# THE PREVENTIVE EFFECT OF HOUSING FIRST ON HEALTH CARE UTILIZATION AND COSTS AMONG CHRONICALLY HOMELESS INDIVIDUALS:

## New Evidence Using Propensity Score Analysis

DECEMBER 2020

MASSACHUSETTS FOUNDATION

MASSACHUSETTS HOUSING & SHELTER ALLIANCE

ANALYSIS GROUP

UNIVERSITY of DENVER JOSEF KORBEL SCHOOL of INTERNATIONAL STUDIES

UMASS MEDICAL SCHOOL

## ABOUT THE AUTHORS

**Kevin Brennan** was a Manager at Analysis Group, Inc. from 2007 through August 2018. Mr. Brennan led the development of technology solutions to enable rapid and efficient analyses of very large databases to investigate legal issues. He provided expert technical advice and analytical services to consultants and clients of Analysis Group's Health Economics practice, and has prepared reports and testified as an expert witness in litigation matters on enterprise information systems and the software industry. Since August 2018, Mr. Brennan has served as Visiting Professor of Practice in the Daverio School of Accountancy at the University of Akron's College of Business Administration.

**Kathryn Buggs** was a Senior Analyst at Analysis Group, Inc. until July 2020. Since Fall 2020, she has been enrolled at the University of Michigan Law School.

**Pete Zuckerman** was a Senior Analyst at Analysis Group, Inc. until July 2020. Since Fall 2020, he has been enrolled in the Doctoral Program in the Department of Statistics at the University of Washington.

**Singumbe Muyeba**, PhD, is Assistant Professor at the University of Denver's Josef Korbel School of International Studies. His research and teaching are in the areas of international development, homelessness and housing poverty in the United States, urbanization, property rights to low-income property among the poor, and the crisis of slums in Sub-Saharan Africa. He has a book and peer-reviewed publications in national and international journals in these research areas. Prior to his appointment at the Korbel School, Professor Muyeba was Research and Evaluation Specialist at the Massachusetts Housing and Shelter Alliance.

**Alexis D. Henry**, ScD, is Associate Professor of Psychiatry Research and Evaluation Unit within the Public and Private Health Solutions arm at Commonwealth Medicine at the University of Massachusetts Medical School.

**John (Jack) Gettens**, PhD, is Assistant Professor of Family Medicine and Community Health Research and Evaluation Unit within the Public and Private Health Solutions arm at Commonwealth Medicine at the University of Massachusetts Medical School.

**Parag Kunte**, MPH, is Senior Biostatistician for the Research and Evaluation Unit within the Public and Private Health Solutions arm at Commonwealth Medicine at the University of Massachusetts Medical School.

## ACKNOWLEDGMENTS

The authors conducted the research underlying this report as part of the Analysis Group, Inc., Pro Bono Program, which enables employees to independently participate in projects that contribute to their communities and research interests. While this research was conducted by the authors in 2019, the design and development of this project has been ongoing since 2014. The authors would like to especially thank Manas Babas, Samantha Flattery, Michaela Johnson, and Erich Trieschman for their material contributions. Results and opinions expressed herein are the work of the authors and do not represent any opinions or positions of Analysis Group, Inc.

The authors would also like to acknowledge the support of Joe Finn, President and Executive Director of the Massachusetts Housing & Shelter Alliance, and Thomas Brigham, Program Director of the Home and Healthy for Good and the Social Innovation Pay for Success initiative.

The study proposal and methodology were reviewed and approved by the University of Massachusetts Medical School (UMMS) Institutional Review Board, and the study was conducted under a Data Use Agreement among MassHealth, UMMS, and the Analysis Group. We are grateful for our partner organizations' support and contributions to the study.

Finally, the authors wish to thank the Blue Cross Blue Shield of Massachusetts Foundation for its generous support throughout the entire project from the original proposal to the preparation of this final report.

Design: Madolyn Allison
Line Editing: Barbara Wallraff

# TABLE OF CONTENTS

# INTRODUCTION

The principal objective of this study is to examine the effect of the Housing First model on Massachusetts' Medicaid program (MassHealth) expenditures. Housing First offers chronically homeless individuals immediate, stable housing to serve as the foundation for the delivery of the range of other support services. In this model, support services are provided after entry into housing and being receptive to them is not a condition for maintaining housing tenure. The individual is provided housing, treatment programs for substance use, psychiatric care, and a lease agreement that is indistinguishable from one obtained on the private market. The Massachusetts Housing & Shelter Alliance (MHSA) administers two statewide Housing First initiatives: the Home and Healthy for Good (HHG) program, which has served over 1,100 formerly chronically homeless individuals since 2005, and the Social Innovation Financing Pay for Success (PFS) program, which has served over 800 formerly chronically homeless individuals since 2015.

This report is a collaborative effort among MHSA, economic consultants supported by Analysis Group, Inc. Pro Bono Program, and the University of Massachusetts Medical School (UMMS). The study methodology integrates MHSA and MassHealth administrative claims data to compare MassHealth expenditures before and after the provision of Housing First services among MassHealth members with a history of chronic homelessness (the "intervention group") to expenditures for a matched group of members who are chronically homeless but have not received Housing First services (the "comparison" group). Pursuant to the terms of an Interdepartmental Services Agreement between MassHealth and UMMS that establishes UMMS as a MassHealth Business Associate, MassHealth has allowed UMMS researchers to access the administrative data for the study.

# REVIEW OF THE LITERATURE

## STUDIES

Studies examining the effects of Housing First on the cost of health care can be observed from the beginning of the current millennium, although rudimentary forms of the intervention can be seen in scholarly studies of psychiatric health as early as 1980 (Stein & Test, 1980; Weisbrod, Test, & Stein, 1980). These early studies demonstrated that treating individuals with severe and persistent mental illnesses while they were based in their homes and allowing them to interact with their home community reduced the number of hospitalizations and length of re-hospitalizations and had a modest cost benefit relative to inpatient services. Studies by Hoult et al. (1983), Mulder (1985), Olfson (1990), and Test (1992) supported the earlier findings, and Burns & Santos (1995) further confirmed them.

Evaluations of nonprofit programs to end homelessness followed in different parts of the country near the turn of the century. From 1994 to 1998, Martinez & Burt (2006) evaluated the cost of support services for individuals with dual psychiatric and substance use disorders, demonstrating that the provision of support services led to a reduction in cost. These studies were followed by a plethora of program evaluation reports by nonprofits and research centers from 2002 onwards (Culhane, 2002, in New York; Moore, 2006, in Portland, Oregon; Massachusetts Housing & Shelter Alliance, 2007, in Massachusetts; Mondello, 2007, in Maine; Linkins et al., 2008, in California; Hirsch et al., 2008, in Rhode Island; Bamberger & Considine-Cortelyou, 2008, in San Francisco; and Nogaski et al., 2009, in Illinois). All of these demonstrated a reduction in costs. Many of these studies were published as program evaluation reports, with a few appearing as articles in peer-reviewed journals.

More recent studies have used much more rigorous methods and been published in peer-reviewed journals. These include Wright et al., 2016, in Oregon; Mackelprang et al., 2014, in Seattle; Buchanan et al., 2009, in Chicago; and Larimer et al., 2009, in Seattle. They all report significant reduction in utilization of services after the provision of hous-

ing. In order to understand the validity of these research findings, it is important to be acquainted with the prevailing methods and to critically evaluate them.

## METHODS

Studies examining the effects of Housing First on the cost of health characterize Housing First participants in terms of entry into permanent supportive housing. The control, or "before," participants are characterized either as street homelessness or extended stays in homeless shelters.

To measure outcomes, studies typically conceptualize and operationalize health services in terms of number of and claims for inpatient hospitalizations, physical health care outpatient visits, emergency room or emergency department visits, ambulance rides, and days of psychiatric hospitalizations. Additionally, they also measure in terms of claims for enrollment in substance use (detox) treatment programs, and prescription drug costs.

The prevailing evaluation method is a cost-benefit analysis. This is typically a one-group, pre-post design where average cost utilization is measured before and after the provision of housing, and the difference is computed to show how much cost reduction occurred. A randomized controlled trial (RCT) is another, more rigorous evaluation method.

A few non-experimental studies use a pre-post control design. In this method, a synthetic control group is added to a pre-post design using propensity score matching. Propensity score analysis is a statistical method used to compute the probability that an individual will receive treatment based on a set of observed factors. The goal is to systematically identify a "double" for each individual in the treatment group. The double is the same in all observed factors and different only in that s/he did not receive housing. This method allows for cost and utilization to be computed for an individual who received housing for the situation in which the double or control group did not receive housing. The control group is either constructed from data collected from homeless individuals on the streets to match those enrolled in a permanent housing program — Flaming et al. (2009) used this in their Los Angeles study — or constructed from administrative data from Medicaid, in which the control is constructed from de-identified data.

The appeal of the before and after design is understandable and, in most cases, justifiable. Before and after designs are easy and useful in assessing the relationship between housing and utilization of services. They are also a less costly way of evaluating programs, do not require highly specialized skills, can be done within a relatively short time frame, and are less demanding administratively.

Despite the appeal of the pre-post design, it cannot tell us that any reductions observed are due to permanent supportive housing. The design does not control for the possibility that observed changes are attributable to something other than housing. In terms of making causal inference, it is the least reliable evaluation method in use. Methods of evaluation range in order of strength from experiments (such as an RCT) to quasi or natural experiments to non-experimental designs and finally to the pre-post design. Pre-post design evaluations are weak in that they do not control for the influence of history. In the year before and after receiving permanent supportive housing, participants are exposed to various influences that go unobserved. For example, natural stabilization of health due to less stressful environments may reduce use of public health services. Also, some individuals may obtain their health services from nonprofits that provide health care services for homeless individuals. Another limitation is that pre-post studies lack a comparison group against which outcomes can be evaluated (see Wright, 2016:23), and the studies therefore cannot construct an approximation of the metrics for the people in housing had they not moved into housing. This makes it difficult to compute an accurate estimate of the size of the effect of supportive housing. Therefore, positive results from a pre-post study should be considered only as a basis for more rigorous designs. As Wright (2016) puts it, the results are intriguing enough to provoke more rigorous empirical investigation.

Construction of a synthetic control group to counteract weaknesses of the pre-post design may control for observed bias but not for unobserved bias. No two individuals are the same. If the data set from which the control group is se-

lected has few observations to match on, the control is less reliable because the "best" match may still be materially different from the study population. The use of scores to facilitate selection into supportive housing leads to selection bias such that the control is in a less acute condition than members of the treatment group, when the assumption is that the two are similar. This is not a problem with respect to statistical tests to establish significance of differences, but it is certainly a problem for accurately estimating the effect size – that is, the differences in cost utilization.

In general, another weakness of many previous studies is that they are subject to recall bias during data collection. They use retrospective questions in order to establish health service utilization before the receipt of supportive housing. The answers depend on the participants' ability to recall the frequency with which they used a specific service. The accuracy of this information is questionable, particularly given that substance use — with which many chronically homeless individuals struggle — has a neurodegenerative effect on the accuracy of working memory (Oscar-Berman & Marinković, 2007:243).

Other weaknesses of these studies are that sample sizes are small (a median of n=98 for 15 studies, with the lowest being n=20), which means explanatory power is compromised. This is further complicated by missing data points that occur because some participants disappear. With small sample sizes, there is weak scientific basis for generalizing results beyond the program's population.

Further, the ethical implications for RCTs makes allocation to treatment and control groups complex (see Buchanan et al., 2009), because random allocation to a group that will not receive the Housing First intervention means continued exposure to health risks and even death for some in this vulnerable population.

## RESULTS

Despite the weakness of the evaluation methods applied to date, there is general agreement that the Housing First model reduces utilization of health services for people experiencing chronic homelessness but less consensus on whether there is adequate evidence that this reduction is significant enough to allow for a scale-up of the model.

The consensus is that Housing First reduces the number of days of inpatient hospitalizations, but there is wide variation among studies in terms of the actual percentage reduction. It ranges from 29 percent to as high as 83 percent. The number of physical health care outpatient visits decreases between 27 percent and 29 percent. Reduction in the number of emergency room (ER) visits ranges from 24 percent to 87 percent; reduction in the number of ambulance rides ranges from 19 percent to 60 percent; the number of psychiatric hospitalizations ranges from 38 percent to 82 percent. The wide variation can be attributed to differences in the location of study sites but also to the weaknesses in the research methods as discussed in the previous section. The general finding, however, is that utilization of health care goes down when people experiencing homelessness are housed using Housing First.

There is less consensus on the relationship between Housing First and change in costs of treatment for chronic illnesses. Some studies report an increase in the costs of substance use treatment, mental health treatment, and prescription drugs, while others report a reduction. Buchanan et al. (2009) report a 45 percent reduction in number of days spent and costs in a nursing home by HIV seropositive clients in Chicago. Pearlman (2006) reports 82 percent fewer detox visits and a reduction in costs in Denver. On the other hand, Mondello et al. (2007) report a cost increase of 35 percent in mental health treatment programs, 22 percent in substance use treatment programs, and 31 percent in prescription drugs in Maine. Moore (2006) reports a substantial increase in outpatient visits for mental health and drug and alcohol treatment in Portland, Oregon, but he does not provide a cost estimate. The study attributes the increase to the greater attention given to physical well-being and a larger number of opportunities for treatment. Logically, if there are more individuals in a program that have substance use disorders, mental illness, and alcohol addiction, then the costs of treatment should increase once they enter housing. It may therefore depend on the nature of the clients.

The important question to ask is whether the costs of administering Housing First programs (including the cost of the actual housing), plus the possible increase in utilization of mental health, alcohol, and substance use treatments,

plus the reduced cost of providing medical services and emergency shelter, are less than the costs prior to entering permanent housing. The available evidence suggests the answer is yes, but we cannot confirm this definitively. What we know is that cost-benefit analyses show a reduction in costs per person per year, but with wide variation. Table 1 below summarizes a selection of these studies.

As Table 1 shows, the average reduction in cost per person per year varies greatly. For studies that consider the cost of administering Housing First programs, there is an average reduction in costs to the public that ranges from $900 to $29,400 per person per year after entry into a Housing First program. For those that do not consider the costs of administering these programs, the range in savings is from $4,000 to $33,125 per person per year. Across states, there is wider variation than there is for studies done within the same state. For example, the two studies done in Seattle resulted in very similar estimates, which may indicate some consistency. Similarly, it may also reflect the estimate from the use of pre-post mean difference estimates. This method tends to inflate the effect size relative to control group designs.

**TABLE 1. COST SAVINGS PER PERSON PER YEAR FOR UTILIZATION OF PUBLIC SERVICES AMONG CHRONICALLY HOMELESS ADULTS IN PERMANENT SUPPORTIVE HOUSING**

| STUDY | CITY AND/OR STATE | COST SAVINGS PER PERSON PER YEAR ($) |
|---|---|---|
| Pearlman (2006) | Denver, Colorado | 15,772 (31,545 over two years) |
| Flaming (2009) | Los Angeles, California | 27,288 |
| Hirsch et al. (2008) | Rhode Island | 7,946* |
| King County (2013) | Seattle, Washington | 33,125 |
| Larimer (2009) | Seattle, Washington | 29,388* |
| MHSA (2006) | Massachusetts | 8,949* |
| MHSA (2015) | Massachusetts | 13,401* |
| Mondello (2006) | Maine | 944* |
| Moore (2006) | Portland, Oregon | 16,299* |
| Nogaski (2009) | Illinois | 4,000 |
| Wright et al. (2016) | Portland, Oregon | 8,724* |

* These figures factor in the costs of providing supportive housing.

## NEW CONTRIBUTIONS FROM THIS STUDY

In conclusion, what has not been done in previous studies examining the effects of the Housing First model is the use of a pre-post control group design that utilizes propensity score analysis on a large sample, using administrative claims data from Medicaid. This study addresses that gap.

# METHODS

## DATA SOURCE

This analysis uses a de-identified data set drawn from MassHealth administrative eligibility and claims data files. These contain information on member demographics; plan enrollment; and inpatient, outpatient, skilled nursing facility, home health agency, hospice, durable medical equipment, and pharmacy claims from 1/1/2004 to 12/31/2017. The data set contains only de-identified information and is compliant with the requirements of the Health Insurance Portability and Affordability Act (HIPAA).

## STUDY DESIGN

A retrospective longitudinal matched cohort design was used. The index date was defined as the Housing First date of enrollment for members in the intervention cohort and as a randomly imputed date for members in the control cohort. Baseline characteristics were evaluated during the two years pre-index date (baseline period). Outcomes were evaluated from the index date up until the one-year post-index date (hereinafter referred to as the follow-up period).

The theoretical backing for the minimum time period within which results of the effects of permanent supportive housing should begin to be observed is not clearly established in the literature. The period between the baseline and the time of entry into housing differs from one study to another. Some studies use one year and others six months. The period between entry and when cost is measured after entry also differs among studies, with studies variously using six months, one year, two years, and five years. Most studies use the six-month, one-year, or two-year mark before and after entry. Choosing to use one year of follow-up data both provides a longer view of the effects of permanent housing than a six-month period and allows a large sample size of members who for whatever reason may not have two uninterrupted years of claims data in a two-plus year follow-up period.

## IDENTIFICATION OF STUDY COHORTS

Members in the Housing First intervention cohort were identified by MHSA. Their medical claims were pulled from the MassHealth claims data file by a UMMS researcher and subsequently de-identified. Of this group, members with at least two years of continuous MassHealth coverage eligibility before and one year after intervention were considered for inclusion in the intervention cohort.

A two-step process was employed to identify members of the control cohort in the MassHealth database. In the first step, a pool of all MassHealth members who had a homeless flag but were not in the intervention cohort was selected. These members were identified using indicators for homelessness. In order to identify members of the control group as chronically homeless, we required that they meet the following conditions:

- Medical diagnosis codes indicating homelessness (ICD-9-CM: V60.0, ICD-10-CM: Z59.0). We required that a member have at least two distinct claims with a primary medical diagnosis code indicating homelessness occurring at least 30 days apart.
- Presence of a homelessness indicator (or "homeless flag"[1]) maintained in the MassHealth database.

1 The "homeless flag" is generated at the time of application and enrollment into MassHealth. If an applicant indicates that they have "no home address" on the MassHealth application, the system automatically assigns the "homeless flag". If, at the time of the annual renewal, the member provides a home address, then the "flag" is turned off. It will miss episodes of homelessness that occur between application/enrollment and annual renewal. In other situations, members may provide a home address (e.g., for a family member) where they might receive mail, but don't reside.

The number of individuals who met the criteria listed above was 44,022. However, accurately identifying chronically homeless members of the control cohort was the greatest conceptual and data-related hurdle in this analysis, in large part because housing status is a byproduct rather than the primary purpose of claims data. Specifically, as with any other analysis that uses claims data, this analysis relies on medical providers to accurately and consistently use medical diagnosis codes to indicate chronic homelessness. Additionally, it relies on the assumption that the homelessness indicator maintained in the MassHealth database member files is accurate and current (i.e., that it is changed when a member becomes homeless and is changed again if they move into permanent housing). Cognizant of the challenges of maintaining indicators of homelessness in the database, we implemented a comparison of the medical diagnosis codes in claims that indicate homelessness with homelessness indicators in the MassHealth membership files to provide additional confirmation that members identified as chronically homeless were, in fact, chronically homeless. This takes us to step two.

In step two, we created a "waterfall" summary of members in the full control cohort based on their meeting of increasingly strict identification criteria. As shown in Table 2 below, 4,687 of the 44,022 total control cohort members (approximately 11%) met the strictest criteria. Then, only the individuals that meet the strict criteria were included in the propensity score matching process. Using nearest neighbor propensity score matching methods, we arrived at a matched sample of 690.

Second, we performed sensitivity tests on the analysis in which we made the criteria stricter for the amount of time between claims with primary or secondary medical diagnosis codes for homelessness. Specifically, we increased the interval between claims that indicated homelessness from 30+ days to 90+ and 180+ days. That is, whereas a member could have previously had only a total of two claims within about a month of each other in which the primary or secondary diagnosis code indicated homelessness, we required in the sensitivity tests that such claims appear over a three- or six-month span. The results of these sensitivity tests, as well as the sizes of the matched intervention and control cohorts, are consistent with the results we report for the main analysis with the 30+ day criteria.

**TABLE 2. CRITERIA FOR IDENTIFYING CHRONICALLY HOMELESS MEMBERS OF THE CONTROL COHORT**

| | IDENTIFICATION CRITERIA | MEMBERS IN CONTROL COHORT |
|---|---|---|
| | **TOTAL** | **44,022** |
| *LEAST STRICT* | No evidence of Community Support Program for those Experiencing Chronic Homelessness (CSPECH) procedure codes in full set of member's claims, which would indicate that the formerly homeless member had been housed | 43,951 |
| | Either (1) homelessness indicator or (2) medical diagnosis code for homelessness in any position[2] | 39,790 |
| | Multiple medical diagnosis codes for homelessness in any position listed in claims 30+ days apart | 14,353 |
| | Medical diagnosis code for homelessness listed as primary or secondary code | 8,898 |
| *MOST STRICT* | Multiple medical diagnosis codes for homelessness listed as primary or secondary code in claims 30+ days apart | 4,687 |

2 The MassHealth claims data include up to nine medical diagnosis codes, ordered as the medical provider input them. The first diagnosis code on each claim is considered to be the primary diagnosis code, the second is secondary, and so on.

## INCLUSION CRITERIA

We define a member's MassHealth eligibility as continuous if we observe either no gaps in eligibility or gaps of less than 30 days throughout the relevant periods outlined below.

The following inclusion criteria were applied for the Housing First intervention cohort:

- At least two years of continuous eligibility (MassHealth) prior to the index date (baseline period) and one year after the index date (follow-up period).

The following inclusion criteria were applied to the non-intervention (control) cohort:

- At least two years of continuous eligibility (MassHealth) prior to the index date (baseline period) and one year after the index date (follow-up period).
- No evidence of CSPECH intervention services. CSPECH is the Community Support Program for People Experiencing Chronic Homelessness, and its intervention would mean that the individual had been housed.

## STUDY MEASURES

Baseline characteristics included demographics, physical and mental comorbidities, health care resource utilization (HRU), and costs. HRU and cost patterns were evaluated during the follow-up period, which was divided between the first six months after the index date and the second six months after the index date. HRU outcomes were classified into the following categories:

- Inpatient visits
- Inpatient days
- ER visits
- Mental health encounter visits
- Other visits

Cost outcomes were classified into the following categories:

- Pharmacy
- Medical
- Inpatient
- ER
- Mental health
- Other

## STATISTICAL ANALYSIS

Continuous variables were described with means and standard deviations (SDs), and categorical variables were described with frequencies and proportions. Unmatched baseline characteristics were compared across cohorts using Wilcoxon rank-sum tests for continuous measures and chi-square tests for proportions. Housing First intervention members were matched 1:1 to members in the control cohort using propensity score models based on key demographics, including age, sex, race, and year of the index date, as well as the Charlson Comorbidity Index (CCI) and total baseline health care costs.

### *Estimating the Propensity Scores*

Propensity score matching is a common tool employed in retrospective observational studies where complete randomization of treatment is not possible. This tool is used to address the inherent selection bias in the observed treatment and control groups, who have unequal distributions of variables that are thought to be correlated with the outcome of interest. This is unlike a randomized control trial, in which the observed, as well as the unobserved, covariates are balanced across groups and causal inferences can be drawn directly from the observed treatment effect. Propensity score matching methods are effective in balancing the observed covariates across groups; they do not, however, guarantee balanced unobserved covariates, which may or may not be correlated with the outcome (Rosenbaum and Rubin, 1984). As a result, propensity score models are typically estimated using observed variables that are thought to affect the outcome. In our study, this outcome is the treatment, defined as being enrolled in a Housing First program.

To estimate the propensity score for each patient in the database, we used a logistic regression model in which treatment status (enrollment in a Housing First program) was regressed on baseline total health care costs, the CCI, and demographics including age and year of the index date. The interpretation of the output of this model is the probability that a given patient would have been in the Housing First cohort, given the covariates of this particular patient, if we were unable to observe their actual enrollment status. Therefore, if two patients from each cohort had the same predicted probability of treatment produced by the propensity score model, it is likely due to the close alignment of their baseline variables. In summary, matching patients in this metric is a useful tool to ensure a balanced distribution of variables thought to be linked with the outcome across cohorts.

There were 1,342 treatment and 44,022 control participants in the unmatched data. Participants were matched on the propensity score using the nearest neighbor method, with the Mahalanobis Metric and with a caliper width equal to 0.2. While no caliper width is universally superior in reducing selection bias, our specific caliper width of 0.2 was determined based on the observed variance of our propensity scores across cohorts. There is evidence that certain caliper widths are more effective at reducing selection bias based on the experimental variance of the predicted propensity score (Cochran and Rubin, 1973). Based on this matching criteria, 690 Housing First participants were matched to 690 control participants with similar characteristics.

### *Propensity Score Matching Results*

Post-match statistical comparisons show no statistically significant difference among the baseline covariates outlined in the propensity score model section, and the measures of central tendency and spread shown in Table 3 indicate the cohorts are well-balanced on the predicted probability of treatment after the nearest neighbor matching method.

**TABLE 3. COHORT-LEVEL POST-MATCH PROPENSITY SCORES: MEASURES OF CENTRAL TENDENCY AND SPREAD**

| PROPENSITY SCORE STATISTIC | INTERVENTION (N = 690) | CONTROL (N = 690) |
|---|---|---|
| Mean | 0.30 (0.16) | 0.30 (0.16) |
| Median [Interquartile Range (IQR)] | 0.305 [0.154 – 0.436] | 0.304 [0.154 – 0.437] |
| Min – Max | 0.033 – 0.619 | 0.033 – 0.626 |

Note: Standard deviations in parentheses.

Using propensity score matching, we were able to estimate the average treatment effect for the treated (ATT). This is the average response to treatment for individuals who were assigned to the treatment, or intervention, cohort. The propensity score matching model does not control for unobservable or unknown factors that may be driving observed variation in the outcomes. The assumption of the propensity score models in using observed characteristics is that there are no unobserved factors correlated with the outcome. The ATT was estimated using a comparison of

means among the matched pairs across the intervention and control cohorts. Categorical measures were compared using McNemar's tests and continuous measures were compared using Wilcoxon signed-rank tests.

# RESULTS

## BASELINE DEMOGRAPHIC AND CLINICAL CHARACTERISTICS

To highlight salient findings, at baseline, members in the matched intervention and control cohorts were aged, on average, 47.8 and 48.7 years, respectively. The age difference was not statistically significant. Male members represented 71.3% and 72.2% of the intervention and control cohorts, respectively. The gender difference was also not statistically significant. At baseline, CCI [mean±SD] was 1.7±2.0 and 1.9±2.2 in the intervention and control cohorts, respectively. There were no statistically significant differences between the groups.

Summary statistics of unmatched and matched cohorts at baseline are presented in Appendix A. The table presents baseline and clinical characteristics of the two cohorts.

## HEALTH RESOURCE UTILIZATION AND COSTS

During the one year of follow-up among matched pairs, members in the Housing First intervention cohort were less likely to have emergency room (63.3% vs 72.6%, P<0.001) and inpatient visits (41.2% vs. 52.3%, P<0.001) than members in the control cohort, and more likely to have mental health encounter visits (80.0% vs. 53.0%, P<0.001).

Although overall medical and pharmacy expenditure on the Housing First cohort was about $1,600 lower per patient per year ($13,294) than that of the control cohort ($14,781), there was no significant difference during the first six months after housing entry. But there were significant differences in individual expenditure categories where the intervention cohort scores were lower. During the first six months after housing entry, the intervention cohort expenditure on mental health encounters was significantly higher, at $9,471 compared with $6,303 for the control cohort. Pharmacy expenditure was also significantly higher, at $1,667 compared with $1,113 for the control cohort. For the category Other, the Housing First cohort had significantly lower expenditures, at $6,437 compared with $7,608. For the remaining categories — medical, ER, and inpatient hospitalizations — the Housing First cohort had lower expenditures than the control cohort but the differences were not statistically significant.

During the second six months after housing entry, the overall medical and pharmacy expenditure was even lower for the Housing First cohort than during the first six months, by nearly $3,700 per patient per year ($12,320 compared with $15,995), but the statistically significant differences are found in the individual expenditure categories. Expenditure on medical was $11,090 for the Housing First cohort and $15,305 for the control cohort, a difference of almost $5,000. The difference in medical expenditure is statistically significant. The expenditure on ER visits was $986 per person per year for the Housing First cohort compared with $1,351 for the control cohort, about $350 lower, and this is a statistically significant difference. There was a significant difference in inpatient hospitalization expenditures, which were $1,281 lower for the Housing First cohort. However, expenses were significantly higher for the Housing First cohort for mental health encounters by $1,752 and for pharmacy by $540.

During the year of the follow-up period, members in the Housing First intervention cohort had 17% lower total all-cause medical spending than members in the control cohort. When comparing the first and second six months of the follow-up period, the differences were 10% and 28% respectively. For emergency care expenditures, members in the intervention cohort had 17%, 6%, and 18% lower spending relative to members in the control cohort over the full year, first, and second six months, respectively. Mental health expenditures for members in the intervention cohort were 35% higher than for members in the control cohort for the full follow-up year (50% and 25% higher in the first

and second six months, respectively). All HRU comparisons had $p<0.001$ and thus are statistically significant, while differences in total cost were not statistically significant. The standardized difference of total health care costs between cohorts was 0.12. Summary statistics of unmatched and matched follow-up characteristics of the two cohorts are presented in Appendix B.

## DISCUSSION

In this retrospective claims-based study, the differences in health care costs and rates of HRU were assessed in a population of MassHealth members. All categories of HRU assessed in the follow-up period were significantly reduced for members in the intervention cohort of participants enrolled in the Housing First program as opposed to the chronically homeless control cohort, except for rates of mental health encounters. While the difference in rates of all-cause inpatient and emergency department visits between groups was relatively large (>10% higher for control group), the largest difference in magnitude was mental health encounters, with 80% in the intervention group and only 53% in the control group. This would suggest that while intervention members are receiving more mental health services than the chronically homeless control members, there may be a preventive effect that leads to lower use of emergency and inpatient hospital services.

This possibility is also reflected in the differences in medical costs between groups. While all-cause costs were almost $5,000 higher for the control group in the follow-up period, this was primarily driven by inpatient costs that were higher by almost $9,000 dollars. Even though intervention members incurred more mental health-related costs ($18,240 vs. $13,529), the savings from lower use of costly inpatient services were high enough to keep the total costs for these members below those for the chronically homeless control cohort.

The results of this study are consistent with conventional wisdom but also challenge some studies and new directions in policy that go against the Housing First model. The results are particularly consistent with Hirsch et al.'s (2008) study in Rhode Island and Wright et al. (2016) in Portland, Oregon. Hirsch et al. (2008) conducted an evaluation of the Rhode Island Housing First pilot program that had housed 50 chronically homeless individuals. Their primary evaluation tool was in-depth interviews with 41 participants at baseline and 63 follow-up interviews. After accounting for support services and housing subsidy, they found an overall reduction of $7,946. Similarly, Wright et al. (2016) used a mixture of survey and administrative claims data on 98 individuals in Oregon to evaluate cost utilization of the Housing First model. They show that overall claims for health care expenditures for people in Housing First were significantly lower ($8,724 less) after entry into permanent supportive housing. Expenditure changes were driven primarily by reductions in emergency and inpatient care. Using stronger methodology, this study has confirmed these findings. Further, the study by Ly and Latimer (2015) cited below confirms these findings. A key difference between these studies and our study is that they consider costs of jail and prison overnights, which are reduced significantly after housing.

The evidence in this study challenges findings of other studies. A recent observational study by Tsai, Gelberg, and Rosenheck (2019) shows that entry into supported housing with linked primary care services was not associated with improvements in physical health-related quality of life (HRQOL). They use before and after evidence from 756 chronically homeless adults. Although Tsai, Gelberg, and Rosenheck found that participants had fewer medical problems and received more preventive procedures over time, there was no statistically significant change in their physical health as measured by the HRQOL scale. The weakness in their study is that it was observational and did not have a control, while our study uses a control group design with propensity score analysis and a larger sample — much more rigorous methods.

By the use of a control group design and propensity score matching, the evidence challenges the regression to the mean argument. The regression to the mean is a statistical idea that if there are extreme scores on a variable the first time it is measured, scores will normalize the next time. Applied to homelessness, the idea is that chronically homeless individuals are at their extreme level of acuity of health at the point of enrollment since they have to meet those criteria. When observed after housing entry, a natural stabilization of health takes place and their scores normalize. The conclusion reached by those who make the regression to the mean argument is that housing is not associated with the subsequent reduction of health care costs, but rather, a natural stabilization of health takes place. However, using a matched sample, the study has shown that without housing, utilization of health services and costs remained higher for the control group. If natural stabilization took place in the absence of housing, both utilization and costs for the control group should have dropped, but they increased instead. This interpretation is consistent with Srebnik, Connor, and Sylla's (2013) study, which reached similar conclusions, albeit with a much smaller sample of 29 in their intervention group and 31 in their control group.

It is important to note that our study does not clear up the ambiguity in results between cost reduction with and without housing costs. Our study does not consider housing costs because of a focus on medical and behavioral health. To be comprehensive in the cost-benefit analysis, the study would need to account for actual claims on jail and prison nights as well as claims for shelter nights spent before and after housing. In their study that reviewed existing literature on Housing First, Ly and Latimer (2015) found that shelter and emergency department costs decreased with participation in Housing First, while impacts on hospitalization and justice costs were more ambiguous. They suggested that while their review casts doubt on whether Housing First programs can be expected to pay for themselves, "the certainty of significant cost offsets, combined with their benefits for participants means that they represent a more efficient allocation of resources than traditional services."

We identify two policy implications of the findings of this study. First, the evidence supports the expansion and scaling up of CSPECH. CSPECH is a Medicaid-funded program that provides community-based support services for chronically homeless individuals in Massachusetts and is the program under which most of the Housing First members in this study received support services. The program was started in 2006 by the Massachusetts Behavioral Health Partnership (MBHP) and MHSA. Under CSPECH, housing agencies can claim reimbursements from MassHealth for support services rendered to chronically homeless individuals, but not for housing. The evidence indicates that the housing component and the support services component under CSPECH together are associated with reductions in utilization of and claims for medical and behavioral health services.

Second, the evidence in this study also challenges current federal policy directions. Regarding policy shifts, the U.S. Interagency Council on Homelessness underwent a leadership change in late 2019 signaling direction that takes policy practice back to its pre-Housing First condition. According to media reports, the new leadership opposes Housing First and prefers what is referred to as "Housing Fourth": that "cities stop giving out food, criminalize sidewalk sleeping, and force homeless residents who want services to move into city-operated facilities in large temporary structures" (Walker, 2019). Although this direction has yet to result in actual observed policy shifts under the new leadership, the signal is strong enough to warrant emphasizing that the policy implications of this study do not support this potential policy shift and direction. The evidence in this study and consensus in the field make a convincing case for Housing First as evidence-based practice and for scaling up at the city, state, and federal government levels.

## LIMITATIONS

The present study has limitations. First, there is no single way to identify chronically homeless members in the MassHealth claims dataset. As noted, we rely on a combination of the homelessness indicator maintained by MassHealth and medical diagnosis codes that indicate homelessness. The homelessness indicator is not rigorously maintained, and it may be activated for a chronically homeless member and not deactivated after that member finds or is placed in permanent housing. Alternatively, medical providers may not use the medical diagnosis codes

indicating homelessness reliably, so the number of claims that indicate homelessness may understate the homeless population. As noted previously, we assessed members of the control cohort using a "waterfall" method of increasingly strict criteria and performed two sensitivity tests of the analysis requiring a longer time interval between claims with diagnosis codes indicating homelessness. The waterfall and sensitivity tests allowed us to avoid relying solely on the homelessness indicator and to tease out potential inaccuracies in the use of medical diagnosis codes to identify homelessness.

Second, for members identified as chronically homeless using the inclusion criteria above, it is not possible to directly identify enrollees who may have received interventions other than Housing First or CSPECH. The sensitivity tests performed also served to eliminate members whose receipt of other interventions would have moved them into permanent housing during the baseline or follow-up periods under consideration.

### AREAS FOR FUTURE STUDY

Since the study focused on effects over one year, it does not specifically address the long-term cost savings of the Housing First model. In most studies conducted to date, the cost estimates focus on the first year following housing, with a few studies continuing longer. Some scholars suggest that the first year of treatment is the most expensive and overall cost savings should increase in subsequent years (Larimer et al., 2009:1354). Evidence backing this claim needs to be supported by a longitudinal study with a comparison group.

Although our study has challenged the idea of regression to the mean, it remains to be established exactly what the role of natural stabilization of health among people experiencing homelessness is and how it is related to the subsequent reduction of health care costs when individuals access housing. In other words, it remains unknown how much variation in costs can be attributed to supportive housing and how much to natural stabilization of health.

Furthermore, it remains difficult to separate the effect of supportive services alone and those of housing alone. The closest approach to this we have is one study done in California (Linkins, 2008) that compared a group that received services only and a group that received both housing and support services. If the membership of the two groups had been essentially the same, we could subtract the cost of support services from the housed group in order to isolate the cost of housing alone, but this was not the case.

Finally, there is need for future research to go beyond medical costs and conduct a comprehensive study that considers effects of Housing First on costs related to shelter use and incarceration, and to take into account the costs of housing provision. Given its reliance on administrative claims data, this study does not estimate the cost of the Housing First housing provision or any relationship between the cost of housing and the cost of medical resource use. Further, it does not address the cost of Housing First housing as compared with the alternative cost of housing the chronically homeless population in shelters or, as is sometimes the case, in jails and prisons.

## CONCLUSION

In the present retrospective claims-based study, members in the Housing First intervention cohort were found to have significantly lower health care resource utilization than members in the control cohort. Mean costs were also lower among the intervention cohort, although not all results were statistically significant at the 5% level. Participants in the intervention cohort appear to have used relatively more mental health care services and relatively less emergency care. These findings are consistent with studies carried out elsewhere in the field. The policy implications of these findings are that it would be beneficial to scale up CSPECH, which provides community-based support services for the chronically homeless in Massachusetts, and to keep Housing First as evidence-based practice that should be expanded at the city, state, and federal government levels.

## REFERENCES


Bamberger, J., and Considine-Cortelyou, H. 2008. Changes in the lives of formerly homeless individuals in supportive housing S/A. Conference Presentation to the 136th APHA Annual Meeting and Exposition 2008

Buchanan, David, Romina Kee, Laura S. Sadowski, and Diana Garcia. 2009. The Health Impact of Supportive Housing for HIV-Positive Homeless Enrollees: A Randomized Controlled Trial. *American Journal of Public Health* 99 (S3): S675–S680

Corporation for Supportive Housing. 2013. Summary of Studies: Medicaid/Health Services Utilization and Costs. Providence, Rhode Island: CSH

Culhane, Dennis P., Stephen Metraux, and Trevor Hadley. 2002. Public Service Reductions Associated with Placement of Homeless Persons with Severe Mental Illness in Supportive Housing. Housing Policy Debates, Volume 13 (1): 107–163

Dickey, Elizabeth (2015). When California's Medi-Cal Will Pay for a Nursing Home, Assisted Living, or Home Care. [Available at http://www.nolo.com/legal-encyclopedia/when-californias-medi-cal-will-pay-nursing-home-assisted-living-home-health-care.html.]

Flaming, Daniel, Patrick Burns, Michael Matsunaga, Gerald Sumner, Manuel H. Moreno, Halil Toros, and Duc Doan. 2009. Where We Sleep: Costs when Homeless and Housed in Los Angeles: Los Angeles: ECONOMIC ROUNDTABLE

Hirsch, Eric, Irene Glasser, Roger Williams, Kate D'Addabbo, and Jessica Cigna. 2008. Rhode Island's Housing First Program Evaluation. United Way of Rhode Island and State of Rhode Island

Hoult, J., I. Reynolds, M. Charbonneau-Powis, P. Weekes, and J. Briggs. 1983. Psychiatric hospital versus community treatment: The results of a randomized trial. *Australian and New Zealand Journal of Psychiatry*, 17,160–167

Larimer, Mary E., Daniel K. Malone, Michelle D. Garner, David C. Atkins, Bonnie Burlingham, Heather S. Lonczak, Kenneth Tanzer, Joshua Ginzler, Seema L. Clifasefi, William G. Hobson, and G. Alan Marlatt. 2009. Health Care and Public Service Use and Costs Before and After Provision of Housing for Chronically Homeless Persons with Severe Alcohol Problems. *JAMA*. 301(13):1349–1357

Linkins, Karen W., Jennifer, J. Brya, and Daniel W. Chandler. 2008. Frequent Users of Health Services Initiative: Final Evaluation Report. California: The Lewin Group

Mackelprang, Jessica L., Susan E. Collins, and Seema L. Clifasefi. 2014. Housing First Is Associated with Reduced Use of Emergency Medical Services. *Prehospital Emergency Care* 18(4): 476–482

MHSA. 2007. Home and Healthy for Good: A Statewide Pilot Housing First Program. Preliminary Report February 2007. Boston: MHSA

MHSA, 2015. Home and Healthy for Good. Permanent Supportive Housing: A Solution-Driven Model. December 2015 Home and Healthy for Good Progress Report. Boston: MHSA

Mondello, Mellany, Anne B. Gass, Thomas McLaughlin, and Nancy Shore. 2007. Cost of Homelessness: Cost Analysis of Permanent Supportive Housing. Portland, Maine: The State of Maine

Moore, T.L. 2006. Estimated cost savings following enrollment in the Community Engagement Program: Findings from a pilot study of homeless dually diagnosed adults. Portland, OR: Central City Concern

Mulder, R. (1985). Evaluation of the Harbinger Program, 1982–1985. Lansing, Michigan: Department of Mental Health

Nogaski, Alyssa, Amy Rynell, Amy Terpstra, and Helen Edwards. 2009. Supportive Housing in Illinois: A Wise Investment. Illinois: The Heartland Alliance Mid-America Institute on Poverty

O'Connell, J.J. and S. Swain. 2005. Rough Sleepers: A Five-Year Prospective Study in Boston, 1999–2003. Presentation, Tenth Annual Ending Homelessness Conference, Massachusetts Housing & Shelter Alliance, Waltham, MA

Oscar-Berman, Marlene, and Ksenija Marinković. 2007. Alcohol: Effects on Neurobehavioral Functions and the Brain. *Neuropsychology Review* (2007) 17:239–257

Pearlman, Jennifer, and John Parvensky. 2006. Denver Housing First Collaborative Cost Benefit Analysis and Program Outcomes Report. Colorado: Colorado Coalition for the Homeless

Srebnik, D., T. Connor, and L. Sylla. 2013. A pilot study of the impact of housing first-supported housing for intensive users of medical hospitalization and sobering services. *American Journal of Public Health*, 103(2), 316–321. doi:10.2105/AJPH.2012.300867

Stein, L., and M.A. Test. 1980. Alternative to mental hospital treatment I: Conceptual model, treatment program, and clinical evaluation. *Archives of General Psychiatry*, 37, 392–397

Test, M.A. 1992. Training in community living. In R.P. Liberman (ed.), *Handbook of psychiatric rehabilitation*, 153–170. New York: Macmillan Publishing Company

Test, M.A. and L. Stein. 1980. Alternative to mental hospital treatment III: Social cost. *Archives of General Psychiatry*, 37, 409–412

Tsai, Jack, Lillian Gelberg, and Robert A. Rosenheck. 2019. Changes in Physical Health After Supported Housing: Results from the Collaborative Initiative to End Chronic Homelessness. *Journal of General Internal Medicine*, 34(9):1703–1708

Walker, Alissa. 2019. Trump's new homeless czar a 'real-life horror.' Curbed Affordable Housing News [Available at www.curbed.com/2019/12/10/21003178/robert-marbut-homelessness-director-trump]

Weisbrod, B. A., Test, M. A., & Stein, L. I. (1980). Alternative to mental hospital treatment II: Economic benefit-cost analysis. Archives of General Psychiatry, 37(4), 400–405

Wright, Bill J., Keri B. Vartanian, Hsin-Fang Li, Natalie Royal, and Jennifer K. Matson. 2016. Formerly Homeless People Had Lower Overall Health Care Expenditures After Moving into Supportive Housing. *Health Affairs*, 35, 20–27

# APPENDIX A. UNMATCHED AND MATCHED BASELINE[1] CHARACTERISTICS

| | Unmatched[2] | | | Matched[2] | | |
|---|---|---|---|---|---|---|
| | HF Cohort N = 1,342 | Control Cohort N = 44,022 | P-value[3] | HF Cohort N = 690 | Control Cohort N = 690 | P-value[3] |
| **Age as of enrollment** | | | | | | |
| Mean, SD | 47.8 ±11.0 | 35.6 ±14.4 | <0.001* | 47.8 ±10.3 | 48.7 ±11.7 | 0.060 |
| Median | 50.0 11.0 | 35.0 14.4 | <0.001* | 49.0 11.0 | 50.5 14.4 | 0.060 |
| Under 65, n (%) | 1,301 (96.9%) | 43,269 (98.3%) | <0.001* | 677 (98.1%) | 641 (92.9%) | <0.001* |
| 65+, n (%) | 41 (3.1%) | 753 (1.7%) | <0.001* | 13 (1.9%) | 49 (7.1%) | <0.001* |
| **Race, n (%)** | | | | | | |
| White | 578 (43.1%) | 18,589 (42.2%) | 0.538 | 286 (41.4%) | 289 (41.9%) | 0.871 |
| Black | 192 (14.3%) | 6,160 (14.0%) | 0.744 | 106 (15.4%) | 108 (15.7%) | 0.879 |
| Hispanic | 72 (5.4%) | 3,638 (8.3%) | <0.001* | 39 (5.7%) | 45 (6.5%) | 0.480 |
| Asian | 10 (0.7%) | 221 (0.5%) | 0.218 | 5 (0.7%) | 3 (0.4%) | 0.480 |
| Indian | 5 (0.4%) | 183 (0.4%) | 0.809 | 2 (0.3%) | 2 (0.3%) | 1.000 |
| Inter | 4 (0.3%) | 352 (0.8%) | 0.040* | 3 (0.4%) | 2 (0.3%) | 0.655 |
| Unknown | 481 (35.8%) | 14,879 (33.8%) | 0.119 | 249 (36.1%) | 241 (34.9%) | 0.657 |
| Male, n (%) | 949 (70.7%) | 24,271 (55.1%) | <0.001* | 492 (71.3%) | 498 (72.2%) | 0.703 |
| **Behavioral health disorders, n (%)** | | | | | | |
| Mental illness | | | | | | |
| Any mental illness | 1,219 (90.8%) | 32,256 (73.3%) | <0.001* | 659 (95.5%) | 642 (93.0%) | 0.047* |
| Schizophrenia | 363 (27.0%) | 7,625 (17.3%) | <0.001* | 194 (28.1%) | 302 (43.8%) | <0.001* |
| Bipolar disorders | 408 (30.4%) | 10,088 (22.9%) | <0.001* | 205 (29.7%) | 278 (40.3%) | <0.001* |
| Depression | 890 (66.3%) | 17,687 (40.2%) | <0.001* | 478 (69.3%) | 424 (61.4%) | 0.002* |
| Anxiety | 694 (51.7%) | 13,594 (30.9%) | <0.001* | 367 (53.2%) | 328 (47.5%) | 0.028* |
| Substance use | | | | | | |
| Any substance use disorders | 1,072 (79.9%) | 22,936 (52.1%) | <0.001* | 586 (84.9%) | 522 (75.7%) | <0.001* |
| Alcohol use disorder | 697 (51.9%) | 11,144 (25.3%) | <0.001* | 396 (57.4%) | 374 (54.2%) | 0.223 |
| Drug use disorder | 935 (69.7%) | 18,092 (41.1%) | <0.001* | 480 (69.6%) | 419 (60.7%) | <0.001* |
| **Charlson Comorbidity Index, mean ± SD** | 1.6 ±2.0 | 0.9 ±1.6 | <0.001* | 1.7 ±2.0 | 1.9 ±2.2 | 0.243 |
| Components, n (%) | | | | | | |
| Myocardial infarction | 81 (6.0%) | 978 (2.2%) | <0.001* | 47 (6.8%) | 42 (6.1%) | 0.583 |
| Congestive heart failure | 92 (6.9%) | 1,337 (3.0%) | <0.001* | 53 (7.7%) | 56 (8.1%) | 0.756 |
| Peripheral vascular disease | 75 (5.6%) | 1,267 (2.9%) | <0.001* | 44 (6.4%) | 62 (9.0%) | 0.066 |
| Cerebrovascular disease | 103 (7.7%) | 1,692 (3.8%) | <0.001* | 61 (8.8%) | 88 (12.8%) | 0.022* |
| Dementia | 15 (1.1%) | 458 (1.0%) | 0.783 | 7 (1.0%) | 34 (4.9%) | <0.001* |
| Chronic pulmonary disease | 486 (36.2%) | 10,918 (24.8%) | <0.001* | 287 (41.6%) | 293 (42.5%) | 0.745 |
| Rheumatologic disease | 26 (1.9%) | 562 (1.3%) | 0.035* | 18 (2.6%) | 11 (1.6%) | 0.194 |
| Peptic ulcer disease | 27 (2.0%) | 506 (1.1%) | 0.004* | 18 (2.6%) | 19 (2.8%) | 0.869 |
| Mild liver disease | 365 (27.2%) | 5,871 (13.3%) | <0.001* | 214 (31.0%) | 176 (25.5%) | 0.023* |
| Mild to moderate diabetes | 144 (10.7%) | 3,022 (6.9%) | <0.001* | 82 (11.9%) | 94 (13.6%) | 0.337 |
| Diabetes with chronic complications | 84 (6.3%) | 1,467 (3.3%) | <0.001* | 38 (5.5%) | 59 (8.6%) | 0.029* |
| Hemiplegia or paraplegia | 26 (1.9%) | 470 (1.1%) | 0.003* | 21 (3.0%) | 18 (2.6%) | 0.602 |
| Renal disease | 71 (5.3%) | 1,018 (2.3%) | <0.001* | 34 (4.9%) | 49 (7.1%) | 0.079 |
| Any malignancy including lymphoma and leukemia | 48 (3.6%) | 868 (2.0%) | <0.001* | 22 (3.2%) | 23 (3.3%) | 0.879 |
| Moderate or severe liver disease | 24 (1.8%) | 440 (1.0%) | 0.005* | 12 (1.7%) | 25 (3.6%) | 0.033* |
| Metastatic solid tumor | 11 (0.8%) | 169 (0.4%) | 0.012* | 7 (1.0%) | 9 (1.3%) | 0.617 |
| HIV/AIDS | 51 (3.8%) | 1,021 (2.3%) | <0.001* | 25 (3.6%) | 30 (4.3%) | 0.492 |

*continued*

| | Unmatched[2] | | | Matched[2] | | |
|---|---|---|---|---|---|---|
| | HF Cohort N = 1,342 | Control Cohort N = 44,022 | P-value[3] | HF Cohort N = 690 | Control Cohort N = 690 | P-value[3] |
| **Medical resource use, by setting** | | | | | | |
| At least one visit, n (%) | | | | | | |
| ER | 1,074 (80.0%) | 27,366 (62.2%) | <0.001* | 597 (86.5%) | 580 (84.1%) | 0.196 |
| Inpatient | 804 (59.9%) | 19,053 (43.3%) | <0.001* | 471 (68.3%) | 460 (66.7%) | 0.510 |
| Mental health encounter | 1,049 (78.2%) | 22,867 (51.9%) | <0.001* | 559 (81.0%) | 429 (62.2%) | <0.001* |
| Other | 1,266 (94.3%) | 37,950 (86.2%) | <0.001* | 681 (98.7%) | 671 (97.2%) | 0.050* |
| Visits, mean ± SD | | | | | | |
| ER | 7.3 ±11.5 | 3.4 ±6.4 | <0.001* | 8.6 ±12.3 | 8.2 ±11.4 | 0.622 |
| Inpatient | 3.9 ±7.2 | 2.1 ±5.1 | <0.001* | 4.7 ±7.6 | 4.9 ±7.8 | 0.918 |
| Mental health encounter | 21.5 ±26.9 | 11.4 ±21.1 | <0.001* | 23.9 ±27.5 | 17.0 ±27.5 | <0.001* |
| Other | 33.6 ±30.0 | 26.0 ±30.3 | <0.001* | 39.1 ±30.4 | 39.8 ±35.3 | 0.796 |
| Days, mean ± SD | | | | | | |
| Inpatient | 20.2 ±36.8 | 10.5 ±28.3 | <0.001* | 23.5 ±37.7 | 25.5 ±44.6 | 0.951 |
| Expenditures, mean ± SD | | | | | | |
| Total (Medical + Pharmacy) | $45,131 ±$68,562 | $22,177 ±$49,829 | <0.001* | $52,831 ±$73,146 | $54,993 ±$80,858 | 0.455 |
| Medical | $41,738 ±$66,191 | $20,225 ±$47,601 | <0.001* | $48,918 ±$71,033 | $50,834 ±$78,267 | 0.510 |
| • ER | $5,093 ±$11,013 | $1,699 ±$4,728 | <0.001* | $5,696 ±$11,546 | $4,238 ±$8,001 | <0.001* |
| • Inpatient | $19,001 ±$42,254 | $7,826 ±$24,867 | <0.001* | $22,619 ±$46,362 | $18,064 ±$36,363 | 0.009* |
| • Other | $17,644 ±$29,185 | $10,700 ±$32,651 | <0.001* | $20,602 ±$30,423 | $28,531 ±$63,084 | 0.132 |
| Mental health encounter | $27,175 ±$49,599 | $9,647 ±$27,735 | <0.001* | $30,226 ±$52,967 | $22,289 ±$43,546 | <0.001* |
| Pharmacy | $3,393 ±$11,874 | $1,953 ±$9,691 | <0.001* | $3,913 ±$12,068 | $4,159 ±$14,315 | 0.339 |

1. The index date for the intervention cohort ("Housing First") is the date of initial enrollment in the program, while the index date for the control cohort is a randomly selected day during their eligibility period. The baseline period is defined as the two years prior to the index date.
2. Patients were matched based on ± 0.25*standard deviation of propensity score of being in their cohort. Propensity score controlled for age, gender, categorical year of index date, race, CCI, and total baseline healthcare costs. Among the matched analysis, both HF and control members were required to have full baseline and follow-up eligibility, and each control member has at least 2 distinct claims with a primary diagnosis code for homelessness. Control patients with evidence of CSPECH services were excluded.
3. P-values were calculated using Wilcoxon rank sum tests for continuous variables and chi-squared tests for categorical variables for the unmatched analysis. P-values were calculated using Wilcoxon signed-rank tests for continuous variables and McNemar tests for categorical variables for the matched analysis. P-values <0.05 are indicated with one asterisk ("*").

Abbreviations: CCI, Charlson Comorbidity Index; CSPECH, Community Support Program for People Experiencing Chronic Homelessness; ER, Emergency room; SD, Standard deviation.

# APPENDIX B. UNMATCHED AND MATCHED FOLLOW-UP[1] CHARACTERISTICS

| | Unmatched[2] | | | Matched[2] | | |
|---|---|---|---|---|---|---|
| | HF Cohort N = 1,342 | Control Cohort N = 44,022 | P-value[3] | HF Cohort N = 690 | Control Cohort N = 690 | P-value[3] |
| **FULL YEAR** | | | | | | |
| ***At least one visit, n (%)*** | | | | | | |
| ER | 624 (46.5%) | 22,689 (51.5%) | <0.001* | 437 (63.3%) | 501 (72.6%) | <0.001* |
| Inpatient | 399 (29.7%) | 14,360 (32.6%) | 0.026* | 284 (41.2%) | 361 (52.3%) | <0.001* |
| Mental health encounter | 891 (66.4%) | 19,579 (44.5%) | <0.001* | 552 (80.0%) | 366 (53.0%) | <0.001* |
| Other | 1,059 (78.9%) | 36,200 (82.2%) | 0.002* | 665 (96.4%) | 652 (94.5%) | 0.091 |
| ***Visits, mean ± SD*** | | | | | | |
| ER | 1.9 ±4.3 | 1.9 ±3.7 | 0.011* | 3.0 ±5.3 | 4.5 ±6.4 | <0.001* |
| Inpatient | 1.2 ±4.0 | 1.1 ±3.0 | 0.026* | 1.8 ±4.7 | 2.6 ±4.5 | <0.001* |
| Mental health encounter | 7.7 ±13.3 | 6.1 ±12.0 | <0.001* | 11.2 ±15.5 | 9.3 ±15.1 | <0.001* |
| Other | 10.4 ±14.2 | 13.9 ±16.9 | <0.001* | 15.3 ±16.1 | 20.9 ±19.5 | <0.001* |
| ***Days, mean ± SD*** | | | | | | |
| Inpatient | 5.1 ±14.5 | 5.7 ±16.8 | 0.038* | 7.7 ±17.7 | 13.4 ±26.7 | <0.001* |
| Expenditures, mean ± SD | | | | | | |
| Total (Medical + Pharmacy) | $17,336 ±$29,956 | $12,220 ±$28,388 | <0.001* | $25,614 ±$36,118 | $30,881 ±$49,643 | 0.316 |
| Medical | $15,305 ±$26,524 | $11,133 ±$27,140 | <0.001* | $22,717 ±$32,029 | $29,072 ±$48,268 | 0.103 |
| • ER | $1,399 ±$4,135 | $951 ±$2,821 | 0.643 | $2,082 ±$5,207 | $2,520 ±$5,348 | 0.055 |
| • Inpatient | $5,441 ±$16,751 | $4,375 ±$15,803 | 0.526 | $8,233 ±$20,773 | $10,359 ±$23,839 | 0.034* |
| • Other | $8,465 ±$13,247 | $5,807 ±$17,596 | <0.001* | $12,402 ±$16,165 | $16,193 ±$37,229 | 0.139 |
| Mental health encounter | $12,613 ±$22,997 | $5,393 ±$17,196 | <0.001* | $18,240 ±$27,407 | $13,529 ±$27,530 | <0.001* |
| Pharmacy | $2,031 ±$10,355 | $1,086 ±$5,929 | 0.014* | $2,896 ±$12,827 | $1,809 ±$7,997 | <0.001* |
| **FIRST 6 MONTHS** | | | | | | |
| Total (Medical + Pharmacy) | $9,554 ±$17,867 | $6,087 ±$15,653 | <0.001* | $13,294 ±$21,486 | $14,781 ±$26,352 | 0.857 |
| Medical | $8,371 ±$15,391 | $5,540 ±$14,991 | <0.001* | $11,627 ±$18,212 | $13,668 ±$25,153 | 0.917 |
| • ER | $782 ±$2,465 | $479 ±$1,624 | 0.006* | $1,096 ±$3,062 | $1,160 ±$2,753 | 0.250 |
| • Inpatient | $2,836 ±$9,360 | $2,172 ±$9,309 | 0.487 | $4,093 ±$10,992 | $4,900 ±$14,134 | 0.276 |
| • Other | $4,752 ±$7,904 | $2,889 ±$9,559 | <0.001* | $6,437 ±$9,505 | $7,608 ±$18,607 | 0.007* |
| Mental health encounter | $7,028 ±$13,322 | $2,670 ±$9,675 | <0.001* | $9,471 ±$15,138 | $6,303 ±$15,501 | <0.001* |
| Pharmacy | $1,183 ±$7,944 | $547 ±$3,308 | <0.001* | $1,667 ±$10,276 | $1,113 ±$6,555 | <0.001* |
| **FIRST 6 MONTHS** | | | | | | |
| Total (Medical + Pharmacy) | $7,782 ±$16,665 | $6,100 ±$16,136 | 0.002* | $12,320 ±$19,895 | $15,995 ±$27,711 | 0.203 |
| Medical | $6,934 ±$15,398 | $5,563 ±$15,470 | <0.001* | $11,090 ±$18,952 | $15,305 ±$27,255 | 0.044* |
| • ER | $616 ±$2,116 | $470 ±$1,626 | 0.025* | $986 ±$2,700 | $1,351 ±$3,276 | 0.007* |
| • Inpatient | $2,605 ±$11,207 | $2,191 ±$9,733 | 0.020* | $4,139 ±$14,384 | $5,420 ±$14,389 | 0.004* |
| • Other | $3,713 ±$6,827 | $2,902 ±$9,755 | <0.001* | $5,965 ±$8,278 | $8,533 ±$20,358 | 0.275 |
| Mental health encounter | $5,570 ±$13,522 | $2,675 ±$10,162 | <0.001* | $8,742 ±$16,660 | $6,990 ±$15,376 | <0.001* |
| Pharmacy | $848 ±$4,970 | $537 ±$3,513 | 0.620 | $1,230 ±$4,987 | $690 ±$2,543 | 0.001* |

1. The index date for the intervention cohort ("Housing First") is the date of initial enrollment in the program, while the index date for the control cohort is a randomly selected day during their eligibility period. The followup period is defined as the one year following the index date..
2. Patients were matched based on ± 0.25*standard deviation of propensity score of being in their cohort. Propensity score controlled for age, gender, categorical year of index date, race, CCI, and total baseline healthcare costs. Among the matched analysis, both HF and control members were required to have full baseline and follow-up eligibility, and each control member has at least 2 distinct claims with a primary diagnosis code for homelessness. Control patients with evidence of CSPECH services were excluded.
3. P-values were calculated using Wilcoxon rank sum tests for continuous variables and chi-squared tests for categorical variables for the unmatched analysis. P-values were calculated using Wilcoxon signed-rank tests for continuous variables and McNemar tests for categorical variables for the matched analysis. P-values <0.05 are indicated with one asterisk ("*").

Abbreviations: CCI, Charlson Comorbidity Index; CSPECH, Community Support Program for People Experiencing Chronic Homelessness; ER, Emergency room; SD, Standard deviation.

MASSACHUSETTS

FOUNDATION