interviews with key awardee staff, members of the workforce, and other stakeholders. In addition, Mathematica convened focus groups with members of the workforce and with program participants. Data from the participant focus group are not part of this report; instead, we provide an in-depth analysis of this focus group in our seventh quarterly report.

The program effectiveness results we present in this chapter are preliminary. Trends in outcome measures that are not evident now may become apparent when we are able to analyze data that reflect a longer period of enrollment in the program. In light of the preliminary nature of our analyses to date, it is important to avoid drawing premature conclusions about the effectiveness of MMC's program based on these analyses.

The questions we address in this report were selected in consultation with our COR. Specifically, we examine questions related to implementation effectiveness in Section B, program effectiveness in Section C, and program and policy context in Section D. We close with an overall summary and conclusion (Section E).[35]

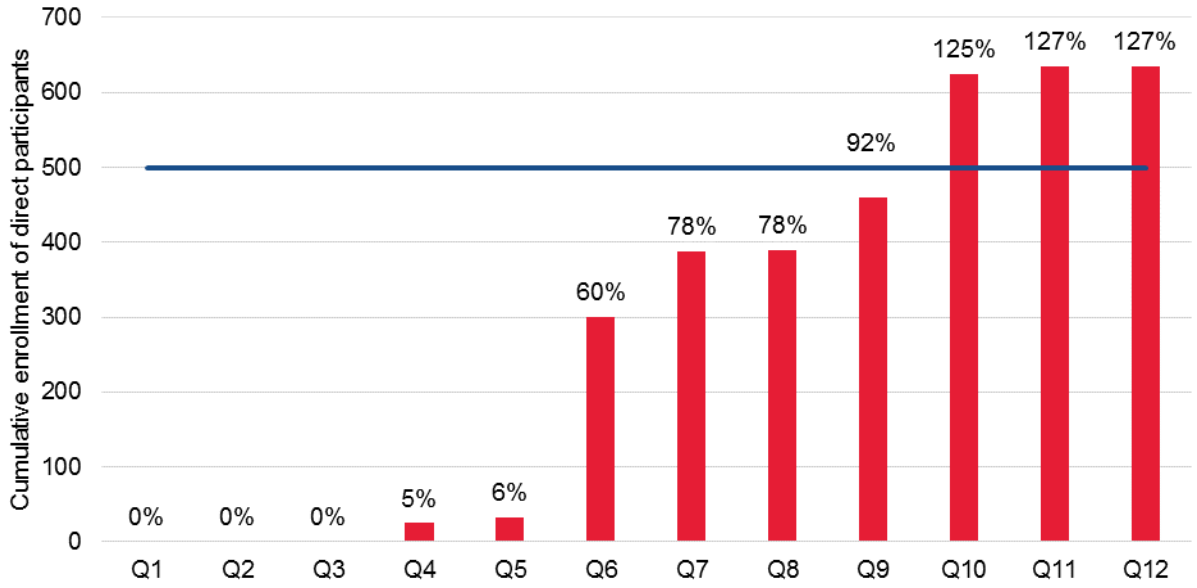
B. Implementation effectiveness

1. Who was enrolled in the program?

By the end of the project's 12th quarter (June 30, 2015), MMC had enrolled 635 direct participants, significantly exceeding its enrollment target of 500. MMC began enrolling participants into the program in Q4 of the award. After a slow start in the first two quarters, enrollment increased rapidly in Q6 and continued to increase over the course of the award. MMC reached its goal of 500 participants in Q10 and ultimately enrolled 635 participants, exceeding its goal in the last three quarters of the award period (Q10 through Q12).

[35] MMC may conduct analyses of data that they have collected on program participants and publish their own reports about program outcomes. These analyses are likely to use different data, focus on different outcome measures, and address different specific questions compared with the data and measures we use in this report to address questions of direct interest to CMMI. As a result, MMC may reach conclusions about program effects that differ substantially from the ones we present in this and future reports.

Figure X.1. Percent of target enrollment achieved by quarter, Q1–Q12—MMC



Source: Quarterly reports submitted to the Lewin Group's reporting website for the HCIA initiative.
Note: As displayed by the horizontal bar, MMC's target enrollment was 500 unique participants.

2. To what extent was implementation timely, conducted as planned, and responsive to site-level constraints? (RQ 17)

MMC and its partner organizations were largely successful in implementing the program as planned and on time. MMC sought to adapt and improve on its original operational plans and encourage significant customization among partner organizations. As a result, the program evolved throughout the award period in response to challenges and lessons learned, particularly in the following areas:

- **Participant recruitment and enrollment.** MMC and its partner organizations originally intended to draw many of the participants from lists of people assigned to the health homes program by the state. Program staff found these lists less useful than expected, because the lists often contained outdated addresses and contact information. As a result, MMC and its partner organizations shifted their strategies in several ways:
  - Program staff identified eligible individuals by using internal MMC administrative data and leveraging their partnerships with Medicare managed care and commercial payers.
  - Care management organizations did outreach within their organizations and with other community-based organizations to identify potential direct participants.
  - As time went on, physical and behavioral health providers who were familiar with the program began making referrals to it.
  - In addition, in the fall of 2014, MMC added program participants through its new partnership with another large health home in Brooklyn administered by Coordinated Behavioral Care (CBC). The CBC health home had preexisting relationships with many of MMC's existing care management partner organizations and agreed to transition to MMC's care coordination platform, use the program's comprehensive assessment tool to

identify participant needs, and adhere to the program's standards of care. As a result, individuals enrolled in CBC's health home who met the program's criteria were included in the program's enrollment numbers, because they benefitted from HCIA-supported protocols, training, and technology. This may account for the large increase in enrollment between Q9 and Q10 (Figure X.1).

- **Staffing changes at partner organizations.** MMC's care management partners were not required to follow a particular staffing configuration and could adapt the program's staffing guidelines, to a degree, to fit their unique organizational needs. An administrator from a partner organization described the approach as "flexible," noting appreciatively that "we've been given some room to figure out what works best." As a result, program staffing shifted throughout the award period. For example:
  - Many of the partners initially assigned care managers the responsibility of reaching out to new participants and following up with enrolled participants. MMC and its partners reported that many of these organizations later changed their staffing models to separate outreach-focused staff from staff who provide care management. As one administrator at a partner organization noted, "[Our organization] found that we do much better when we starkly divide the outreach, intake, and enrollment tasks and responsibilities from the care coordination tasks and responsibilities."
  - In some cases, care management organizations also used different terminology to refer to their care management staff. For example, some partner organizations referred to their care managers as "care coordinators" in accordance with the organization's existing staff titles.
- **Provider engagement.** Initially, MMC intended to fully engage each participant's existing providers in the program by training these providers to use the care coordination platform to collaborate and share information with the other members of the participant's care team. However, program leaders ultimately found it challenging and resource-intensive to engage providers who saw only a small number of program participants or who worked in small practices with limited administrative and technological support. In response, MMC shifted its provider engagement strategy as follows:
  - Program staff began to identify preferred providers who served a critical mass of program participants and targeted those providers for care coordination platform "dashboard" training and full engagement in the program.
  - For participants who were not already connected to physical and behavioral health providers, MMC created guidelines for care managers to refer participants to preferred providers who already worked within the program's care model.
- **Care coordination platform.** Program administrators and frontline respondents described changes in the program's expectations for providers' use of the platform, or dashboard, in response to the aforementioned challenges to engaging participants' existing providers and the providers' reluctance to use the care coordination platform as originally intended. Physical and behavioral health care providers reported that although they sometimes used the care coordination platform to view information on participants' care (for example, medication prescribed by other providers), they did not often enter information into the platform or use it to communicate with other care team members as expected. Instead,

providers often relied on the care manager for updates on participant care and for communication with other providers. As a result, the program leaders' vision for the care coordination platform evolved over time in the following ways:

- Care management staff took on the responsibility to update care plans and coordinate among members of the care team, shifting these tasks away from providers. As one program leader noted, "Early on, we had hopes that everyone on the care team would embrace the [platform] as a way to document things in the care plan. What we've realized is that providers are not documenting on it, and they're not going to document on it. It's really a care management platform and what we want providers to do is look at it and review it. "
- To support this new approach, program staff worked with the health IT contractor (GSI Health) to create a provider-specific landing page, making it easier for providers to find relevant information on the participants in their care.
- Some care management staff reported their frustration with the burden that was placed on them to do more coordination and documentation as expectations for the platform evolved. One respondent from a partner organization noted, "It's become a telephonic model and it's not supposed to be." Another respondent said that for care management staff, entering information into the platform "takes away from client care, which is a big, constant problem."
- As of May 2015, program staff were also exploring options that would allow providers to access the care coordination platform directly from their existing electronic health records system (in other words, transitioning to a single sign-on process).

- **Care standards.** Program leaders worked with care management partners throughout the award period to refine the program's standards of care and develop new guidelines when needed. For example, program leaders developed specific guidance for responding to critical events such as hospitalizations or arrests:
  - Care management staff were expected to conduct a clinical case conference with all of a participant's care team members (that is, the primary care provider, psychiatrist, therapist, and care manager) when a hospitalization or other critical event occurred.
  - In some cases, these conferences included others involved in the participant's care, such as a parole officer or housing contractor.

Program leaders also developed guidelines to address specific challenges as they arose. For example, the program created guidelines on the steps outreach staff should take when facing a language barrier or when they were unable to locate participants.

- Total cost of care model. As described in our 2014 annual report, program leaders reported initial delays in model development due to challenges associated with obtaining Medicaid claims data from New York State. Despite these delays, program leaders successfully worked with the actuarial contractor (Milliman) to develop the model, which incorporated the cost of physical health services, behavioral health services, and care management. Program leaders reported they plan to use the model to negotiate rates with managed care organizations to support continuation of the program under part of the state's behavioral health Medicaid Health and Recovery Plans.

3. How did unexpected events support or conflict with successful implementation of the innovation? (RQ 8)

As described in our 2014 annual report, dual eligible enrollees were eligible for the Medicaid Health Homes Program, which narrowed the expected pool of direct participants in the HCIA component of the program. As a result, MMC leaders reduced the program's enrollment goal for HCIA direct participants from 2,000 to 500, while increasing the indirect enrollment target from 5,500 to 7,000. The program maintained the same diagnostic and geographic enrollment criteria for the larger target population of 7,500 participants.

In addition, two program leaders left the organization in the third year of the program's implementation. The program's research and evaluation manager departed in November 2014, and the program's administrative director left in January 2015. Other program leaders stepped in to assume responsibility for the day-to-day operations of the program, and new staff were brought on to ensure that the program's goals would still be met.

C. Program effectiveness

In this section, we present the results of impact analyses for three of the four core outcome measures: total expenditures, hospitalizations, and emergency department (ED) visits. We do not present results for hospital readmissions due to small sample sizes. These outcome measures include all Medicare-covered psychiatric and non-psychiatric services provided to the analysis population, with the corresponding expenditures. First, we review our analytic methods. We then report the findings and limitations of the analysis.

1. Overview of quantitative methods and population

Our analysis includes only participants who have Medicare as their primary payer, are enrolled in Medicare Parts A and B, and are not enrolled in Medicare Advantage. Our analytic population includes both Medicare-only participants and participants dually enrolled in Medicaid and Medicare (the data do not allow us to distinguish between full and partial Medicaid enrollees).

For the purposes of the evaluation, we designated the intervention group as Medicare enrollees who enrolled in the program and who had evidence of bipolar disorder, schizophrenia, and/or depression in their Medicare claims data. Enrollees were flagged as having a condition if they had at least one inpatient or two or more outpatient Medicare claims (not including prescription drugs) with the relevant diagnosis in the two years before enrollment (see Appendix A for details). In addition, we required at least six months of fee-for-service (FFS) Medicare data in the year before enrollment in the program. We then used propensity score matching methods to select a comparison group. Specifically, we took the following steps:

- We developed a comparison group of Medicare enrollees with schizophrenia, bipolar disorder, or depression residing in Philadelphia, Pennsylvania. We selected Philadelphia because it is a major metropolitan area that is close to New York, but in a state that did not implement a Medicaid health home program. We chose an alternative state that did not implement a Medicaid health home program so that we could create a comparison sample

that included full dual eligibles who were not eligible for health homes.[36] Potential comparison group members were identified as those who resided in Philadelphia and had evidence of schizophrenia, bipolar disorder, or depression in Medicare claims data at any time in the year before the start of the MMC program or through the last enrollment month included in these analyses (March 2014). More information on the time periods used for the analysis is in Appendix A.

- 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 they had been in the intervention group. The pseudo-enrollment month was the first month between February 2013 and March 2014 that followed a month when the enrollee had at least one of the relevant diagnoses (depression, schizophrenia, and bipolar disorder) and a visit with a primary care provider.
- We required potential members of the comparison group to have at least six months of FFS, Parts A and B Medicare enrollment with Medicare as primary payer in the year before the pseudo-enrollment month.
- We matched up to seven potential comparison group members to each member of the intervention group. We first exact matched on schizophrenia diagnosis, bipolar disorder or depression diagnosis, dual eligibility status, and gender. We then fit propensity score models using other important characteristics predictive of intervention group status, such as month and year of enrollment (or pseudo-enrollment), number of months with Parts A and B FFS Medicare with Medicare as primary payer in the baseline period, number of qualifying SMI diagnoses (1–3), age, race, HCC score, and the number of hospitalizations and ED visits in the one to 6 months before enrollment and the 7 to 12 months before enrollment.[37]
- Detailed information on the matching methodology is available in Appendix A.

2. Quantitative findings on program effectiveness

Our impact analyses showed no statistically significant results; therefore, any differences between the intervention and comparison groups in the form of changes over time may be due to chance (Table IX.1). The current results only reflect impacts observed during the initial year of program participation. It is possible that the effects of the program might accrue over a longer period of participation.

[36] All full dual eligibles in New York are eligible for health homes. Given the similarity of health homes to the HCIA program, including full dual eligibles from New York in the comparison group would bias our results toward the null.

[37] We included separate variables for the number of hospitalizations and ED visits for 1-6 months prior to enrollment and 7-12 months prior to enrollment. We included the variables in this way as patients with serious mental illness may have cyclical or episodic periods of high need and we wanted to capture which portion of the year such episodes occurred in.

Table X.1. Impact estimates for MMC, annual results

	Regression-adjusted mean for the intervention group	Regression-adjusted mean for the comparison group	Estimated impact[a]		
			Value	Percent	P-value
Total expenditures per patient					
Baseline	$28,469	$28,313	N.A.	N.A.	N.A.
IY1	$25,280	$20,035	5,088	25	0.29
Hospitalizations per patient					
Baseline	1.4	1.6			
IY1	0.9	0.9	0.2	23	0.58
Readmissions per patient[b]					
Baseline	N.A.	N.A.	N.A.	N.A.	N.A.
IY1	N.A.	N.A.	N.A.	N.A.	N.A.
ED visits per patient					
Baseline	1.7	1.9	N.A.	N.A.	N.A.
IY1	1.4	1.5	0.2	14	0.61
Number of observations	472	578	N.A.	N.A.	N.A.
Number of unique patients	236	289	N.A.	N.A.	N.A.

Source: Mathematica analysis of FFS Medicare claims data for baseline and program periods March 2012–September 2014

Note: To be included in the analysis individuals in both the intervention and comparison group had to have at least six months of enrollment in FFS Medicare, Parts A and B with Medicare as primary payer in the year before enrollment (intervention group) or pseudo enrollment (comparison group). We applied weights to adjust for the number of months in each period that an enrollee was observable—that is, alive, enrolled in FFS Medicare, Parts A and B, with Medicare as primary payer—as well as the number of enrollees in the comparison group matched to each enrollee in the intervention group. All regression models controlled for age (linear and squared), race (white, black and other), gender, dual eligibility status, whether 12 months of baseline data were available, SMI diagnoses alone or in combination (schizophrenia only, bipolar disorder only, depression only, schizophrenia and bipolar, schizophrenia and depression, depression and bipolar or all three conditions), disability status, and HCC score.

[a] We derived the impact estimates in Stata using the lincom command to compare the difference between the intervention and baseline period means for the treatment and comparison groups.

[b] Readmission rates were not analyzed due to too few observations and too few events.

IY1 = Intervention Year 1; N.A. = Not applicable.

3. Limitations of the quantitative analysis

The results presented in this section are preliminary and have important limitations:

- They reflect the effects of the program in the first year. As noted, it may take more than one year for the program to have an effect. Consequently, longer-term findings may differ from the one-year results.
- The current analysis is limited to Medicare FFS enrollees and Medicare covered services. Therefore, our findings may not be generalizable to the large group of indirect participants with only Medicaid insurance or to direct participants with only commercial insurance.

- We excluded intervention participants with childhood emotional disturbance and psychotic disorders from our analysis. Five intervention participants were dropped based on a diagnoses of psychotic disorders that did not meet program enrollment criteria. No participants were dropped due to childhood emotional disturbance.
- We were unable to find suitable matches for 34 members of the intervention group who passed our initial criteria (at least six months of FFS Medicare data in the year prior to enrollment in the intervention), and who had schizophrenia, bipolar disorder, or depression. We dropped these participants from the analysis. More information on these participants is included in Appendix A. Exclusion of these members may have affected our results.
- Our sample size is small. The small size of the analysis population may result in greater sensitivity to outliers and model specification. We may obtain more stable results in subsequent analyses with larger samples.
- Differences in Medicare expenditures between our intervention and comparison groups may be partially driven by differences in geographic adjustments to Medicare payments between New York City and Philadelphia.[38] Because we did not make any adjustments to Medicare spending data to account for geographic differences in payment, we interpret impact estimates on Medicare spending very lightly.

In light of the preliminary nature of our findings to date, it is important to avoid drawing premature conclusions about the overall impact of MMC's program.

4. Qualitative findings on program effectiveness: From the perspective of key informants, to what extent does the program improve desired health outcomes and quality of life for participants? (RQ 23 and RQ29)

Key informants we spoke to believed the program had numerous positive effects on its participants' health, mental health, and quality of life. For example:

- Several respondents (both administrators and frontline workforce members) reported anecdotally that the program reduced the number of participants' hospitalizations as well as their use of the emergency department.
- One program administrator noted that these changes have a particularly profound effect on participants with SMI, because many hospitalizations for these participants were involuntary and, as a result, often seen as "confinement." She commented, "I think [being] free is the biggest thing that [participants] notice. From a [participant's] perspective, what's changed in my life is that I'm outside now."

Many respondents also said the program addressed social determinants of health that were critically important to participants. For instance:

[38] For example, CMS' proposed wage index and capital geographic adjustment factor for acute care hospitals in fiscal year 2012 were 1.07 and 1.05, respectively, in Philadelphia and 1.32 and 1.21, respectively, in New York-White Plains-Wayne, NY-NJ (source: https://www.cms.gov/Medicare/Medicare-Fee-for-Service-Payment/AcuteInpatientPPS/Wage-Index-Files-Items/CMS1239640.html?DLPage=2&DLEntries=10&DLSort=1&DLSortDir=descending.

- One program leader commented that "without housing, you can't deal with any of the other issues around medication adherence or keeping appointments or just organizing your life."
- Frontline staff also highlighted the importance of the program's focus on factors such as housing and transportation, stating "social [issues] have to be stabilized first" before the participant will engage on psychiatric or medical issues.
- As one provider emphasized, "If [a participant] can't get to the doctor, then you can't expect improvement in any measure."

Respondents also highlighted the care managers, citing participants' strong relationships with them and describing the care managers' efforts to coordinate the participants' providers. For example:

- One program administrator remarked that the program "makes people feel like they have some kind of team that's working with them and someone looking after them," noting that this helps participants "manage some of those [complex medical] issues a little better."
- Another respondent corroborated this viewpoint, saying that participants see their care manager as "a person who cares and show[s] up and is taking a vested interest in my wellness," which influences the participant to "follow through with this doctor's appointment and pick up the phone when this person calls or return messages."
- Several respondents mentioned that participants appear to be missing fewer appointments as a result of this engagement.

Program leaders commented that although they think the program has positively affected participants thus far, they expect the program's biggest effects on service use, health, and quality of life to emerge in the longer term. One program leader with experience implementing a similar model remarked that this program has not "had the run time yet to fully realize the value of the intervention," due to the slow start in enrolling participants. This leader expected the program to have a more substantial impact over a longer period of time.

D. Context

1. To what extent did organizational features support or conflict with implementation? (RQ 104)

Most respondents credited MMC's institutional support and leadership as important facilitators of the program's implementation. For example:

- One respondent described MMC and its partner organizations as "very progressive and inclusive," noting that program leaders "have their ears to the ground" for opportunities to improve care.
- Another respondent described MMC as a leader in collaborating with a variety of stakeholders in the community. This makes the organization well positioned to "understand the unique makeup of the treatment world … and all the different entities that share a stake in many of the lives of the people we serve."

- Several administrative and frontline staff described MMC as an institution where leaders have long fostered a patient-focused culture. One frontline staff member referred to this culture as "truly patient-centered" and observed that "people here have always believed in working together to try and help the patient."
- Program administrators also noted the importance of having "enlightened leadership at the [hospital] level that's willing to take a risk" and who were "willing to attempt a program like this, understanding that there may be short-term losses but giving us an opportunity to experiment with new systems for a long-term gain."

In addition, several administrative and frontline respondents from MMC and its partner organizations praised the program's collaborative governance structure and focus on continuous improvement:

- As one clinical leader noted, "We've been using different committees to fine-tune [the program], to get feedback and really understand the growing pains of a new program and how to implement changes that will benefit [participants]."
- Another program administrator who participates in the program leadership's various committee meetings said, "When we hear that something isn't working, our response isn't [to say] 'Okay, let's try it for another month.' It's 'Okay, hold it. Let's stop what we're doing and figure out where the bottlenecks are and try to make it better.'"
- Program staff also reported that program leaders are uncommonly receptive to input and feedback from partners and program staff. For example, one respondent from a partner organization described the program as "a unique experience to work with this network [of organizations] and have an ongoing dialogue about what we can do better for patients."

Some respondents noted that the experience of some care management organizations initially posed a barrier to program implementation for several reasons. For example:

- Many frontline and supervisory staff came to the program from case management positions, which often involved limited interaction with clinical providers and different expectations for the services to be provided. As one staff person put it, the program made it clear that "we're really responsible for health, whereas back in [the case management] days, it was like we did everything for the clients."
- Transitioning to this new role required a shift in mindset. As one program leader explained, MMC had to "get people to focus more on performance and [participant] outcomes, as opposed to just seeing [participants] in order to check off the boxes."
- Program leaders helped to support this shift by hosting biweekly case conferences during which care managers received advice and feedback from the program's clinical leaders. There was also a training series in coordinating care that was provided by the program's training partner (SEIU 1199 Training and Employment Fund).

2. To what extent did the policy and political environment support or conflict with implementation? (RQ 106)

Program leaders noted that changes in the state policy environment supported program implementation. New York State implemented its health homes program around the same time that MMC received HCIA funding; as a result, the health homes program provided the foundation and impetus for HCIA-funded activities. Program leaders noted that although the health homes program only funded care management services, its goal of coordinating care management with clinical care aligned with MMC's plans for its HCIA funding. As a result, MMC and its partner organizations reported that their program evolved alongside the state health home program over the past three years. Several respondents noted that the state sometimes cites MMC's program as an example of the direction in which the state hopes other health homes will move. For example, New York State recently adopted MMC's outreach protocol as the standard for all health homes within the state. Program leaders expected state policy to support continued program implementation after HCIA funding ends, citing opportunities such as the state's participation in the federal Medicaid Delivery System Reform Incentive Payment program.

Nearly all respondents mentioned the lack of available, affordable housing in New York City as a major barrier to program implementation. This issue had multiple effects on program staff and participants:

- Outreach and care management staff reported spending a great deal of time trying to find and follow up with participants who do not have stable housing and are constantly moving from place to place (for example, homeless shelters, friends' and families' couches). One respondent noted, "It's not uncommon for [participants] to say, "We can talk about my doctor's appointments and my medicine after you find me a place to live. Until you do that, I don't really want to talk to you."
- Respondents reported that in addition to the housing shortage, participants often struggled with the complex housing application process and landlords' reluctance to accept individuals with SMI into housing units.

Respondents mentioned several ways in which New York State attempted to address this challenge over the award period. For example:

- One administrator from a partner organization noted appreciatively, "New York State has been very active in recognition of the impact of social determinants, and [the state] is actually allocating federal [and] state dollars towards the development of more supportive housing and other programs to help address those social determinant issues."
- Other respondents remarked that New York State is working to tie eligibility for supportive housing services to enrollment in the health homes program in hopes of prioritizing housing access for health home participants.

Respondents, nevertheless, overwhelmingly agreed that finding affordable and available housing for participants is an ongoing challenge.

E. Summary and conclusions

MMC and its partner organizations generally implemented the program on time and as expected; however, they adapted some aspects of the program to address challenges and enhance the program's effectiveness. For example, after realizing that it was not feasible to engage all the providers of every participant, leaders streamlined the program's approach to engage a more limited network of preferred providers. Program leaders also adapted their vision for the care coordination platform in response to providers' feedback about their limited use of this technology.

Most respondents thought the program had improved participants' health outcomes and quality of life. Program staff believed the program reduced the number of hospitalizations and improved the social determinants that may affect participant outcomes. However, nearly all respondents cited housing as a critical and persistent need, and the lack of stable housing is a barrier to program implementation among some participants.

Respondents also indicated that MMC was unique in its ability to convene multiple partners and stakeholders and in its ability to seek out opportunities to improve care. Program leaders created a robust, inclusive governance structure to solicit feedback from administrative and frontline staff at the partner organizations. MMC and its partner organizations described the program as closely aligned to the policy direction in New York State, and they expect state initiatives to continue to support the program's activities beyond the award period.

The preliminary quantitative findings for Medicare enrollees in MMC's health home program are inconclusive. The results of one-year impact analyses on three core outcome variables showed no statistically significant results and may be due to chance. It is possible that the program will generate longer-term impacts that will become apparent in future analyses.

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

XI. VALUEOPTIONS

A. Introduction

Staff with the Massachusetts Behavioral Health Partnership (MBHP), a company owned by ValueOptions that contracts with the Commonwealth of Massachusetts to manage behavioral health benefits for Medicaid recipients, used HCIA funding to test the effectiveness of recovery support navigators (RSNs) and patient incentives to reduce costs associated with the repeated use of detoxification services. Brandeis University's Institute for Behavioral Health served as MBHP's research partner in this work. RSNs worked with individuals who had been admitted two or more times to detoxification facilities to ensure their access to medically necessary services and to improve their engagement with the community and the integration of their care. Program staff also offered financial incentives (gift cards) to a subset of participants to reinforce target behaviors during the first few months following enrollment in the program.[39]

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. At discharge from these facilities, individuals offered an opportunity to enroll in the program and those referred to here as direct participants (or simply participants) were assigned to one of two groups:

- **RSN+I.** In this group, participants were offered RSN support plus incentive payments (described below).
- **RSN only.** In this group, participants were offered only RSN support.

MBHP staff assigned eligible individuals to these groups at the facility level using a midpoint, crossover design[40] 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; then, 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, and eligible individuals from the other two facilities were assigned to the RSN+I group.

MBHP has nine other detoxification facilities in its system. All of these facilities provide a service called the community support program (CSP) that is very similar to the RSN service.[41] There were two key differences between the RSN and CSP programs: (1) RSNs were trained on evidence-based treatment for substance use disorders, readiness-to-change assessments, and

[39] ValueOptions received a no-cost extension from CMMI through 12/31/2015 for continuation of evaluation activities only.

[40] Individuals were eligible to be enrolled as participants (also called the intervention group in the review of quantitative results) or as members of the comparison group if they were between the ages of 21 and 64, enrolled in Medicaid, and had been admitted two or more times to detoxification facilities in the year leading up to the discharge at the time of enrollment, with that discharge included as one of the required two times.

[41] Throughout this chapter, providers in the CSP are referred to as CSPs and RSN providers are referred to as RSNs.

motivational interviewing, whereas CSPs were not; and (2) the RSN payment model was a case-rate model, and the CSP program used a fee-for-service model. Consequently, the project team decided to construct a comparison group composed of individuals discharged from seven of the facilities that offer the CSP. MBHP expected the program, in the short term, to improve its direct participants' social support and their attitudes toward recovery, and in the long term, to reduce their addictive behaviors, enhance overall health, and improve their patient experience relative to the comparison group. The total enrollment goal was 2,300 individuals, including individuals assigned to the comparison group.

The findings in this report are based on quantitative and qualitative data collected or received by February 1, 2015, and June 1, 2015, respectively, and enrollment data reported throughout the award period. Data sources included:

- MBHP-provided program enrollment data, including information on the date of enrollment and demographics, for participants and members of the comparison group who were discharged from a detoxification program and then enrolled in the program or assigned to the comparison group between March 2013 (program start date) and December 2013.
- Data on baseline assessments of substance use and health status among participants and comparison group members between March 2013 and December 2013 provided by MBHP.
- Data provided by MBHP on Medicaid eligibility; medical, pharmacy, and dental claims; and behavioral health encounters between March 2012 and September 2014.[42]
- Qualitative data, including telephone interviews and an in-person site visit in April 2015. Mathematica conducted in-depth interviews with key awardee staff, 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 comparison sites. Data from the participants at the RSN and comparison sites are not part of this report; instead, we provide an in-depth analysis of these focus groups in our seventh quarterly report.

The questions we address in this chapter were selected in consultation with our COR. Specifically, we examine questions related to implementation effectiveness in Section B, program effectiveness in Section C, and the program and policy context in Section D. We close with an overall summary and conclusion (Section E).[43]

[42] MBHP provided these data through December 2014, and the analyses included data on services provided through September 2014 to allow for a three-month runout period for receipt and processing of billed claims.

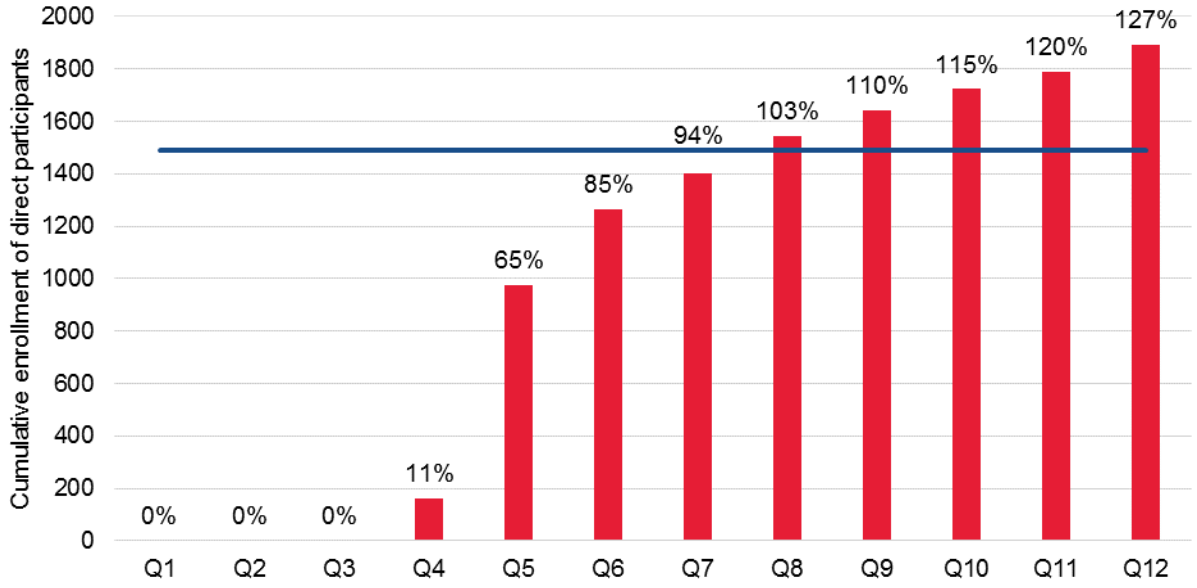
[43] ValueOptions may conduct analyses of data that they have collected on program participants and publish their own reports about program outcomes. These analyses are likely to use different data, focus on different outcome measures, and address different specific questions compared with the data and measures we use in this report to address questions of direct interest to CMMI. As a result, ValueOptions may reach conclusions about program effects that differ substantially from the ones we present in this and future reports.

B. Implementation effectiveness

1. Who was enrolled in the program?

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 by over 25 percent (Figure XI.1). In the first three quarters, MBHP started up and piloted the project. Enrollment began in the fourth quarter with 162 direct participants and increased sharply to include 976 direct participants in quarter 5. In quarters 6 through 12, cumulative enrollment of direct participants steadily increased each month, with MBHP reaching its target enrollment in quarter 8.

Figure XI.1. Percent of target enrollment achieved by quarter, Q1–Q12—ValueOptions



Source: Awardee's self-reported enrollment data, Lewin quarterly reports.

Note: As displayed by the horizontal bar, ValueOptions's target enrollment was 1,492 unique participants.

2. To what extent was implementation timely, conducted as planned, and responsive to site-level constraints? (RQ 17)

MBHP was largely successful in implementing the program as planned and on time, although enrollment lagged at first until the site changed both the enrollment criteria and the definition of a participant, as described in B.3. The awardee made several changes to the original program protocol based on its experiences implementing the program over the course of the award period. These changes included:

- **Modifications to the RSN requirements.** After the first year of the program, MBHP modified the requirements for the RSN role—which originally was to have been a peer support position that required personal experience with substance use and recovery. Most RSNs were hired from the pool of CSPs at detoxification facilities. As the CSP position does not require personal experience with substance use and recovery, MBHP decided not to require this experience either.

- **Revisions to the consent form for study participants.** MBHP determined that the original consent document did not include the necessary language about data sharing between MBHP and CMS. Project staff began using a revised version of the consent document and sought consent from 401 individuals who had already signed the previous consent document. Although the project team had already collected data from these 401 individuals, as of the end of Quarter 11, only 137 had been re-contacted to provide their consent to the revised document.[44]
- **Modifications to follow-up interview procedures.** The program sought to conduct follow-up interviews with participants between five and eight months after enrollment into the study; these interviews included questions about their substance use, functional health and well-being, and experience with care. MBHP had reported in the eighth quarter that the rate of follow-up interview completion was 19 percent. The goal was 60 percent. The process of follow-up was time-consuming and burdensome for the RSNs, who were not paid for the significant hours of extra time they spent on the follow-up data collection. To address this shortfall, MBHP hired a part-time research assistant to help with the follow-up. This position was essential to tracking down members for follow-up, which was made more difficult than expected by the transient nature of the participant population. MBHP also increased participants' compensation for follow-up interviews from $15 to $30 to encourage participants to complete the interview; conducted interviews by phone when in-person interviews were not possible; reimbursed participants for travel to the interview location; provided $5 gift cards to participants for updating their contact information; and offered to provide participants with pre-paid phone minutes for interview time. In their 11th quarterly report, MBHP indicated the follow-up rate for members had increased from 19 to 24 percent.
- **Implementation of learning collaboratives.** In April 2014, MBHP implemented learning collaboratives for the RSN staff. MBHP and Brandeis staff noted that the learning collaboratives provided an opportunity for RSNs to network and share best practices in a face-to-face setting. They indicated that these opportunities were more useful than monthly conference calls; they said they would have incorporated the learning collaboratives earlier had they known how helpful the collaboratives would be.

Respondents thought the implementation of the program was generally responsive to site-level constraints. The program's adaptability was especially important to its success given the diversity of the sites involved. Although MBHP has provided guidance, training, and technical assistance to each detoxification provider, it has also attempted to give sites "room to breathe" when it comes to operational details in delivering services. For example, MBHP encouraged sites to develop their own protocols for distributing the incentive payments to participants. Workforce respondents noted that MBHP solicited their input when making changes to program procedures and protocols, which fostered buy-in from staff. In addition, some sites made adjustments to innovation protocols to account for internal guidelines or policies. For example, one site that typically discouraged clinical staff from touching the participants decided not to abide by the training guidelines that encouraged RSNs to touch participants to show support and compassion.

[44] Mathematica did not receive data from ValueOptions for the 264 individuals who did not sign the revised consent document for the project.

Finally, there were issues with distributing incentives to people at detoxification facilities and other facilities where participants lived (including rules for at least one facility that prevented distribution of incentives on-site), so incentives were not always distributed to participants immediately following an RSN contact, as indicated in the protocol.

3. How did unexpected events support or conflict with successful implementation of the innovation? (RQ 8)

Effective January 1, 2014, MassHealth implemented a change in the distribution of its members among the state's managed care organizations. The redistribution resulted in a decrease in the number of individuals for whom MBHP was the responsible payer. Staff reported that 50 to 60 percent of HCIA participants were no longer MBHP members, and thus lost access to RSN services. This led to a struggle for MBHP to enroll the targeted number of participants into the program. There were staff layoffs at some sites, and RSNs with small RSN caseloads operated as both RSNs and CSPs. RSNs indicated that using both payment models at one time was difficult because it was hard to juggle the two payment systems.

To improve its low enrollment numbers, MBHP took a number of steps. First, it modified the enrollment criteria to allow individuals who initially declined to enter the program to be approached again. MBHP also reversed its earlier decision to exclude from the target population individuals who were civilly committed to the detoxification program. There was initially concern about the involuntary commitment status for these individuals, but MBHP eventually decided they could be enrolled and counted toward the program's target number of direct participants (although assessment data were not collected for these individuals). Finally, MBHP changed the definition of a direct participant to include members who consented to baseline and follow-up interviews, members who did not consent to study interviews but agreed to receive RSN or RSN plus incentive services, and members who did not consent to study interviews but agreed to receive CSP services from a trained RSN.

C. Program effectiveness

In this section, we present impact analyses for two core outcome measures, emergency department (ED) visits and total expenditures, and for two additional outcomes that had high rates of utilization among both the participants and comparison group members: short-term residential treatment stays and days of intensive day treatment. Participants (referred to in reviewing these results as the intervention group) and comparison group members had low rates of hospitalizations in both the pre- and post-HCIA periods; thus we did not analyze impacts on hospitalizations and readmissions.[45] First, we review our data and analytic methods. We then report the findings and limitations of the analysis. Next, we discuss qualitative findings about the perspectives of key informants on program effectiveness.

[45]Behavioral health services are administered by MBHP, a Medicaid managed care plan, and are therefore exempt from the Medicaid Institution for Mental Diseases exclusion. For this reason, we expect to observe all utilization of behavioral health services in the data provided by MBHP, with the exception of any that occurred during periods when the individual was not enrolled in MBHP.

1. Overview of quantitative methods and population

This analysis is based on the following data, all provided by MBHP:

- Program enrollment data on date of enrollment, age, and gender, which were used as control variables in the impact regressions
- Self-reported baseline data on alcohol and drug use, and baseline scores on the Physical and Mental Health Composite Scale derived from the SF-12 survey; we used these baseline data as control variables in the impact regressions
- Medicaid claims for medical, pharmacy, and dental services, as well as encounters for all behavioral health services that were used to construct the four outcome measures for the impact analysis: total expenditures, ED visits, short-term residential treatment stays, and days of intensive day treatment.

Our analysis included data for the intervention group, who received program services, and the comparison group. Individuals who met all eligibility criteria for the program, but who received treatment as usual from the CSP-only comparison sites, formed the comparison group for this analysis.

To ensure the intervention group had sufficient exposure to the program, we limited our analyses to all direct participants and comparison group members who were continuously enrolled in MBHP for nine months after they enrolled in the program. As a result of this requirement, the analyses represent only about one-quarter of the direct participants included in data received from MBHP as of February 2015. We examined impacts for this subgroup for the first year of program participation. Because participants enroll in the program at different times, resulting in different post-program periods and because Medicaid claims and encounter data are included in the analyses through September 2014, the analyses include members of the intervention and comparison groups who enrolled between March 2013 and December 2013.

Because analyses are based on only the first 12 months of the 33-month program for about a quarter of the participants, trends in outcomes that are not evident now may become apparent as we conduct more analyses of Medicaid data in the future. In light of the preliminary nature of our findings, it is important to avoid drawing premature conclusions about the effectiveness of this program.

Detailed information on the methodology is available in Appendix A.

2. Quantitative findings on program effectiveness

We estimated impacts using a difference-in-difference approach, which compares the change in outcomes between the baseline and the program period for the intervention group relative to the comparison group. Overall, the results indicate a statistically significant negative impact on ED visits in the first year of the program. (Table IX.1).[46] Impacts on expenditures, days spent in intensive day treatment, and short-term residential treatment stays were not

[46] The pre-program period includes the year before date of enrollment in the program.

statistically significant. These analyses have important limitations and should be interpreted with caution.

- In the first year of the program, we found substantial, statistically significant decreases in ED visits for members of the intervention group relative to the results for comparison group members. The estimated impact on ED visits was two visits per person (27 percent) and it was statistically significant (p-value=0.02).
- Expenditures in the first year of the program increased for both groups, and the difference in the increase between the two groups was not statistically significant (p-value= 0.14). The increase for the intervention group was an estimated $3,516 (or 12 percent) lower than the increase for comparison group members, but the lack of statistical significance means that we cannot rule out that this difference occurred by chance.[47]
- The use of intensive day treatment in the first year of the program increased for both participants and comparison group members, but the difference in the increase between the two groups was not statistically significant (p-value=0.12) and could be due to chance.
- We found no effect on short-term residential treatment stays in the first year of the program.

Table XI.1. Impact estimates for ValueOptions, annual results

	Regression-adjusted mean for the intervention group	Regression-adjusted mean for the comparison group	Estimated impact[b]		
			Value	Percent	P-value
Total expenditures per patient					
Baseline year	$21,867	$19,984	N.A.	N.A.	N.A.
IY1	$25,513	$27,146	$-3,516	-12	0.14
Short-term residential treatment stays per patient					
Baseline year	3.0	2.7	N.A.	N.A.	N.A.
IY1	4.1	3.6	0.3	7	0.5
Days of intensive day treatment per patient					
Baseline year	4.2	5.4	N.A.	N.A.	N.A.
IY1	8.0	6.2	3.0	59	0.12
ED visits per patient					
Baseline year	6.7	5.7	N.A.	N.A.	N.A.
IY1	5.5	6.5	-2.1**	-27	0.02
Number of observations	478	210	N.A.	N.A.	N.A.
Number of unique patients[a]	239	105	N.A.	N.A.	N.A.

Source: Mathematica analysis of MBHP-provided program enrollment data, Medicaid administrative data, and baseline assessment data covering the period from March 2012 through September 2014.

[a] To be included in the analysis, members of both the intervention and comparison group had to be enrolled for at least nine continuous months in MBHP.

[b] We derived the impact estimates in Stata using the lincom command to compare the difference between the intervention and baseline period means for the treatment and comparison groups. All regression models control for age, gender, calendar month of enrollment, baseline Physical and Mental Health Composite Scale scores derived from SF-12

[47] Results are based on a GLM model to parallel the observed distribution of expenditures in the analysis population and reduce sensitivity to outliers. Sensitivity testing indicates an OLS model would find the same direction of the effect and confidence level.

(provided by MBHP), and baseline alcohol and drug use severity indices (constructed by Mathematica using compositing methodology available at: http://www.tresearch.org/wp-content/uploads/2012/09/CompositeManual.pdf).

** Impact estimates are statistically different at the .05 level, two-tailed test.

IY1 = Intervention Year 1; N.A. = Not applicable.

The results presented in this section are preliminary and have important limitations:

- These findings reflect the period from March 2013 to September 2014 and are limited to the first year of the 33-month program. Due to ValueOptions' receipt of a no-cost extension, HCIA funding for the program will continue until December 31, 2015. The sample size for this analysis is insufficient to detect small changes in outcomes. As more participants enroll and more data become available for the full period of the program, findings may change.
- In addition, these findings are based on a small proportion of the participants—those with at least nine continuous months of coverage by MBHP after the start of the program. Therefore, findings are not generalizable to all program participants, such as those with shorter or discontinuous enrollments in MBHP. Further, the members of the intervention and comparison groups who are included in the analysis do not all have a full 24 months of coverage by MBHP corresponding to the full analysis period. Estimates for those individuals with less than 24 months of enrollment were pro-rated to reflect a full period of enrollment, and analysis weights were applied to adjust for the proportion of the analysis period during which the individual was observed. However, enrollment is nearly full for both the intervention and comparison group members, implying minimal bias, if any, to the impact estimates.[48]

In light of the preliminary nature of the current findings, it is important to avoid drawing premature conclusions about the implementation or outcomes of ValueOptions' program.

4. Qualitative findings on program effectiveness: From the perspective of key informants, to what extent does the program improve desired health outcomes and quality of life for participants? (RQ 23 and RQ29)

Overall, key informants indicated the RSN program had several positive effects on health outcomes and quality of life for participants, and no negative effects. Positive program effects included improved health and better connections to health providers. For example:

- One RSN explained, "I had a client who was almost dying when we first met him—his HIV was over the roof, liver function was very bad, but by the time we discharged him he was fine and doing so well."
- Another RSN noted, "I have a client who is on Vivitrol [a medication used to prevent relapse after opioid detoxification], which is a huge help … The psychiatrist and mental health [care] has been great. He still has cirrhosis but his liver function has been great. When he came through detox it was completely yellow. Ever since he stopped drinking, his health has been great."

[48] Three-fourths of the intervention and comparison group members were enrolled for nearly 90 percent of the 24-month analysis period.

- Key informants also indicated that the training RSNs received on motivational interviewing may have been particularly helpful in improving participant outcomes, as it helped the workforce to better understand their roles and more actively help participants achieve their recovery goals. As one respondent 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 also described how the program helped to improve participants' quality of life. In many cases, RSNs provided participants with reassurance that they were not alone in the recovery process; or the RSNs were an extra source of support as participants worked toward their goals. In other instances, RSNs helped participants with non-medical needs, such as obtaining public benefits or navigating the justice system. For example, one RSN helped a participant who was living in a halfway house acquire visitation with her child.

While respondents were very enthusiastic about the RSN program, they cautioned that positive outcomes were not always possible due to the nature of the participant population. For example:

- One respondent indicated that outcomes were largely dependent on the motivation level of each participant, saying "If the person is willing, they are going to do it. The ones that are willing to make their lives better and utilize services take full advantage of this. The ones that are not willing, it doesn't matter. This program is very good for people who are willing to do some footwork."
- Other RSNs explained that while some participants do not achieve a full recovery, the program may still help them to improve their lives: "I think the people who actually get engaged and get connected with all this other stuff, they do very well. Whether or not they stay sober is not always the thing, but the fact that they got that far."
- Another respondent indicated that substance use treatment facilities sometimes stigmatize their patients, and that RSNs can provide a respite from this by connecting participants to care in a less judgmental way.

Respondents' assessments of the usefulness of the financial incentive and payment reform components of the program were mixed. Many respondents thought the incentives did not benefit participants who would not have recovered otherwise. However, some described the structured nature of the incentive program as potentially beneficial for some participants. For example:

- One RSN explained, "I think the only thing that is more beneficial … is the treatment plan for RSNs is a little more structured for us and I like the fact that … the treatment goals … are right there in front of them … this is where the incentive plays in a positive way."
- Respondents from MBHP also noted that, although literature on financial incentives describes positive effects for short-term outcomes, there has been less research on the effects of incentives on longer-term goals, like substance use recovery.
- Respondents indicated that payment reform provided an opportunity to provide participants with more services and to sustain the services over time, which can improve participant

outcomes. Payment reform also allowed RSNs to account for time spent on training and other professional development activities.

D. Context

1. To what extent did organizational features support or conflict with implementation? (RQ 104)

Respondents indicated that MBHP, Brandeis, and the provider sites worked together well, and that this cohesion facilitated program implementation. As one respondent explained, "The teamwork has been amazing. From our [CMS] project officer … to the RSN workforce and MBHP and Brandeis in the middle … It's been such a nice group to work with. We've had to change things along the way. Nobody yells at you. They understand. Everyone's working hard together." Providers 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.

Although the strong relationship between MBHP and the providers facilitated program implementation, MBHP's role as an insurance payer and its consequent lack of control over staffing at the treatment sites presented challenges. MBHP respondents believed the program could have served more participants if more RSNs were on staff. However, individual 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. While the sites provided similar services, they had different management styles and organizational structures, which made standardizing services for research purposes difficult. As one respondent explained, "there's a limit to how much you can control what they do and standardize what they do … for research, that's really hard." However, working with a large number of providers also had positive effects. One respondent indicated that the high visibility of the RSN study increased attention on the RSN/CSP program, even among the comparison sites, and renewed overall interest in the program among detoxification facility clients.

The characteristics of each provider site also supported program implementation. For example, in each site, RSNs largely worked collaboratively, which helped them to provide services and boosted morale. As one respondent explained, "it's almost like we're a little family, and we encourage. We build up. We support …We help each other." Some RSNs reported that having a detoxification facility within their organization helped with recruitment by allowing easy access to and interaction with patients and staff at the facility.

2. To what extent did the policy and political environment support or conflict with implementation? (RQ 106)

Several state-specific policies in Massachusetts affected implementation of this project. Importantly, Massachusetts Medicaid pays for the RSN service, eliminating any need for new funding. However, because the state does not have a formal accreditation program for recovery coaches, RSN training was paid for by program funds.

The substance abuse treatment culture in Massachusetts also affected program implementation. In Massachusetts, substance abuse treatment is largely focused on inpatient detoxification facilities. As a result, this program recruited potential RSN participants while they were recovering in these facilities. One respondent noted this may have hampered recruitment,

because of the challenges individuals face in such an early stage of recovery. The respondent explained that recruiting participants who were further along in the recovery process may have improved the program, saying, "I think engagement would be better. I think that the uptake might be better. I think that you would have fewer people who are agreeing to services and then immediately changing their mind because they are thinking more clearly when they're approached for services." Respondents also noted that residential recovery homes[49] were often resistant to giving outside service providers access to their residents, presenting a challenge for RSNs who served participants in group homes.

E. Summary and conclusions

MBHP implemented the RSN program relatively smoothly, effectively working with its research partner, Brandeis, and a large number of partner sites to provide the innovation components to participants. The company adjusted well to an unanticipated decrease in its potential study population and was able to achieve its target enrollment. It also adapted to challenges by implementing changes when needs arose, particularly by hiring a research assistant to help with participant follow-up and instituting a learning collaborative for the workforce. Respondents were enthusiastic about the program and perceived that the program had a positive impact on the services received by participants.

Overall, the preliminary quantitative findings indicate a large and statistically significant reduction in ED use. Impacts on other outcomes were not statistically significant.

The results presented in this chapter are preliminary; they are based on only the first 12 months of the 33-month program and on approximately one-quarter of program participants. Therefore, it is important to avoid drawing premature conclusions regarding the implementation or outcomes of ValueOptions' program based on the preliminary findings presented here.

[49] A residential recovery home is one type of residential treatment facility. These facilities provide longer-term services compared to detoxification facilities, which are also classified as residential treatment facilities by MassHealth and are not a covered benefit paid for by MassHealth.

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

XII. VINFEN

A. Introduction

Vinfen, a community-based behavioral health service provider, is using HCIA funds to implement a program offering a behavioral health home intervention that integrates primary health care with behavioral health care for adults in the metropolitan Boston area with serious mental illness. This project redesigns care processes and introduces self-management techniques with a goal of reducing use of medications and acute care services, thereby reducing expenditures.

Vinfen's program has three key components:

1. Integration of a nurse practitioner into four existing psychiatric rehabilitation outreach teams (one Vinfen team and one team for each of its three provider partners) to coordinate care, manage clinical care, and provide primary care. On each team, a nurse practitioner partners with a health outreach worker (HOW) to provide outreach, self-management training, and other health interventions.
2. The Health Buddy telehealth system, a small device used by a subset of participants that allows the nurse practitioner and HOW to monitor participants who need ongoing medical attention.
3. The Integrated Illness Management and Recovery (IIMR) curriculum, which the HOWs use to train participants in behavioral strategies that will improve their health and help them manage their illness.

Vinfen is implementing the program in partnership with three other community-based behavioral health providers: Bay Cove Human Services, North Suffolk Mental Health, and Brookline Mental Health. At each of the four organizations, one existing psychiatric rehabilitation outreach team was selected to implement the program; a nurse practitioner was added to each of the four teams. The Commonwealth Care Alliance (CCA), a nonprofit Medicaid managed care entity, provides and supports the four nurse practitioners assigned to the program teams. For the HCIA program, the four program teams serve participants who were already assigned to the psychiatric rehabilitation outreach teams before the start of the HCIA award as well as participants assigned to the teams after the award began. These teams were selected to deliver the program services because they were serving individuals perceived to have the most complex health conditions. Although all participants served by these teams have access to the program services, evaluation data are available only for participants who provide written consent. Vinfen has identified four teams to serve as a comparison group, but no quantitative data on the members of the comparison group are available for our evaluation.

Vinfen is also partnering with research staff from Dartmouth College, which developed the IIMR curriculum, to train and support the HOWs in using the curriculum and to conduct health outcomes analyses. Bosch Healthcare provides the Health Buddy telehealth system, as well as ongoing technical assistance and support. JEN Associates is partnering with Vinfen to conduct a cost analysis based on Medicaid and Medicare claims data.

The four program teams served 470 individuals in total; of these, Vinfen anticipated that about 85 percent, or 400 individuals, would engage with at least one component of the intervention. By the end of the 12th program quarter, 216 participants had used the program's services.

The findings in this report are based on qualitative data collected or received by June 1, 2015, and enrollment data reported throughout the award period. Data sources included:

- Enrollment data submitted by Vinfen to the reporting website maintained by the Lewin Group for the HCIA initiative
- Qualitative data, including phone interviews and an in-person site visit in March 2015. Mathematica conducted in-depth interviews with key awardee staff, members of the workforce, and other stakeholders. In addition, Mathematica convened focus groups with members of the workforce and with groups of program participants and non-participants. Data from the latter groups are not part of this report; instead, we provide an in-depth analysis of the participant and non-participant focus groups in our seventh quarterly report.

The time line for receiving quantitative data from Vinfen remains uncertain, and the data Vinfen may eventually provide will not support rigorous impact analyses.[50] As a result, we are not able to report at this time on the four core measures (total expenditures, hospitalizations, hospital readmissions, and emergency department visits), and any future presentations will be limited to aggregate data provided by Vinfen for the program group only, possibly supplemented by limited analyses of self-reported data on hospitalizations for the program group.

The research questions we address in this report were selected in consultation with our COR. Specifically, we examine questions on implementation effectiveness in Section B, on program effectiveness in Section C, and on the program and policy context in Section D. We close with an overall summary and conclusion (Section E).[51]

B. Implementation effectiveness

1. Who was enrolled in the program?

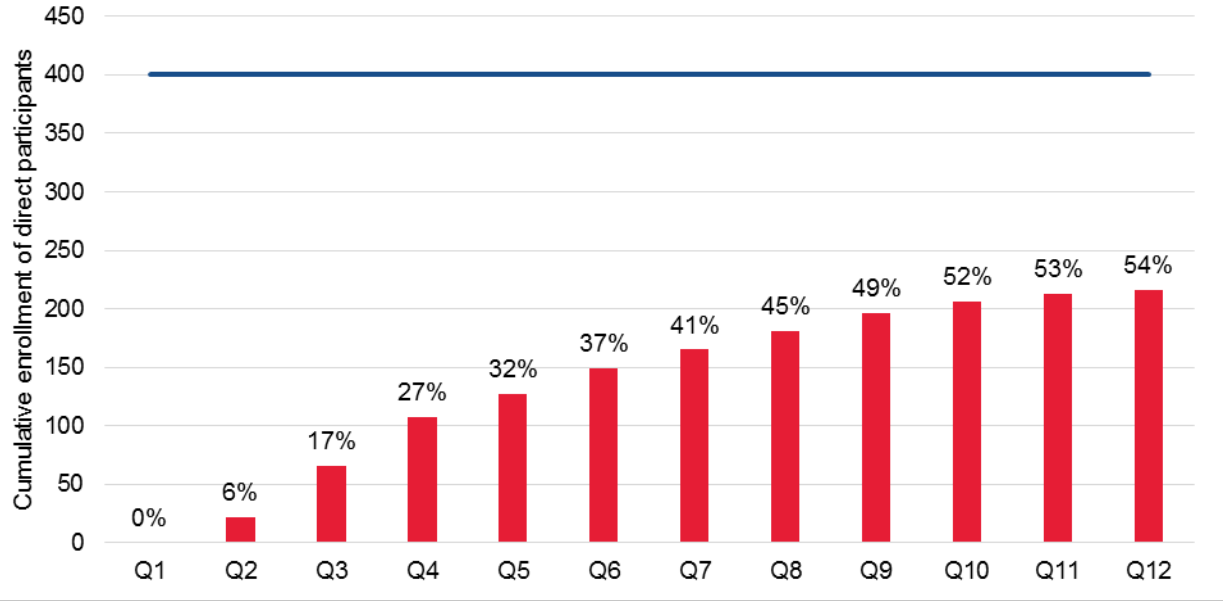
By the end of the 12th quarter (June 30, 2015), Vinfen and its partners had provided program services to 216 direct participants, about 54 percent of its enrollment target (Figure

[50] During production of this report, MassHealth, which oversees the Massachusetts Medicaid program, approved Vinfen's request to access participant-level Medicaid data, but Vinfen will not release person-level Medicaid data to Mathematica for this study. Instead, Vinfen will give Mathematica aggregated, unadjusted data on the four CMMI core outcomes only for Medicaid-enrolled participants who are part of the program group *and* provide consent. In addition, Mathematica anticipates that Vinfen will provide aggregated, unadjusted Medicare data on the four core outcomes for consenting participants assigned to the program group who are also Medicare enrollees. When these data will be provided to Mathematica is unclear. The third source of data for this awardee, project assessment data, includes participant-level demographic information, self-reported hospitalizations, and data on the health of all consenting program group participants; Vinfen provided Mathematica with these data in July 2015.

[51] Vinfen may publish their own reports about program outcomes. These reports are likely to use different data, focus on different outcome measures, and address different specific questions compared with the data and measures we use in this report to address questions of direct interest to CMMI. As a result, Vinfen may reach conclusions about program effects that differ substantially from the ones we present in this and future reports.

XII.1). Although program services were available to all of the approximately 470 participants served by the four program teams, not all were expected to use these services; Vinfen estimated that approximately 85 percent of the 470 participants served by the program teams, or roughly 400 participants, would engage with at least one of the intervention components. Enrollment began in the second quarter with 22 direct participants and increased sharply to a total of 66 direct participants in quarter 3 and 107 direct participants in quarter 4. In quarters 5 through 12, cumulative enrollment of direct participants steadily increased each month.

Figure XII.1. Percent of target enrollment achieved by quarter, Q1–Q12--Vinfen



Source: Quarterly reports submitted to the Lewin Group's reporting website for the HCIA initiative.
Note: As displayed by the horizontal bar, Vinfen's target enrollment was 400 unique participants.

2. To what extent was implementation timely, conducted as planned, and responsive to site-level constraints? (RQ 17)

Vinfen was largely successful in implementing the program as planned and did so in a timely manner. However, the awardee made several changes to the original protocol based on feedback from staff and participants and on lessons learned in the early phases of implementation:

- **Revised outreach team structure and roles.** Vinfen redefined roles to help staff work more efficiently with the Health Buddy system, and modified its outreach team structure for some teams.
  - As noted in the first annual report, initially both the HOW and the nurse practitioner were tasked with monitoring Health Buddy data. Because this approach created confusion about roles and duplication of efforts, Vinfen staff worked with Bosch to review sample cases, streamline the Health Buddy work flow, and assign distinct roles to each position: HOWs reviewed Health Buddy responses daily and notified the nurse practitioner about high clinical values.

  - In the second year of the project, Vinfen also 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 within the last six months of the project; instead, for some outreach teams, Vinfen used RNs hired by CCA. 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-need populations—that is, those more likely to live in a group home and to lack consistent access to primary care services.
  - Although these changes demanded flexibility from both HOWs and primary care staff, key respondents from the workforce agreed that the modified team structure generally worked well and supported CCA's internal model, in which a small number of nurse practitioners support a larger pool of RNs.

- **Introduced mobile and web-based Health Buddy options.** In the third year of the project, Vinfen launched mobile and web-based versions of the Health Buddy operating system.
  - The web-based system included "brain games" to help participants maintain and improve their cognitive skills, as well as links to relevant reading materials (including guidance on medication management, healthy recipes, and exercise tips), which were customized to each participant's condition.
  - Vinfen pre-loaded the mobile version of the system on a basic cell phone provided to participants. Vinfen also pre-loaded phones with five contacts selected by the HOW and the participant (for example, the participant's HOW, primary care doctor, psychiatrist, a close family member, and an emergency service). Vinfen programmed the phones to allow calls only to these five contacts.
    - Key respondents reported that the mobile Health Buddy system was especially useful for working with participants who were transient.
    - One HOW noted that the mobile Health Buddy system was the only way she could stay in touch with participants who were homeless: "It's been great because we have a couple of people [who] are homeless so it helps us stay in contact with them … and then they do their Health Buddy session right on the phone."
    - Key respondents perceived that participants were more receptive to the mobile Health Buddy than the web version of the system. Respondents suggested that this may in part be because members of this population may lack consistent access to a computer and in some cases have limited computer literacy.

- **Updated IIMR protocol.** In the second year of the project, Vinfen made changes to the protocol for administering the IIMR curriculum in response to feedback from workforce members.
  - The revised curriculum's modules included more repetition of concepts and more visuals and could be reviewed in a nonlinear fashion as long as the basic outline was covered.

  - One key respondent explained that these changes were made based on input from staff and participants: "Because our client population needed things to be much more streamlined, we needed better visuals [and] much more repetition."
  - To ensure consistent provision of services, Vinfen also worked closely with one of the provider sites that implemented the IIMR curriculum in a group setting. All other sites implement the curriculum in an individual setting.

- **Modified service delivery strategy.** As discussed in the first annual report, Vinfen relaxed its service delivery strategy in January 2013 with the hope of increasing participant engagement.
  - Under the new strategy, participants had to use at least one of the three innovation components (IIMR curriculum, Health Buddy system, or nurse practitioner services), but no longer had to use all three.
  - Key respondents explained that the modified strategy allowed the workforce to gradually introduce program components once they had established a relationship with a participant.
  - Most participants used one or two program components, with a smaller subset eventually using all three components. Some participants also stopped and started use of program components based on progress toward goals or other events in their lives.

3. How did unexpected events support or conflict with successful implementation of the innovation? (RQ 8)

As noted in the first annual report, the October 2013 implementation of Massachusetts's demonstration for dual Medicare-Medicaid enrollees, known as One Care, influenced several aspects of the service delivery model for CCA, including staffing, procedures for providing care, and training requirements.[52] In turn, these changes limited the capacity of HCIA-funded nurse practitioners, who were employed by CCA. In the sixth quarter, Vinfen offered training for HCIA-funded and non-HCIA-funded staff at each of the provider sites to explain One Care provider options and new services introduced under the program, and to clarify the role of integrated care and care coordination.

Key respondents reported that CCA's shift to the One Care program made retaining nurse practitioners more difficult. In part, this was because implementation of the One Care program resulted in a large demand for nurse practitioners within CCA. Vinfen also had to replace nurse practitioners who left the project, and (as noted above) it assigned RNs to some outreach teams in their place. Although outreach team members reported that the revised team structure worked well, these changes required increased flexibility from all team members.

Reductions in staff at CCA also limited the support and supervision given to nurse practitioners. Respondents emphasized the importance of having proper support for nurse

[52] 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.

practitioners in the outreach team setting: "[We need to] have clear direction and support, especially in the behavioral health field, so that when you're working with particularly challenging participants they're getting assistance around what kinds of interventions or strategies will work with a particular person."

Vinfen also experienced turnover in HOW staff in the second and third years of the project, primarily due to staff promotions or decisions to return to school. Key respondents at Vinfen perceived that HOW turnover reflected the experience and training that staff gained through work on the program, rather than dissatisfaction with the job itself. As one person put it, "A lot of their training … and the support and supervision we give them has actually made them very marketable as community health workers." One Vinfen respondent noted that the turnover rate among HOWs was comparable to the industry standard for similar positions nationwide.

Staff turnover posed a challenge for the members of the outreach team, in part because the teams were structured so that the HOW and nurse practitioner worked as a pair. A key respondent from Vinfen observed that when a nurse practitioner left one outreach team, the team's relatively new HOW felt "a little bit more lost because she doesn't have her go-to pair." These changes were also disruptive for participants, who had to build a trusting relationship with the new staff assigned to them.

C. Program effectiveness

We were not able to conduct an impact analysis of the four core measures because data were not available. We do, however, report on perceptions of the program's effectiveness from the perspectives of key awardee staff, members of the workforce, and other stakeholders.

1. From the perspective of key informants, to what extent does the program improve desired health outcomes and quality of life for participants? (RQ 23 and RQ 29)

Overall, key respondents reported that participant health and well-being improved as a result of the program. For example:

- Anecdotally, some participants experienced improved clinical indicators such as blood pressure and hemoglobin A1c values. Respondents stressed, however, 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.
- Respondents 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. An outreach team leader provided an example of how a nurse practitioner worked closely with a participant to address several significant health issues: "We have a client who was a long-term alcohol user, who was diagnosed with congestive heart failure, and is a smoker, and did not have a good relationship with her primary care physician. [The nurse practitioner] was able to meet with this client and . . . get her to trust her enough to be weighed … every day (to check for fluid retention) and [get] regular blood work. We were able to find out that she was in acute renal failure … more drinking would have led to her death. With that

information, we were able to try to get a Section 35 [a court order for involuntary commitment] in order to get her treatment to avoid a very negative medical consequence."

- Another outreach team member described how the Health Buddy system sparked conversations about health issues and functioned as a helpful 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."
- Respondents also reported the program helped some participants access care they would not have received otherwise, such as dental services, wellness visits, or routine procedures like 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."

The program also trained participants to use available services more appropriately. According to respondents, Health Buddy data supported anecdotal evidence that the program reduced the number of hospitalizations and emergency department visits for participants. Key respondents perceived that participants became better at recognizing a true health emergency, as distinct from an ongoing symptom. As an outreach team leader explained, participants were better able to explain their symptoms and work with staff to determine an appropriate course of action: "There was someone who was chronically in the ER a couple times a week. She was able to call and say this is what I'm feeling, and the nurse practitioner was able to suggest solutions rather than her going straight to ER or to crisis management."

Respondents also perceived that the program empowered participants to make positive changes in their health behaviors. Anecdotally, participants had fewer missed appointments than before they entered the program, and were better able to communicate with staff about their health. For example, an outreach team member described how the Health Buddy system helped a participant remember to take her medication daily: "One client is taking her medication regularly now because she sees the light blinking on the Health Buddy and she remembers. Some people are happy to know that when they enter information into the Health Buddy, there is someone on the other end paying attention." Outreach team members cautioned, however, that participants' self-reported data did not always reflect health status accurately, particularly among those with chronic conditions: "because people in crisis, [their symptoms are] going to remain constant, and if I reacted upon that, it just would not be beneficial for them or their care."

Finally, workforce staff emphasized that participants were able to lead fuller lives because they were not as consumed with physical or mental health issues. One workforce member described how the program helped participants feel more connected to the community and cared for by program staff: "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."

D. Context

1. To what extent did organizational features support or conflict with implementation? (RQ 104)

Respondents generally viewed Vinfen as supportive of the program and its goals, and felt that the organization's mission and values helped to guide program implementation. Respondents believed buy-in from senior leaders at Vinfen was a particularly important facilitator of the program. In turn, the HCIA award helped Vinfen senior leaders become more open to new ideas and ways of doing things. 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]."

Several Vinfen respondents also stressed that the designation of a formal project manager was essential to program implementation: "A project manager is key; we've actually added that component to a lot of other projects because in the absence of that, there's no one person [in charge]—it doesn't have to be full-time depending upon the length of the project … but at the beginning you have to have somebody who is knowledgeable about initiating and implementing new projects."

Respondents agreed that the program's cross-organizational nature presented significant implementation challenges. For example:

- HOWs were employed by one organization, while nurses were employed by another. Outreach team members did not share a standard electronic medical record system, which made it difficult to share participant information.
- Nurse practitioners perceived that their employer (CCA) did not understand their role on the project and thus did not provide adequate support or supervision.
- 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."

However, because the implementation involved multiple organizations, its impact may ultimately have a broader 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."

2. To what extent did the policy and political environment support or conflict with implementation? (RQ 106)

Key respondents indicated that other ongoing work in Massachusetts supported and may help sustain elements of the awardee's program. As noted in the first annual report, Vinfen initially saw the HCIA innovation as a useful "run-up" to the One Care Health Care program for

dual Medicare-Medicaid enrollees, which launched in October 2013. During the spring 2015 site visit, key respondents from Vinfen reported plans to sustain parts of the innovation through One Care, including provision of integrated physical and mental health services. These plans would allow participants who are enrolled in One Care to continue receiving integrated services through this program, so Vinfen enrolled as many participants as possible in One Care before the end of the HCIA award.

E. Summary and conclusions

Vinfen implemented its behavioral health home program relatively smoothly, effectively working with partner sites to provide the innovation components to participants. The awardee modified several components of the program in response to feedback from staff and participants, including changes to the service delivery structure that required participants to use at least one, rather than all, of the innovation components; clarification of staff roles; introduction of mobile and web-based options for the Health Buddy system; and improvements to the IIMR curriculum. Vinfen also adapted to implementation of One Care, the state's demonstration for dual Medicaid-Medicare enrollees, by modifying the outreach team structure to better reflect the One Care model and by providing training for innovation staff on new health care options and services offered under the program. Although respondents were enthusiastic about the HCIA program and perceived that it had a positive impact on the well-being of participants and the services that they received, we were unable to calculate the four core impact measures for this report because limited data were available.

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

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.

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.

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, Samuel 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.

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.

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.

Office of Management and Budget. "Standards and Guidelines for Statistical Surveys." September 2006. Available at [https://www.whitehouse.gov/sites/default/files/omb/inforeg/statpolicy/standards_stat_surveys.pdf]. Accessed July 17, 2015.

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.

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.

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.

APPENDIX A:

TECHNICAL METHODS

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

A. Description of data sources

The data sources used for the impact analyses vary across awardees. The analyses for ICSI, KMHS, and MMC primarily used CMS Medicare administrative data, but we also used finder files to identify intervention participants for ICSI and MMC. The analysis for ValueOptions used program enrollment data, Medicaid administrative data and survey data provided by MBHP. Lastly, CHCS provided the program enrollment and survey data we used to analyze its program. In this section, we provide a general overview of these data sources.

1. CMS Medicare administrative data

For ICSI, KMHS, and MMC, our analysis of core outcome measures used CMS Medicare administrative data. We obtained data files through the CMS's Virtual Research Data Center (VRDC). We extracted all final action claims with dates of service from January 2009 through September 2014[53] for individuals for whom the HIC or SSN included in the ICSI and MMC finder files matched to a BENE_ID in the VRDC cross-reference files. For KMHS, data were extracted for all individuals with a mental health visit billed by KMHS as described in Section C.2 below. 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 Enrollment Database (EDB). The EDB includes information on date of birth, gender, most recent county of residence, enrollment in Medicare Advantage (MA), and third party insurance coverage.

2. Medicaid claims and enrollment data provided by MBHP

The Medicaid administrative data provided by MBHP were of three types. First, the data included Medicaid fee-for-service claims for medical and pharmacy services. Second, the data also contained encounter claims for behavioral health services provided under the managed behavioral health provider, MBHP. The behavioral health claims listed a paid amount which represents the negotiated payment rate paid to the service provider. Lastly, the Medicaid data included an eligibility segments file with records that listed the beginning and end dates for MBHP enrollment spans for a given Medicaid enrollee. These files consisted of data for all MBHP members who met the study eligibility criteria (i.e., they were from 18–64 years old and had at least two detoxification treatments in the past 12 months).

3. CHCS survey and enrollment data

CHCS provided Mathematica with an enumeration file in November 2014 including a record for all participants and control group members. This file included the person's date of enrollment in the intervention. It also included Medicaid and Medicare program IDs, if applicable. However, many CHCS program participants and control group members were uninsured. The file also included Social Security Number (SSN), date of birth, and gender.

[53] Claims were extracted between February and April 2015 depending on the awardee.

Lastly, the file included indicators of employment status, highest level of education completed, and living situation. All this information was reported as of the enrollment date.

CHCS also provided Mathematica with a survey assessment data file in May 2015 including multiple records for most, but not all, participant and control group members. Specifically, the file included baseline assessment records for all 261 participants and 259 control group members, as well as six-month follow-up records for 214 participants and 208 control group members. The data elements in this file included a unique survey ID and assessment round indicator (e.g., baseline, 6-month follow-up, 12-month follow-up), as well as assessment scores for a variety of mental and physical health indicators.

CHCS has agreed to provide electronic medical record information on service use for participants and control group members in addition to the enumeration and survey files. We have received an initial extract of these data, but decided not to include analysis of these data in the current report because we are still working with CHCS to classify services into more meaningful categories including categories for intervention services. We plan to report findings on service use in future reports when this classification scheme is complete and more data are available.

B. Specifications of measures

We use multiple types of measures in these analyses. If it was possible to calculate the core measures identified by CMS and these measures were appropriate to the intervention, we used them. Our specifications for these measures are described in the first section below. Section B.2 describes other outcome measures used in the impact analyses for different awardees. For ValueOptions, short-term residential treatment stays and days of intensive day treatment were examined instead of inpatient stays and readmissions because these services were more common among intervention participants. For CHCS, because claims data were not available for the participants, survey-based outcome measures were examined. In addition to these other outcome measures, in Section B.2 we also provide specifications for the control variables included in the impact models.

1. Core measures

CMMI has 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 emergency department (ED) visits. These measures were calculated for 12-month periods. For individuals included in our analyses for less than the full analysis period, measures were prorated and regression models were weighted based on the proportion of the analysis period for which the individual was represented. In this section, we provide detail on the data and analytic methods used to develop these core outcome measures for all awardees for which core measure findings are presented in this report (ICSI, KMHS, MMC, and ValueOptions). The data sources available for this analysis varied across awardees; however, we developed specifications to standardize across the data sources to the extent feasible. Similarly, the population analyzed and how the intervention is targeted varies across awardees. Our approach to selecting the population to include in the analyses varied with the characteristics of the intervention as described below.

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 Medicare or Medicaid 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. When an individual is not enrolled in Medicare/Medicaid, if Medicare/Medicaid is not their primary insurance, if they are not covered by Medicare Part B, or if they are enrolled in a managed care plan, their health expenditures and utilization are not consistently observable in the administrative data available for this analysis. Thus, our analysis must be limited to patients and time periods during which sufficient data is available to calculate the core measures. The methods for identifying the patients and periods with observable data varied across awardees. We first describe the approach used for CMS Medicare administrative data. Then, we describe the approach used for ValueOptions Medicaid administrative data.

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. ICSI and MMC each 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. KMHS patients were identified based on having a claim for a mental health service provided at KMHS. More information on identification of these patients is provided in Section C.2.

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 Medicare Advantage, 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 2009 to September 2014 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 through C.3 below for additional exclusion criteria that were applied during the development of the comparison groups for these awardees.

Step 3: Define baseline and intervention periods. Baseline and intervention years were defined for each intervention participant or comparison group member relative to their enrollment date or the intervention start date.[54] The first intervention year was defined as the enrollment or intervention start month and the 11 months following that month. Where applicable the second intervention year was the 12 months following the first intervention year. The first baseline year was the 12 months prior to the enrollment or start month and additional baseline years were identified by moving backward from the first baseline year. For each individual included in the analysis the proportion of each baseline and intervention year 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.

ii. ValueOptions Medicaid administrative data

Like the approach taken for ICSI, KMHS, and MMC in the Medicare data, baseline and intervention years were defined for each intervention participant and comparison group member in ValueOptions' Medicaid data relative to their enrollment date. The first intervention year was counted 359 days forward from the enrollment date and the baseline year was counted 360 days backward from the day before the enrollment date.[55] Within each analysis year, MBHP enrollment spans included in the Medicaid administrative data were identified. These spans varied in length and were defined with begin and end dates. If an individual was enrolled continuously for a minimum of nine months of the first intervention year, they were included in the analysis. Records for individuals who did not meet this criteria were excluded from our analysis. For each individual included in the analysis the proportion of the baseline and intervention year for which the individual was enrolled in MBHP was calculated. This proportion was used to pro-rate expenditures and utilization 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 or Medicaid expenditures, hospitalizations, hospital readmission, and emergency department visits. We summarized each of these measures

[54] Pseudo-enrollment or intervention start dates were defined for comparison group members as described in Section C below.

[55] A year was defined as consisting of 360 days to be consistent with previous analyses where a month was defined as consisting of 30 days and a quarter as consisting of 90 days.

monthly[56] for each individual in the analysis population. Then, we aggregated sets of months for annual analysis.

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 (pmt_amt) 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 admission 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.

For ValueOptions Medicaid data medical (including institutional and outpatient services), pharmacy and behavioral health claim types were included in the analysis. 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, expenditures were counted on a single day. The day on which costs were counted varied depending on the type of claim. Costs for inpatient stays, short-term residential treatment stays, and intensive day treatments were counted on the discharge date and costs for ED visits were counted on the first date of service. Costs for all other claims were counted on the last date of service. Expenditures were excluded from this analysis if they occurred in months during which the associated patient was deemed ineligible for the analysis.

ii. Hospitalizations

Hospitalization counts were developed for ICSI, KMHS, and MMC only. The specifications for these 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. For Medicare administrative data inpatient hospital claims were identified by CLM_TYPE=60 (inpatient). 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 the following provider identifier codes:

- Long term care hospital stay—Last four characters of Medicare provider ID between ‘2000’ and ‘2299’
- Rehabilitation hospital stay—First character of PROVIDER is an ‘R’ or ‘T’ or last four characters are between ‘3025’ and ‘3099’

[56] For ValueOptions 30-day periods were used in lieu of months. Twelve 30-day periods were aggregated to form an analysis year.

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.

Step 3: Combine claims that represent the same stay and combine transfer stays with initial stays. For Medicare, we identified and combined initial and interim claims into one discharge. Interim claims had (1) the same admission date (ADMSN_DT) as the initial claim, (2) an admission date (ADMSN_DT) that was equal to the discharge date (DSCHRGDT) from the initial or another interim claim and the status on the other (previous) claim was "still a patient" (STUS_CD = 30), or (3) a claim with an admission date (ADMSN_DT) that was equal to one day after the discharge date (DSCHRGDT) of the initial or another interim claim and the status on the other previous claim was "still a patient" (STUS_CD = 30). Such claims were combined to count as a single stay.

Next, we identified and combined initial and transfer claims into a single stay. For Medicare, patients were transferred from stays with a STUS_CD equal to 02 (transferred to another short-term hospital), 66 (transferred to a CAH), 05 (another type institution for inpatient care), 43 (federal hospital), or 65 (psychiatric hospital or unit)]. These claims were combined with associated transfer claims with the same BENE_IDs, where the provider (PROVIDER) was not the same and the admission date (ADMSN_DT) was equal to the transfer out stay's discharge date (DSCHRGDT) or discharge date plus one (DSCHRGDT + 1). Initial stays were combined with transfer stays which follow and count as a single stay.

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 readmission counts were developed for KMHS only. 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 the following stays:

- Stays that ended in death—STUS_CD=20 (expired) or 41 (expired-hospice)
- Discharges with a principal diagnosis of pregnancy or condition originating in the perinatal period (ICD-9 code 630-679, V22, V23, V28, 760-779, V21, V29-39)
- Stays 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.

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

This section describes the steps in the analyses of ED visits. ED visit counts are CMMI priority measure 62. This measure includes only ED visits that do not lead to an inpatient stay. The measure does include observation stays that do not lead to admission for Medicare patients.

In the Medicare administrative data, we identified ED claims as claims with claim type=Outpatient that have any revenue center code on any line item with the value:

- 0450 = Emergency room-general classification
- 0451 = Emergency room-emtala emergency medical screening services (eff 10/96)
- 0452 = Emergency room-ER beyond emtala screening (eff 10/96)
- 0456 = Emergency room-urgent care (eff 10/96)
- 0459 = Emergency room-other
- 0981 = Professional fees-emergency room

Then, we excluded line items that have a procedure code equal to 70000–79999 or 80000–89999, which identify lab/imaging services. As long as one line item meets the above criteria without being excluded as lab/imaging, we count the entire claim as an ED visit.

Observation stays were identified based on having any line item on the claim with revenue center = 0760 OR 0762 and CPT-code = G0378 and the unit count greater than or equal to 8. If any claim line item met these criteria then the entire claim was identified as an observation stay.

In the ValueOptions claims data the same set of revenue center codes were used to identify ED services as were used in the Medicare data analysis. All medical claims not classified as inpatient were reviewed for these codes. If these codes were present and one of the following criteria were met, the claim was considered an ED visit:

- The procedure code was in the range of 10040–69979 and the place of service was hospital emergency department; or
- The procedure code was 99281–99285.

ED visits that led to inpatient stays (i.e., ones that share the same start date with an inpatient stay) were excluded.

For both the Medicare and ValueOptions analyses 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 annual analysis periods. For individuals with less than 12 months eligible for a given 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.[57]

2. Other measures

In this section we describe the methods for creating other dependent and independent variables included in our analyses. We discuss the variables for each awardee in turn below.

a. CHCS

For CHCS, we are unable to calculate CMMI's four core measures because few CHCS participants are enrolled in Medicaid or Medicare. Instead CHCS provided data files including variables related to key program outcomes, such as psychological stress, capacity for self-management, feelings of hope, and capacity for life change; we provided impact estimates for these outcomes. In this section, we first describe the methods associated with these four outcome variables. Then, we describe the methods for the control variables included in the regressions.

CHCS administers an assessment package to each of its participants and control group members at three time points: at enrollment, six months after enrollment, and twelve months after enrollment. The package includes a number of survey instruments to cover the breadth of outcomes specified in the CHCS driver diagram, which include improved self-management capacity, reduced psychological stress, improved readiness to address health issues, and increased feelings of hope. After consulting with staff at CHCS and UT-Austin, composite scores were chosen 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. 2009) and populations with severe mental illness (Pahwa et al. 2012).
- **University of Rhode Island Change Assessment (URICA).** The URICA is used to assess motivational readiness for change. Respondents 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). Responses within each stage are averaged to create a stage-level score.[58] These four

[57] For KMHS and MMC, weights for comparison group members were also based on the number of comparison group members associated with the same participant.

[58] One of the eight questions in each of the four stages is omitted from the calculation of the stage-level score.

stage scores are then combined to create a score indicating motivational readiness to change, which is equal to the sum of the contemplation, action, and maintenance scores, minus the pre-contemplation score. Ranging from 2 to 14, higher scores represent higher levels of readiness. This readiness 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).

- **Short Form 36 Health Survey, Version 1 (SF-36).** The SF-36 was developed to assess quality of life by using domains including physical functioning; physical and emotional role limitations; general health; pain; emotional well-being; social functioning; and energy or fatigue. Improvements on this measure may mean that CHCS is achieving its goal of improving the participant's capacity for self-management. Ratings of selected items within each domain[59] are combined using a factor weighting method,[60] into a composite measure (the SF-6D) 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). Further, 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).
- **Adult Hope Scale (AHS).** The AHS assesses the participant's feelings of hope on two subscales. The first subscale, pathways, measures the ability to plan routes to achieve desired goals. The second scale, agency, measures 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). Item ratings from separate sets of four items are summed to derive subscale scores, and the two subscale scores are summed to create a global score of hope, ranging from 8 to 64; higher scores represent 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 psychiatric patients (Brouwer et al. 2008) and among traumatic injury survivors (Creamer et al. 2009).

Because participants and control group members could have missing scores for some items for a given assessment, we imputed item scores using the following methods:

- For the BSI-18 global distress score, we averaged the available information from 18 items. This occurred for 53 participant records (8 percent) and 47 control group records (8 percent).

[59] Specifically, 11 items are used from the following domains: physical functioning (three questions), physical and emotional role limitations (two), pain (two), emotional well-being (two), social functioning (one), and energy or fatigue (one).

[60] This method involves using factor coefficients calculated by Brazier and Roberts (2004) to weight scores in each of six dimensions 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 dimensions is at the "most severe" level to take account of any additional effect on health state.

- For the URICA readiness score, we averaged the available information from the seven items within each of the four domains. This occurred for 28 participant records (4 percent) and 32 control group records (5 percent).
- For the AHS global hope score, we averaged the available information from the four items within each of the two subscales. This occurred for 7 participant records (1 percent) and 7 control group records (1 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 missing scores.

Additionally, if more than 30 percent of item scores used in calculating a composite score were missing for a given assessment, we recoded that composite score to missing.[61] However, this did not occur for any composite scores for any participants or control group members.

The control variables included in the CHCS regression models are listed in Table A.1 along with the specifications for the variables. All variables were derived based on the program enrollment and survey data provided by CHCS.

Table A.1. Impact analysis model control variable specifications—CHCS

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.
Interaction between intervention period and treatment	Interaction between intervention period and treatment indicator variables.
Enrollment date	Categorical variable for member's enrollment month and year. Enrollment dates span between February 2013 (reference) and September 2014.
Age	Categorical variable indicating member's age group at time of enrollment. Categories include: 18–34 (reference); 35–44; 45–54; 55 and older.
Sex	Categorical variable of member's sex. Categories include: female or transgender (reference); male.
Employment status	Categorical variable of member's employment status. Categories include: employed (reference); unemployed.
Insurance status	Categorical variable of member's insurance status. Categories include: neither Medicare nor Medicaid (reference); Medicare; Medicaid; both Medicare and Medicaid.
Education status	Categorical variable of member's education status. 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. Categories include: homeless (reference); not homeless; incarcerated or other living situation.

[61] This item response rate threshold is based on a similar benchmark defined by the Office of Management and Budget (2006).

b. ICSI

The control variables included in ICSI's pre-post analyses are listed in Table A.2 along with the specifications for the variables.

Table A.2. 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 first year post intervention
Race	Categorical variable indicating the individual's race. Categories include: White; Black; Other (reference).
Age	Continuous variable indicating age on the first day of the observation period.
Sex	Categorical variable of member's sex. Categories include: female (reference); 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
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, ESRD, or disability and ESRD)
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 first day of the baseline period.
Depression	Indicator variable for presence of depression prior to the start of the intervention period
Acute myocardial infarction[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
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

Variable name	Specification
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

c. KMHS

The control variables included in the CHCS regression models are listed in Table A.3 along with the specifications for the variables. All variables were derived based on the program enrollment and survey data provided by CHCS.

Table A.3. Impact analysis model control variable specifications—KMHS

Variable name	Specification
Intervention period	Categorical variable indicating time period of observation. Categories include: baseline period (pre-intervention; reference); first year post intervention; second year post-intervention.
Treatment indicator	Categorical variable indicating treatment status. Categories include: comparison group member (reference); KMHS intervention participants.
Interaction between intervention period and treatment	Interaction between intervention period and treatment indicator variables.
Time period	Categorical variable indicating the calendar year of the initial month of observation period. Categories include: 2010 (reference); 2011; 2012; 2013; 2014.
Race/ethnicity	Categorical variable indicating the individual's race/ethnicity. Categories include: White non-Hispanic (reference); Black non-Hispanic; Hispanic; Other or unknown race
Age	Continuous variable indicating age when first used mental health service at KMHS or a comparison facility in the analysis period (January 2010–September 2014).
Age squared	Continuous variable measuring age as defined above squared.
Sex	Categorical variable of member's sex. Categories include: female (reference); male.

Variable name	Specification
Most common mental health diagnosis	Categorical variable for most common mental health diagnosis in first three months in analysis period receiving services at KMHS or comparison facility. Categories included: schizophrenia (295.xx); recurrent and severe depression excluding bipolar disorder (296.33, 296.34); other depression excluding bipolar disorder (296.2x; 296.3x, 296.5x, 296.6x, 298.0, 300.4, 309.1, 311.xx); bipolar disorder (296, 296.0x, 296.1x, 296.4x, 296.7x, 296.80, 296.81, 296.82, 296.89); psychosis (298.xx); anxiety disorder (300.xx); adjustment reaction (309.xx); other (all other diagnoses; reference)
Alcohol disorder	Indicator for alcohol disorder diagnosis code (291.x, 303, 303.0, 303.9, 305.0) on any claim in 12 months prior to first mental health visit in analysis period at KMHS or comparison facility
Drug disorder	Indicator for drug disorder diagnosis code (292.x, 304.x, 305, 305.2-305.9) on any claim in 12 months prior to first mental health visit in analysis period at KMHS or comparison facility
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 based on disability
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.
HCC Score	Continuous variable measuring HCC risk score 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
Time in care	Continuous variable measuring months of since first mental health visit at KMHS or comparison facility as of first month of observation period.
Time in care squared	Square of time in care variable as defined above.

d. MMC

The control variables included in the MMC impact regression models are listed in Table A.4 along with the specifications for the variables.

Table A.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); nine months post enrollment
Treatment indicator	Categorical variable indicating treatment status. Categories include: comparison group member (reference); 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) to 1Q2014.
Race	Categorical variable indicating the individual's race. Categories include: White (reference); Black; Other.
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); 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

Variable name	Specification
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 first day of observation period.
Alcohol disorder	Indicator variable for alcohol use disorder diagnosis on any claim in 12 month pre-period.
Drug disorder	Indicator variable for drug disorder diagnosis on any claim in 12 month pre-period.
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

e. ValueOptions

The treatment population for ValueOption's intervention had few hospitalizations in the pre-intervention period. Therefore, the CMMI core measures for hospitalizations and hospital readmissions were inappropriate for measuring program impacts. The members of the treatment population were frequent users of residential and intensive day treatment for substance use disorders in the pre-intervention period. Thus, instead of inpatient stays and readmissions, for ValueOptions, we examined program impact on short-term residential treatment stays and days of intensive day treatment. Here, we first describe our methods for creating these variables. Then, we describe the specifications for the control variables included in the impact regression models.

To estimate the number of short-term residential treatment stays and days of intensive day treatment we started with all claims with type=inpatient. Then we identified three distinct service types based on revenue code: inpatient, residential, and intensive day treatment. Residential stay claims were identified based on the revenue code 907 and intensive day treatment claims were identified based on the revenue code 1002.

After eliminating duplicate and denied claims, we combined claims that represent the same residential stay or the same intensive day treatment using an approach similar to that used in the Medicare data with one exception. Because patient discharge status codes in the ValueOptions data were not reliable enough to distinguish between when a patient was transferred to another facility and when an entirely separate stay began, all claims which had an admission date on, before, or the day after the previous discharge date were considered to be part of the same stay (or treatment). For example, if a residential claim ended on March 31 and the next residential claim began on April 1, those claims are considered to be part of the same stay. However, if the second claim began on April 2 instead of April 1, it would be considered as part of a separate stay. To identify the number of days of an intensive day treatment, the difference between the discharge date on the last claim and the admission date on the first claim of the treatment plus 1 was used. Short-term residential treatment stays and intensive day treatments were counted in the intervention year or baseline year based on the discharge date.

The control variables included in the ValueOptions impact regression models are listed in Table A.5 along with the specifications for the variables.

Table A.5. Impact analysis model control variable specifications—ValueOptions

Variable name	Specification
Intervention period	Categorical variable indicating time period of observation. Categories include: 12 months pre-intervention (reference); 12 months post-intervention.
Treatment indicator	Categorical variable indicating treatment status. Categories include: comparison group member (reference); intervention participants.
Interaction between intervention period and treatment	Interaction between intervention period and treatment indicator variables.
Enrollment time period	Categorical variable indicating the month the individual enrolled in the intervention. Categories include each month from March through December 2013
Age	Continuous variable indicating age calculated as 2013 minus the year of birth for the post-intervention period and 2013 minus the year of birth minus one for the pre-intervention period.
Age squared	Continuous variable measuring age as defined above squared.
Sex	Categorical variable of member's sex. Categories include: female (reference); male.
Baseline physical health SF-12	The SF-12 v2 is a multipurpose short form survey with 12 questions, selected from the SF-36 Health Survey that ask about health experiences. A subset of these questions from the baseline survey were combined, scored, and weighted to create the physical health composite score that measures physical functioning.
Baseline mental health SF-12	The SF-12 v2 is a multipurpose short form survey with 12 questions, selected from the SF-36 Health Survey that ask about health experiences. A subset of these questions from the baseline survey were combined, scored, and weighted to create the mental health composite score that measures mental functioning.[62]
Baseline alcohol use severity index	This index is a composite constructed from questions from the Addiction Severity Index (ASI) Lite, a semi-structured instrument designed to provide important information about aspects of a patient's life which may contribute to his/her substance abuse disorder (methodology available at: http://www.tresearch.org/wp-content/uploads/2012/09/CompositeManual.pdf).
Baseline drug use severity index	This index is a composite constructed from questions from the Addiction Severity Index (ASI) Lite, a semi-structured instrument designed to provide important information about aspects of a patient's life which may contribute to his/her substance abuse disorder (methodology available at: http://www.tresearch.org/wp-content/uploads/2012/09/CompositeManual.pdf).

C. Matched comparison group methods

In this section we discuss the methods used to identify the comparison group for the KMHS and MMC analyses.

1. KMHS

In this section we describe the steps taken to select the intervention and matched comparison groups for KMHS and provide diagnostics to assess balance between the matched groups.

Beginning on January 1, 2013 KMHS altered the approach to care delivery for all patients of its adult and child care teams. The change provided for integrated mental health and substance

[62] Ware JE, Jr, Kosinski M, Keller SD. "A 12-Item Short-Form Health Survey: Construction of Scales and Preliminary Tests of Reliability and Validity." Medical Care. 1996;34(3):220–33.

abuse treatment. Also, the care teams began to coordinate behavioral and physical health treatment. We used matching techniques to develop a matched comparison group for KMHS' Medicare FFS enrolled patients. 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). Because all KMHS patients are considered intervention participants, we identified patients of other mental health treatment facilities in the state of Washington as the potential pool of comparison patients. Then from within this pool we identified individuals most closely matched to KMHS patients to include in comparison population. Constructing the matched comparison group involved several steps which we detail below.

Step 1: Identify facilities similar to KMHS in Washington state. Using SAMHSA'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,[63] or being in a geographic area of similar size; however, this would reduce the number of facilities from which to identify potential comparison group members to only five and would not allow for sufficient number of potential comparison clients well-matched to KMHS clients. The current analysis period includes calendar years 2010 through 2014. Of the 24 facilities initially identified, we excluded 7 facilities because they did not serve Medicare clients in all five 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 2: Identify treatment and potential comparison group members. Using Medicare data for calendar years 2010 through 2013, we initially identified all individuals who receive a mental health service at KMHS or one of 16 potential comparison facilities.[64] 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.9 were selected for our initial analysis population. It should be noted that on January 1, 2013 the CPT codes used to bill

[63] 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.

[64] 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.

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.[65] Medicare enrollment and claims data for January 2009 through March 2014 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 in January 2010 or later. Most frequently billed mental health diagnosis at treatment initiation was measured in the initial month of mental health treatment and the two subsequent months.

Table A.6. Codes used to identify mental health services (KMHS)

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.7. ICD-9 Mental health diagnosis codes (KMHS)

Principal diagnosis codes	Label
295.00 to 295.95	Schizophrenia spectrum disorders
297.0 to 298.9	Other psychotic 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	Depressive disorders
296.90, 296.99, 293.83, 300.9	Other mood disorders
V62.84, V62.85	Suicidal or homicidal ideation
E950, E951, E952, E953, E954, E955, E956, E957, E958, E959	Injury from suicide attempt or suicide
300.00 to 300.11, 300.20 to 300.3, 309.81	Anxiety disorders
300.12 to 300.15, 300.6	Dissociative disorders
300.7 to 300.89	Somatoform disorders
301.0 to 301.9	Personality disorders
307.40 to 307.49	Sleep disorders
312.0 to 312.23, 312.4 to 312.89, 313.81	Disruptive behavior disorders (CD, ODD)
312.30 to 312.39	Impulse control disorders not elsewhere classified

[65] 72 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.

Principal diagnosis codes	Label
302.0 to 302.9	Sexual and gender identity disorders
299.00 to 299.91	Autism Spectrum disorders
307.1, 307.5, 307.51	Eating disorders
314.00 to 314.01	ADHD
307.20 to 307.3, 313.0 to 313.3, 313.82 to 316	Other disorders diagnosed in childhood
648.4	Mental disorders in pregnancy
V65.2	Person feigning illness
V71.09	Observation for other suspected mental condition
780.09	Other alteration of consciousness
V15.41, V15.42, V15,81, V17.0, V60.0, V62.29, V62.4, V62.81, V62.89	Social/contextual circumstances
All other codes in the range of 290.0-299.91 and 300.00-316 (not specified above)	Other mental health diagnoses

We restricted the analysis population to those residing in the local area of the analysis facilities to assure the patients had the potential to access the facilities. 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 received services. Five percent of the analysis population (both intervention and comparison group) was excluded based on this restriction.

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 16 percent, 9 percent, and 1 percent of the analysis population, respectively.

When this step was complete the analysis population included 936 KMHS intervention participants and a pool of 6,692 individuals who received mental health services from comparison facilities.

Step 3: Match treatment participants at individual level. The next step involves creating a matched comparison group, matching at the individual-level using characteristics derived from Medicare enrollment and claims data. As discussed above, the goal is to select individuals for the comparison group who are as similar as possible to individuals in the intervention group. An individual's propensity score is the probability of belonging to intervention group estimated based on a logistic regression model. Up to five comparison group beneficiaries were matched to each treatment group beneficiary.

We matched as many comparison group beneficiaries to each treatment group beneficiary as possible. When a treatment beneficiary is difficult to match (that is, has few similar comparison beneficiaries), the algorithm conducts a pair matching; when there is an abundance of comparisons for a treatment beneficiary, the algorithm matches multiple comparisons. The statistical goal is 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. The matching algorithm included the following characteristics: age group (18–44, 45–54, 55–64, 65+), gender, disability status, and HCC score.[66] We also enhanced the matching so that we exactly matched on important characteristics: 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, and Medicare/Medicaid dual enrollment status. Where possible, we also exactly matched on the most commonly billed mental health diagnosis at treatment initiation. Because there were a limited number of comparison beneficiaries with specific diagnoses we were only able to exact match on diagnosis code for 72 percent of treatment group members. [67] We tested the sensitivity of our findings to this limitation and found them generally robust to exclusion of pairs that were not an exact match on diagnosis.[68]

[66] 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.

[67] Adding a variable to the set of exact matching variables increases the quality of the balance on that variable, but tends to do so at the expense of identifying fewer comparison beneficiaries per treatment beneficiary and poorer matches on other match variables, unless the match variables are highly correlated.

[68] Excluding pairs not matched on diagnosis, regression results had the same sign and level of significant with one exception. This exception was that the estimated decline in hospitalizations in the second year of the intervention was significant at the .10 percent level when all observations were included but was not significant when pairs that were not matched on diagnosis were excluded.

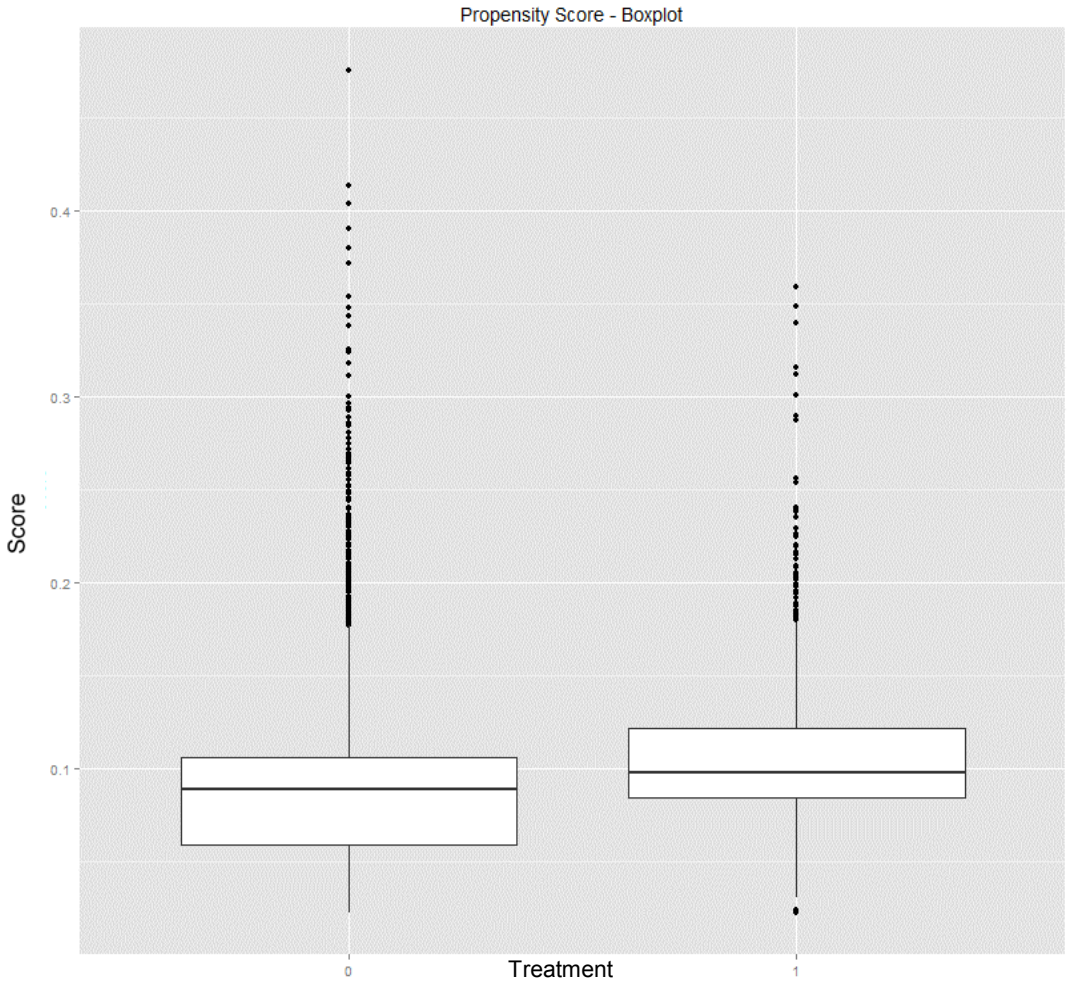
When this step was complete the analysis population included 936 KMHS intervention participants and 3,407 individuals in the comparison group. The reduction in the size of the comparison population relative to the previous step was due to individuals who were not matched to an individual attributed to 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.

Before matching, we examined the ratio of comparison beneficiaries to intervention beneficiaries by exact matching stratum in order to understand how difficult it would be to match at least one comparison beneficiary to every treated beneficiary. For example, if many strata had a low ratio of comparison beneficiaries to treatment beneficiaries prior to matching, we might consider reducing the number of variables used for exact matching. As noted above, we initially planned to use most commonly billed mental health diagnosis in the first three months during which the person was receiving treatment at KMHS or a comparison facility for an exact matching criteria for all treatment beneficiaries. However, there was a low ratio of comparison beneficiaries to treatment beneficiaries in some strata with this requirement, thus this requirement was only maintained in strata where there was a sufficient ratio.

We then graphically compared the propensity score distributions for all treatment and comparison beneficiaries prior to matching, looking for overlap in the propensity scores for the treatment and comparison groups. See Figure A.1 for an example. This test confirmed that there was a pool of comparison group beneficiaries that were available for matching with propensity scores similar to those scores observed in the treatment group.

Figure A.1. Propensity score distributions (KMHS)



Note: Figure presents boxplots created using the estimated propensity scores for the comparison and treatment 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.

After we conducted matching, we examined the number of comparison beneficiaries matched to each treatment beneficiary (Table A.11). 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 a small number of potential comparison group members with the same most common mental health diagnosis.

Table A.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	0:1
Number of matched sets	214	76	73	43	530	3,285

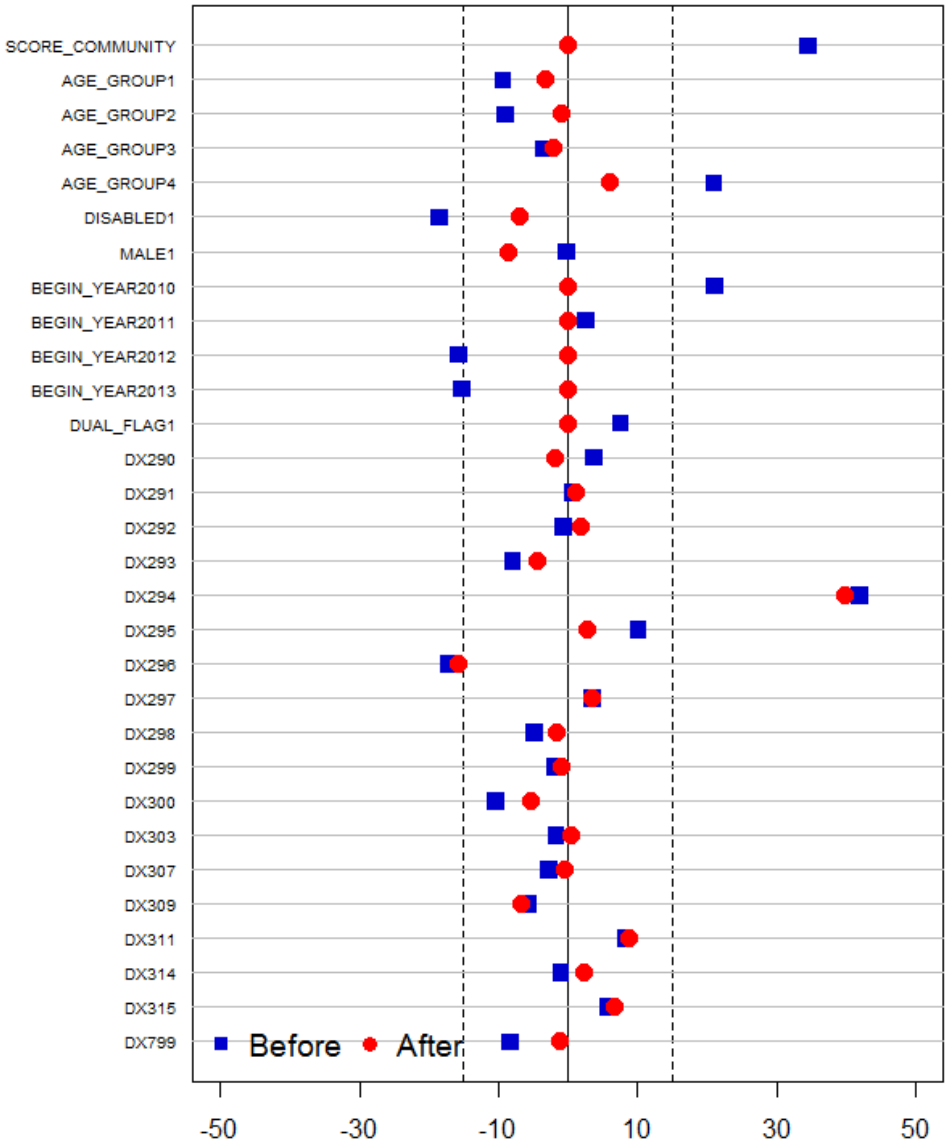
Note: Each cell indicates the number of treatment beneficiaries matched to the number of comparison beneficiaries indicated for that column. In this example, most of the treatment beneficiaries (530) were matched to 5 comparison beneficiaries.

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 163.6 and a *p*-value of < 0.01 indicating an imbalance exists. We explored the source of the difference. Removing most common diagnosis and gender from the set of matching variables assessed resulted in an insignificant p-value for the omnibus test of 0.311 indicating balance. To address the potential imbalance, we control for both gender and most common diagnosis[69] in our regression model. Also, as noted, we tested the sensitivity of the model to inclusion of pairs that were not matched on diagnosis.

To further investigate imbalance between treatment and matched comparison groups, we evaluated how matching affected the balance on all matching variables (Figure A.2) 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 balance table that shows both absolute and standardized differences between treatment and comparison groups before and after matching.

[69] The regression model included indicators for schizophrenia (295.xx); recurrent and severe depression excluding bipolar disorder (296.33, 296.34); other depression excluding bipolar disorder (296.2x; 296.3x, 296.5x, 296.6x, 298.0, 300.4, 309.1, 311.xx); bipolar disorder (296, 296.0x, 296.1x, 296.4x, 296.7x, 296.80, 296.81, 296.82, 296.89); psychosis (298.xx); anxiety disorder (300.xx); adjustment reaction (309.xx); other (all other diagnoses; reference).

Figure A.2. Balance plot comparing the standardized difference for each matching variable before and after matching (KMHS)



Note: Blue markers show the standardized difference before matching; red markers show the standardized difference after exact matching and propensity score modeling.

We found one variable outside the 0.25 bounds—specifically, diagnoses starting with 294 (which corresponds to persistent mental disorders due to conditions specified elsewhere including dementia) and one additional variable outside the 0.15 bounds—296 (which correspond to bipolar, depressive, and other mood disorders). The imbalance on diagnosis and how we addressed it was noted above. We provide more detail on the means and adjusted and standardized difference for the matching variables in Table A.12 below.

Table A.9. Balance table before and after matching (KMHS)

	Before matching					After matching				
	Comparison	Treatment	adj.diff	std.diff	p	Comparison	Treatment	adj.diff	std.diff	p
Score community	1.2813	1.6651	0.3838	0.3441	0	1.6652	1.6651	-0.0001	-0.0001	0.9986
Age_Group1: 18-44	0.327	0.2831	-0.0438	-0.0939	0.0072	0.2981	0.2831	-0.0149	-0.032	0.1627
Age_Group2: 45-54	0.228	0.1902	-0.0379	-0.0909	0.0092	0.1938	0.1902	-0.0037	-0.0088	0.7492
Age_Group3: 55-64	0.1638	0.1506	-0.0131	-0.0356	0.3072	0.1587	0.1506	-0.008	-0.0218	0.4903
Age_Group4: 65+	0.2812	0.3761	0.0948	0.2089	0	0.3494	0.3761	0.0267	0.0587	0.0108
Disabled	0.7826	0.7051	-0.0774	-0.1852	0	0.7343	0.7051	-0.0291	-0.0697	0.0066
Male	0.4553	0.4541	-0.0013	-0.0025	0.9422	0.4966	0.4541	-0.0426	-0.0855	0.0056
Begin year: 2010	0.503	0.6079	0.1049	0.2104	0	0.6079	0.6079	0	0	1
Begin year: 2011	0.1627	0.172	0.0093	0.0251	0.4727	0.172	0.172	0	0	1
Begin year: 2012	0.162	0.1047	-0.0573	-0.1585	0	0.1047	0.1047	0	0	1
Begin year: 2013	0.1723	0.1154	-0.0569	-0.1534	0	0.1154	0.1154	0	0	1
Dual flag	0.7209	0.7543	0.0334	0.0748	0.032	0.7543	0.7543	0	0	1
Most common DX: 290	0.0137	0.0182	0.0044	0.0372	0.2865	0.0205	0.0182	-0.0024	-0.0201	0.5169
Most common DX: 291	0.0009	0.0011	0.0002	0.0057	0.8709	0.0007	0.0011	0.0003	0.0106	0.818
Most common DX: 292	0.0013	0.0011	-0.0003	-0.0076	0.8267	0.0004	0.0011	0.0006	0.0177	0.4386
Most common DX: 293	0.0078	0.0011	-0.0067	-0.0807	0.0208	0.0047	0.0011	-0.0037	-0.0442	0.1408
Most common DX: 294	0.0472	0.1453	0.0981	0.4193	0	0.0522	0.1453	0.0931	0.3981	0
Most common DX: 295	0.25	0.2938	0.0438	0.1005	0.004	0.2823	0.2938	0.0115	0.0265	0.1184
Most common DX: 296	0.4419	0.3568	-0.085	-0.1719	0	0.4344	0.3568	-0.0776	-0.1568	0
Most common DX: 297	0.0058	0.0085	0.0027	0.0347	0.3194	0.0058	0.0085	0.0027	0.0348	0.3492
Most common DX: 298	0.0454	0.0353	-0.0102	-0.0495	0.1561	0.0386	0.0353	-0.0033	-0.0162	0.4377
Most common DX: 299	0.0019	0.0011	-0.0009	-0.0204	0.5584	0.0015	0.0011	-0.0004	-0.01	0.5271
Most common DX: 300	0.046	0.0246	-0.0215	-0.1053	0.0026	0.0356	0.0246	-0.011	-0.0542	0.0099
Most common DX: 303	0.0018	0.0011	-0.0007	-0.0176	0.6146	0.0009	0.0011	0.0002	0.0039	0.9122
Most common DX: 307	0.0024	0.0011	-0.0013	-0.028	0.4216	0.0013	0.0011	-0.0002	-0.0045	0.8658
Most common DX: 309	0.0774	0.062	-0.0154	-0.0584	0.094	0.0801	0.062	-0.0181	-0.0685	0.0001
Most common DX: 311	0.0244	0.0374	0.013	0.082	0.0188	0.0236	0.0374	0.0138	0.0868	0.0079
Most common DX: 314	0.0039	0.0032	-0.0007	-0.011	0.7515	0.0018	0.0032	0.0014	0.0226	0.3829
Most common DX: 315	0.0001	0.0011	0.0009	0.0568	0.1039	0	0.0011	0.0011	0.066	0.0833
Most common DX: 799	0.0182	0.0075	-0.0108	-0.0834	0.0169	0.0091	0.0075	-0.0016	-0.0123	0.6074

Step 5: Enhanced with 2014 data. More recent data became available for this analysis after the matching. Thus, we supplemented our analysis file with data for April to September 2014. Then, we imposed the same exclusion criteria described above on the new analysis period (April to September 2014). [70] Imposing these restrictions reduced the analysis population by 9 percent to 874 and 3,090, treatment and comparison group members, respectively.

Step 6: 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. (4) Lastly, individuals were dropped from our analysis if they died.

2. MMC

In this section we describe the steps taken to select the treatment and matched comparison groups for MMC impact analyses, and assess balance between the matched groups.

Beginning in October 2012, Maimonides Medical Center (MMC), a tertiary care center in southwest Brooklyn, New York, enrolled participants in a program funded by HCIA designed to improve the care of people with targeted serious mental illness (SMI) diagnoses (schizophrenia, bipolar disorder, depression, childhood emotional disturbance, and psychotic disorders). MMC is collaborating with members of the Brooklyn Care Coordination Consortium, a group of more than 20 social service agencies and medical institutions. The program engages a participant's existing medical, mental health, and community service providers, who are supported by HCIA-funded care management staff, to create a multi-disciplinary care team. Care team members share information through the Care Coordination Platform (CCP), thus giving participants a virtually co-located, coordinated medical home.

Before the HCIA funding, New York State (NYS) granted MMC status as a Medicaid health home, and MMC began providing health home services to full-benefit Medicaid enrollees. The HCIA award in turn gave MMC (1) the capacity to focus specifically on providing the same care management services to individuals with Medicare, commercial insurance, or no insurance, and (2) start-up funding for training and technology infrastructure that supports both the Medicaid

[70] Some individuals who had met all the requirements for the analysis in the January 2010 to March 2014 period did not meet these criteria in the new period (April to September 2014). For example, an individual may have decided to enroll in Medicare Advantage or an individual entitled to disability may have lost entitlement if they began to work.

health home and the HCIA program. For our quantitative evaluation, we focused on MMC's Medicare FFS enrolled patients, the population for which we had outcomes data available for this annual report. Medicare-only participants are considered "direct participants" because they benefit from HCIA funding for both care management services and the start-up funding for improved infrastructure, whereas participants with Medicaid benefits are considered "indirect participants" who benefit from HCIA funding only for improved infrastructure. Our analytic population includes both Medicare-only participants and participants dually enrolled in Medicaid and Medicare. Although some of the dual eligibles in our analyses are "indirect participants" because they have full Medicaid benefits, we included these participants in our analyses for three reasons: (1) they increase our power to detect effects of the intervention; (2) they do not introduce any bias into our analyses—that is, services provided by the Medicaid health home are the same as intervention services provided under HCIA for non-Medicaid enrollees; and (3) we had no data to distinguish dual eligibles based on whether they have full or partial Medicaid benefits.

We used matching techniques to develop a comparison group for MMC's Medicare FFS 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 Philadelphia with schizophrenia, bipolar disorder, or depression as the potential group of comparison patients. We focused on this subset of all qualifying conditions, because only five treatment group members in Medicare FFS had a psychotic disorder and no treatment group members had a disturbance of emotions specific to childhood and adolescence. We chose a comparison group outside of NYS for several reasons. First, NYS' health home program was implemented across NYS and full dual eligibles are able to enroll. Therefore, potential comparison group members in NYS who are full dual eligibles may be 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 are not participating in the intervention but who receive services from a HCIA-participating provider, and thus may indirectly benefit from the intervention. By going outside of NYS to choose the comparison group, this potential contamination is avoided. Like NYS, Philadelphia is a major east coast metropolitan area. Unlike NYS, Pennsylvania, including Philadelphia, did not start a health homes program. In addition, we did not identify any major changes to the Medicaid program during the relevant time period, thus care was more stable in Philadelphia making it a good comparison. From within the general pool of Medicare FFS enrollees with SMI in Philadelphia we matched individuals to MMC's participants. Constructing the matched comparison group involved several steps which we detail below.

Step 1: Identify treatment group participants. MMC provided Mathematica with data on their enrolled Medicare and Medicaid populations. The first Medicare FFS enrollee in the intervention enrolled in February 2013. We included beneficiaries in the treatment group if they enrolled between February 2013 and March 2014, provided they had at least six months of FFS

Medicare data in the year prior to enrollment in the intervention (N=379). From Medicare enrollment and claims data, we extracted enrollment history, demographics, health status (based on diagnoses reported on claims) and Medicare health care utilization, as measured by the four core outcomes (number of ED visits, number of inpatient hospital stays, number of re-admissions, and total expenditures) as well as number of visits with primary care providers, during the 12 months prior to enrollment in the intervention (or, the "pre-period") for all FFS participants. We excluded N = 107 participants from the sample because they were enrolled in Medicare FFS Parts A and B with Medicare as primary payer for less than six months in the 12 months before enrollment (or the "pre-period") or because they did not have a diagnosis for at least one of the three qualifying conditions in the claims data. 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. The specific diagnoses codes used are listed in Table A.13 below). As noted above, we excluded five people with psychotic disorders. In addition, we dropped individuals reported as enrolled in the intervention who did not have any of the qualifying conditions based on our definition. Our final pre-matching treatment group included 270 participants.

Table A.10. Diagnoses codes used to identify targeted conditions in treatment and comparison groups (MMC)

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,

Step 2: Identify potential comparison group clients. Potential comparison group beneficiaries included those residing in Philadelphia with one or more of the three primary target mental health conditions: depression, schizophrenia, or bipolar disorder. We used the same criteria to define the potential comparison group as we did the treatment group (as described above). Specifically, we required six months of FFS Medicare data in the year prior to pseudo-enrollment (pseudo-enrollment is the control group equivalent of enrollment date), and at least one relevant diagnosis (defined in the same way as for the treatment group). Beneficiaries also had to reside in zip codes in the city of Philadelphia during the same timeframe. We excluded any potential comparison group members whose current OR original eligibility is by ESRD.

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 intervention if they had been in the treatment group. This pseudo-enrollment month allows us to define a 12-month pre-period and the post-period timeframe, similar to pre- and post-periods for intervention participants. The pseudo-enrollment month was the first month between February 2013 and March 2014 following a month where the beneficiary had at least one of the relevant diagnoses (depression, schizophrenia, and bipolar disorder) and a visit with a primary care provider. These criteria aim to ensure potential comparison beneficiaries have at least one target condition during

the pre-period and that they have some engagement with the healthcare system. Comparison group members for whom we could not set a pseudo-enrollment date (for example, those without a primary care visit or those who were not diagnosed with one of the target conditions prior to December 2013) were dropped from the comparison pool. Our potential comparison group included 13,320 beneficiaries who met all inclusion criteria and for whom we were able to set a pseudo-enrollment date.

Step 3: Match treatment participants at individual level. The next step involves creating a matched comparison group, using individual-level characteristics during the pre-period derived from enrollment and claims data. As discussed above, the goal is to select comparison group beneficiaries who are as similar as possible to the treatment group beneficiaries on observable characteristics. A beneficiary's propensity score is the probability of belonging to treatment group estimated from this model.

We matched up to seven comparison group beneficiaries to each treatment group beneficiary. When a treatment beneficiary is difficult to match (that is, has few similar comparison beneficiaries), the algorithm conducts a pair matching; when there is an abundance of comparisons for a treatment beneficiary, the algorithm matches multiple comparisons. The statistical goal is 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. The matching algorithm first exact matched on gender, dual status, schizophrenia diagnosis, and presence of depression or bipolar disorder.[71] We then fit a propensity score model with the following characteristics: age, race (white, black, and other), original reason for entitlement code (e.g., disability, aged), month of enrollment (or pseudo enrollment date for the comparison group), number of months observable in the pre-period, HCC score,[72] and number of qualifying conditions (i.e., one, two, or three of the qualifying conditions). We also included the number of primary care visits, and pre-period levels of two of the four core outcomes measures (hospitalizations and ED visits), broken out into variables for 1 to 6 months prior to enrollment and 7 to 12 months prior to enrollment. We chose not to include total spending due to potential differences in Medicare geographic adjustments to payment in New York City versus Philadelphia that might lead to different levels of spending for individuals with the same acuity. We also did not include readmissions because individuals with values of zero include both those who had no admissions and those with one or more admissions but no readmissions, which makes it challenging to interpret whether the treatment and matched comparison group are balanced on this variable.

[71] We combined depression and bipolar disorder together to reduce the number of exact match categories to facilitate matching. These two conditions have similar expected spending and utilization, per HCC algorithm (see footnote below)

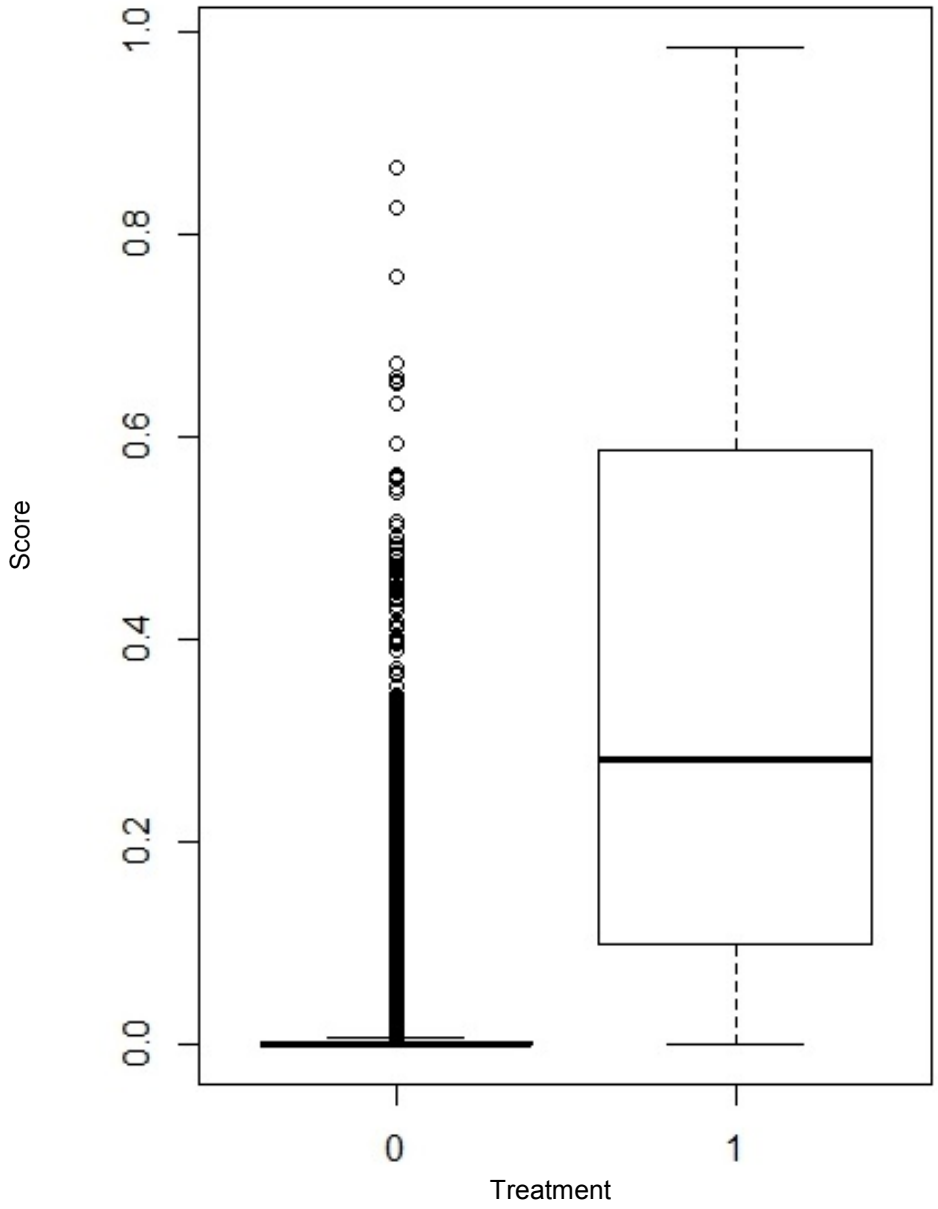
[72] HCC score = Hierarchical Condition Category Score. The HCC model was developed to risk adjust Medicare payments to Medicare Advantage plans by assessing expected expenditures of enrollees. The HCC score provides a proxy of overall health status, as sicker individuals are expected to cost more than healthier individuals.

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 may be to match at least one comparison beneficiary to each participant. For example, if many strata have low ratios of potential comparison beneficiaries to participants prior to matching, we might consider reducing the number of variables used for exact matching. As noted above, we initially planned exact match on both bipolar diagnosis and depression diagnosis. However, there was a low ratio of comparison beneficiaries to treatment beneficiaries in some strata with this requirement, thus we combined these conditions into one variable.

We graphically compared the propensity score distributions for all treatment and comparison beneficiaries prior to matching, looking for overlap in the propensity scores for the treatment and comparison groups. The distribution of propensity scores differed between the two groups, although there was overlap in the scores. Most comparison beneficiaries had propensity scores of zero, while the mean propensity score for the treatment group was above 0.2. In addition, there were few potential comparison beneficiaries in the upper 25th percentile of propensity scores among the treatment group (Figure A.3). As a result, we knew that we would be able to find at least one match for many treatment group participants, but it might be difficult to match some treatment group participants in the upper tail of the propensity score distribution.

Figure A.3. Propensity score distributions for treatment and potential comparison groups (MMC)



Note: Figure presents boxplots created using the estimated propensity scores for the comparison and treatment groups, 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.14). 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 know that treatment group members were actively recruited and that this selection criteria could not be mimicked for the

comparison group, in addition to differences in distribution of propensity scores. As shown in the matching table below (Table A.15), there were differences in the month of enrollment/pseudo-enrollment between the treatment and comparison group that likely affected the propensity scores. For example, 51 percent of comparison group members had their pseudo-enrollment date set to February 2013, less than one percent of the treatment group members enrolled in this month. This difference reflects in-part our inability to mimic the enrollment process, which involved outreach by MMC and its partners directly to individuals identified as having qualifying conditions to encourage them to enroll.

Table A.11. 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	0:1
Number of matched sets	223	3	2	1	7	13,029

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 (219) were matched to one comparison beneficiary.

We were unable to find matches for 34 treatment beneficiaries. As shown in Table A.15 below, the majority of these individuals had at least two of the targeted conditions, and were high users of primary care services. In the 1 to 6 months prior to enrollment, these beneficiaries had an average of 21 primary care visits, and in the 7 to 12 months prior to enrollment they had an average of 17 primary care visits. In contrast, potential comparison group beneficiaries that did not get matched only had an average of 6 primary care visits in the 1 to 6 and 7 to 12 months prior to enrollment. We were unable to find comparison group members to make strong matches who matched this high level of primary care utilization.

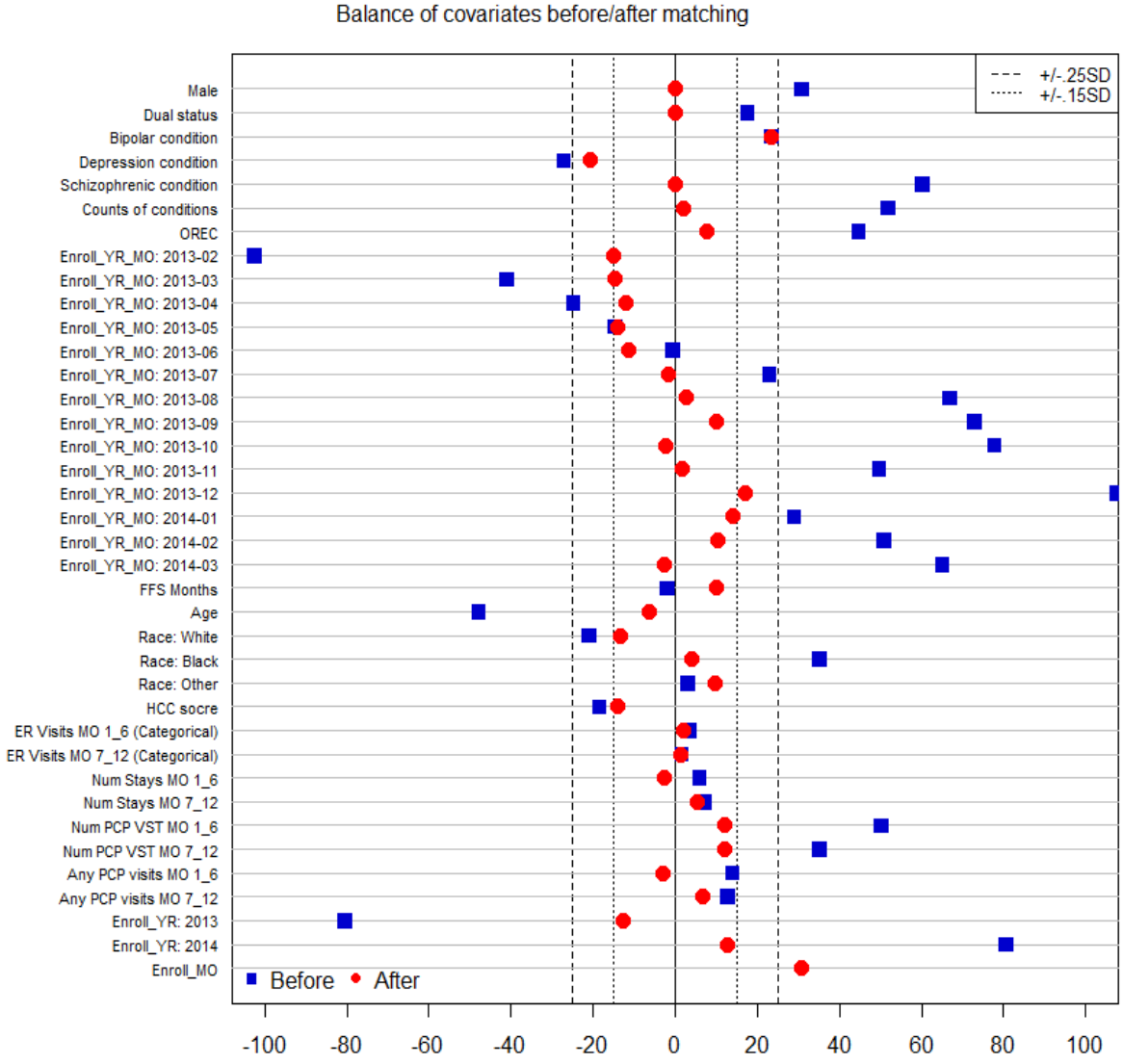
Table A.12. Characteristics and service use of unmatched treatment and comparison beneficiaries (MMC)

	Treatment		Comparison	
	N	Percentage	N	Percentage
Total	34	100	13,029	100
Demographic characteristics				
Male	20	58.8	5275	40.5
Female	14	41.2	7754	59.5
Dual eligible	24	70.6	7415	56.9
Medicare only	10	29.4	5614	43.1
Qualifying conditions				
Bipolar	17	50	4,940	37.9
Schizophrenia	26	76.5	3,908	30.0
Depression	14	41.2	6,615	50.8
Service use	Mean	Range	Mean	Range
Number of primary care visits 1 to 6 months before enrollment	21	12-39	6	0-51
Number of primary care visits 7 to 12 months before enrollment	17	1-33	6	0-47

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 90.12 and associated *p*-value of <0.01.

To further investigate imbalance between treatment and matched comparison groups, we evaluated how matching affected the balance on all matching variables (Figure A.4) by comparing the absolute and standardized differences 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. 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 and +/- 0.25 standardized differences to visually inspect whether we obtained good balance for each variables, and using balance table that shows both absolute and standardized differences between treatment and comparison groups before and after matching.

Figure A.4. Balance plot comparing the standardized difference for each matching variable before and after matching (MMC)



In addition to the exact match variables (with zero absolute and standardized differences), we ideally wanted the core outcomes measures (for example, ED visits and inpatient stays) and individual measures of depression and bipolar (which we combined for exact matching) to fall within +/- 0.15 standardized differences and all other variables within +/- 0.25 standardized differences. Most variables met these conditions. However, the individual depression and bipolar binary variables fell within 0.25 standardized difference (and not within 0.15 as desired). One variable, calendar enrollment month fell outside the 0.25 bounds. However, further assessment of this variable suggests the absolute difference is relatively small (Table A.16). Specifically, the absolute difference in mean enrollment month after matching is 1.06, meaning that treatment group members' mean enrollment date was 1.06 months after comparison group pseudo-enrollment month. Although race fell within 0.15 standardized differences, there are more white individuals in the comparison group. We will adjust for baseline differences after matching on all these variables in our impact models.

Table A.13. Balance table before and after matching (MMC)

	Before matching					After matching				
	Comparison	Treatment	adj.diff	std.diff	p-value	Comparison	Treatment	adj.diff	std.diff	p-value
Male	0.41	0.56	0.15	0.31	0.00	0.55	0.55	0.00	0.00	1.00
Dual status	0.57	0.66	0.09	0.17	0.00	0.65	0.65	0.00	0.00	1.00
Bipolar	0.38	0.49	0.11	0.23	0.00	0.38	0.49	0.12	0.23	0.00
Depression	0.51	0.37	-0.14	-0.27	0.00	0.47	0.37	-0.10	-0.21	0.01
Schizophrenia	0.30	0.58	0.28	0.60	0.00	0.55	0.55	0.00	0.00	1.00
Number of conditions	1.19	1.44	0.25	0.52	0.00	1.40	1.41	0.01	0.02	0.68
Original reason for entitlement (OREC)	0.67	0.88	0.21	0.45	0.00	0.87	0.90	0.02	0.08	0.33
Enrollment year and month										
2013-02	0.51	0.01	-0.51	-1.03	0.00	0.04	0.01	-0.03	-0.15	0.00
2013-03	0.15	0.01	-0.15	-0.41	0.00	0.04	0.01	-0.03	-0.15	0.01
2013-04	0.08	0.01	-0.07	-0.25	0.00	0.04	0.02	-0.02	-0.12	0.04
2013-05	0.04	0.01	-0.03	-0.15	0.02	0.03	0.01	-0.02	-0.14	0.10
2013-06	0.03	0.03	0.00	-0.01	0.92	0.05	0.03	-0.02	-0.11	0.24
2013-07	0.03	0.07	0.04	0.23	0.00	0.08	0.08	0.00	-0.02	0.86
2013-08	0.02	0.13	0.10	0.67	0.00	0.12	0.13	0.01	0.03	0.78
2013-09	0.02	0.14	0.11	0.73	0.00	0.09	0.12	0.03	0.10	0.28
2013-10	0.02	0.14	0.12	0.78	0.00	0.16	0.15	-0.01	-0.02	0.81
2013-11	0.02	0.09	0.07	0.50	0.00	0.09	0.09	0.00	0.02	0.85
2013-12	0.02	0.17	0.15	1.07	0.00	0.11	0.16	0.05	0.17	0.07
2014-01	0.02	0.06	0.04	0.29	0.00	0.03	0.06	0.03	0.14	0.14
2014-02	0.01	0.08	0.06	0.51	0.00	0.06	0.08	0.03	0.10	0.29
2014-03	0.01	0.07	0.06	0.65	0.00	0.07	0.06	-0.01	-0.03	0.79
FFS months	11.72	11.70	-0.02	-0.02	0.74	11.56	11.70	0.14	0.10	0.22
Age	60.30	52.22	-8.08	-0.48	0.00	52.35	51.42	-0.93	-0.06	0.45
Race: White	0.38	0.28	-0.10	-0.21	0.00	0.36	0.29	-0.06	-0.13	0.12
Race: Black	0.06	0.15	0.09	0.35	0.00	0.15	0.16	0.01	0.04	0.64
Race: Other	0.56	0.57	0.01	0.03	0.62	0.50	0.55	0.05	0.10	0.23
HCC score	1.89	1.58	-0.30	-0.19	0.00	1.69	1.50	-0.18	-0.14	0.08
ER visits 1_6 months pre-enrollment	0.72	0.81	0.08	0.03	0.59	0.77	0.82	0.05	0.02	0.83
ER visits 7_12 months pre-enrollment	0.66	0.69	0.04	0.02	0.80	0.65	0.67	0.02	0.01	0.89

	Before matching					After matching				
	Comparison	Treatment	adj.diff	std.diff	p-value	Comparison	Treatment	adj.diff	std.diff	p-value
Num stays 1_6 months pre-enrollment	0.49	0.56	0.06	0.06	0.34	0.55	0.53	-0.03	-0.03	0.73
Num stays 7_12 months pre-enrollment	0.39	0.45	0.07	0.07	0.24	0.38	0.43	0.05	0.06	0.52
Num PCP VST 1_6 months pre-enrollment	5.95	8.85	2.89	0.50	0.00	6.47	7.10	0.63	0.12	0.00
Num PCP VST 7_12 months pre-enrollment	5.47	7.34	1.87	0.35	0.00	5.41	5.98	0.57	0.12	0.01
Any PCP visits 1_6 months pre-enrollment	0.87	0.92	0.05	0.14	0.02	0.92	0.91	-0.01	-0.03	0.62
Any PCP visits 7_12 months pre-enrollment	0.86	0.90	0.04	0.13	0.04	0.87	0.89	0.02	0.07	0.40
Enroll_YR: 2013	0.96	0.80	-0.16	-0.81	0.00	0.84	0.79	-0.05	-0.13	0.19
Enroll_YR: 2014	0.04	0.20	0.16	0.81	0.00	0.16	0.21	0.05	0.13	0.19
Enroll_MO	3.99	10.24	6.25	1.97	0.00	9.21	10.27	1.06	0.31	0.00

D. Impact model methods

We implemented an experimental (for CHCS only) or quasi-experimental design for selected awardees to assess program impacts. From a technical evaluation perspective, Table A.17 illustrates the basic models we used to estimate these impacts, depending on whether we have two time points of data or more than two, thus allowing for a time series analysis. A time series analysis was conducted only for KMHS.

Table A.14. General equations for estimating impacts

Design	Number of time points	Design diagram	Estimating equation
Pre-post with *no* comparison group	2	X1 O X2	$Yi = \alpha + \gamma ti + \varepsilon i$
Pre-post with comparison group	2	X1 O X2 Y1 Y2	$Yi = \alpha + \beta pi + \gamma ti + \delta (pi \bullet ti) + \varepsilon i$
Time series with comparison group	>2	X1 O X2…O Xn Y1 Y2 …. Yn	$Yi = \alpha + \beta pi + \gamma t1 + \gamma t2 + \ldots + \gamma tn + \delta 1 (pi \bullet t1) + \delta 2 (pi \bullet t2) + \ldots + \delta n (pi \bullet tn)$

Notes: (1) Yi is the outcome for the ith person; p is a dummy variable for participation in the intervention (0 = comparison group, 1 = treatment group); t is a dummy variable indicating the data collection time point (0 = pre, 1 = post, for example); (2) α, β, γ, δ are all unknown parameters; and εi is a random, unobserved "error" term that contains all determinants of Yi that our model omits (α = constant term).

β = treatment group-specific effect, to account for average permanent differences between treatment and comparison groups; γ = secular change over time in outcome common to intervention and comparison groups; δ = true effect of treatment, accounting for both participation in the intervention and time trends.

To align with the observed distributions of expenditures in the analysis populations, we assumed a GLM model with a gamma distribution and log link function when modeling total expenditures. The outcomes for hospitalizations, hospital readmissions, ED visits, residential stays, and days of intensive day treatment were measured with count data. Therefore we estimated the models for these outcomes assuming a negative binomial model or zero-inflated negative binomial models. Standard errors were adjusted to account for multiple observations for the same beneficiary.

Power analysis was not conducted for the current report. Findings from power analyses are sensitive to assumptions about modeling and the standard deviations of the outcomes in the invention and comparison populations. The intervention participants for these programs are unique and high-risk, and thus, we did not have a reliable reference population upon which to base power analysis before analyzing data on these populations. In addition, we had no control over the size of the participant populations available to analyze. All findings in this report should be considered preliminary as estimates will be refined as data becomes available for additional participants and later intervention time periods.

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

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