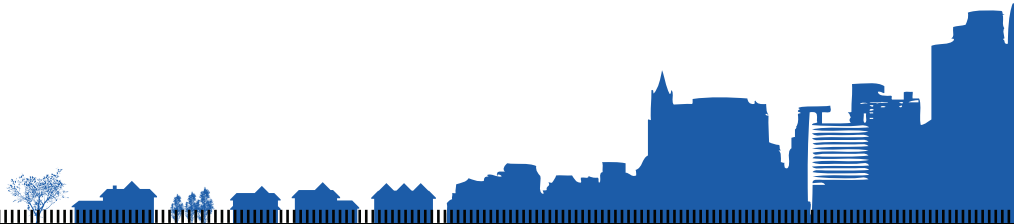


Research Design and Analysis Plan for Impact Study

Evaluation of the Supportive Services Demonstration



1.	Introduction	2
1.1	Site Selection and Random Assignment.....	3
1.2	Overview of IWISH Model.....	6
1.3	Evaluation Components.....	7
2.	Research Questions and Outcome Measures	8
2.1	Research Question 1: Impact of IWISH on Utilization of Unplanned Hospitalizations and Other Acute Care	10
2.2	Research Question 2: Impact of IWISH on Utilization of Primary Care and Other Non-Acute Healthcare Services	11
2.3	Research Questions 3 and 4: Impact of IWISH on Housing Exits and Impact on Transfers to Long-Term Care Facilities	13
2.4	Cross-Cutting Measures	15
2.5	Patient, Property, and Community Characteristics.....	16
3.	Data Sources	18
3.1	HUD Data	18
3.2	CMS Medicare Data	19
3.3	State Medicaid Data.....	22
3.4	Public Use Data	23
3.5	Linking Data Sets	24
4.	Analysis Approach	25
4.1	Experimental Sample.....	25
4.2	Experimental and Quasi-Experimental Analyses	26
5.	Impact Study Limitations	35
	References	37

1. Introduction

This document presents the research design and analysis plan for the impact study component of the Supportive Services Demonstration (SSD) Evaluation. SSD is a three-year demonstration sponsored by the U.S. Department of Housing and Urban Development (HUD) to test the impact of housing-based supportive services on the healthcare utilization and housing stability of low-income adults aged 62 and over living in HUD Office of Multifamily Housing-assisted properties in seven states.

The model being tested is called Integrated Wellness in Supportive Housing (IWISH). The IWISH model is implemented by an enhanced service coordinator, known as a Resident Wellness Director (RWD), and a Wellness Nurse (WN), both of whom work at the property. The IWISH model funds these staff with the goal of connecting residents to the appropriate services in the community to support them as they age in place. IWISH also provides guidelines for how the RWD and WN work with residents, emphasizing a person-centered approach with a detailed health and wellness assessment and service tracking through a web-based client management system.

HUD's Office of Policy Development and Research designed the evaluation as a cluster randomized trial. Within seven states, HUD randomly assigned eligible properties to treatment and control groups, in which 40 properties implement the IWISH model (the treatment) and 84 properties conduct business as usual. The evaluation monitors the implementation of IWISH over a three year period and compares housing and healthcare utilization outcomes for residents living in the treatment group properties to those in the control group properties. HUD contracted with Abt Associates Inc. to conduct the evaluation.

Motivating the SSD evaluation is HUD's interest in identifying housing-based supportive service models that promote successful aging in place, produce savings to the Medicare and Medicaid systems, and are viable from a property management perspective. Identifying such models is critical for meeting the needs of the growing population of low income, older Americans. Key measures of IWISH's success will be whether the model reduces hospitalizations, promotes the use of appropriate services such as Medicare or Medicaid covered primary care or home and community based services (HCBS), and delays transfers to costly settings such as nursing homes and other long-term care facilities. All of these outcomes reduce the cost burden to Medicare and Medicaid of low-income elderly residents of HUD multifamily housing. The evaluation will also test for impacts on housing stability.

HUD established an expert panel for the evaluation to help ensure that the study findings are relevant to the field and provide critical information to Congress and other policy makers. The panel's role is to guide the study team in producing reliable, credible, quantitative evidence for Congress and stakeholders about the impact of the SSD model on costly healthcare utilization and transitions to nursing home care. As of the writing of this report, the expert panel members (in alphabetical order) are:

- **Mara Blitzer**, Director of Housing Development, Mayor's Office of Housing and Community Development, City and County of San Francisco
- **Melanie Brown**, Technical Director, Division of Community Systems Transformation, Disabled and Elderly Health Programs Group, CMCS, CMS

-
- **Bruce Chernof**, President and CEO, SCAN foundation
 - **Partha Deb**, Professor of Economics, Hunter College, CUNY
 - **Tim Engelhardt**, Director, Federal Coordinated Health Care Office, CMS
 - **Kosuke Imai**, Professor in the Department of Government and the Department of Statistics at Harvard University
 - **Cindy Mann**, Partner, Manatt Health
 - **Sandy Markwood**, CEO, National Association of Area Agencies on Aging
 - **J. Michael McWilliams**, Professor, Harvard Medical School
 - **Janice C. Monks**, Founder, President/CEO, American Association of Service Coordinators
 - **Emily Rosenoff**, Acting Director, Division of Long-Term Care Policy, Office of Disability, Aging, and Long-Term Care Policy, ASPE
 - **Kamillah Wood**, Senior Vice President of Health and Housing at Stewards of Affordable Housing for the Future (SAHF)

The expert panel has reviewed the full Research Design and Data Collection and Analysis Plan for the evaluation, from which the material in this document is excerpted.

The rest of this chapter provides an overview of the SSD's site selection and random assignment procedures, core elements of the IWISH model, and evaluation components. Chapters 2 through 5 then focus on the impact study, describing the research questions and outcome measures, data sources, analysis approach, and study limitations.

1.1 Site Selection and Random Assignment

Selection of properties to participate in the SSD began in January 2016 when HUD published a Notice of Funding Availability (NOFA) announcing the availability of \$15 million in funds for the Supportive Services Demonstration for Elderly Households in HUD-Assisted Multifamily Housing and inviting owners of eligible federally assisted multifamily properties to apply.

HUD screened more than 700 applications for eligibility and ranked states for the demonstration based on three factors: the number of properties with and without a current service coordinator; the rate of county-level fee-for-service enrollment among Medicare beneficiaries; and the number of initiatives funded by the Centers for Medicare & Medicaid Services (CMS) to improve care coordination for elderly persons. In each case, a higher number gave a state a greater chance of being selected for the demonstration.

Based on the rankings, HUD identified 182 properties across seven states (California, Illinois, Massachusetts, Maryland, Michigan, New Jersey, and South Carolina) as eligible for random assignment. HUD determined that it could fund up to 40 properties to implement the demonstration approach (treatment sites) and limited the demonstration to seven states in order not to spread the treatment sites too thinly and to permit a sufficient number of control sites in each state.

Prior to random assignment, HUD stratified the properties by Community-Based Statistical Area (CBSA) to help ensure that the treatment and control groups would be balanced on characteristics

such as access to and cost of healthcare services and access to social services that could affect demonstration outcomes and would be expected to vary by CBSA. HUD also assigned “importance weights” to each property within each stratum. The importance weights took into account the rate of Medicare fee-for-service (FFS) participation for its county and the property’s budget request in the response to the NOFA. HUD used the importance weights to help ensure that the final sample would have higher-than-average FFS participation rates and somewhat lower-than-average budget requests. HUD prioritized counties with high FFS participation rates because of the importance of Medicare FFS claims data as a data source for the evaluation. The budget request was a consideration because of constraints on the overall amount of funding available to fund IWISH implementation at 40 properties.

HUD ranked the properties using the importance weights and then selected a total of 131 properties across the strata.¹ HUD then used random sampling to allocate the selected properties in each stratum into treatment, active control, and passive control groups. The three groups play different roles in the demonstration and evaluation:

- **Treatment Group.** The 40 properties in the treatment group enter into cooperative agreements with HUD to implement the IWISH model. They receive grant funds that primarily pay for hiring one full-time RWD and one part-time WN for roughly every 100 residents of the property. The RWD and WN work on site for the three-year demonstration period. Treatment group properties also receive training and technical assistance from The Lewin Group, the implementation contractor for the SSD, and \$15 per unit per month to support programming.
- **Active Control Group.** The 40 properties in the active control group do not implement the IWISH model but enter into cooperative agreements with HUD to receive \$5,000 to participate in the evaluation.²
- **Passive Control Group.** The 48 properties in the passive control group do not enter into cooperative agreements with HUD for the SSD. HUD simply notified them that they did not receive SSD grant funding.

Exhibit 1-1 shows the distribution of the 131 properties after random assignment. Of the 43 properties assigned to the treatment group, 40 were assigned directly to the treatment group and 3 were assigned to a waitlist for the treatment group. Similarly, of the 48 properties assigned to the passive control group, 10 were designated as waitlist properties for the active control group. The waitlist properties were numbered so that replacements would be chosen in the order they were selected. HUD’s desired distribution for the final sample was 40 treatment properties, 40 active control properties, and as many in the passive control group as possible.

¹ The number of properties selected per stratum was equal to: $(3 \times 44) / 182 \times (\text{number of properties in the stratum})$, rounded to the nearest integer. Due to rounding, 131 properties were selected instead of 132 ($= 3 \times 44$).

² Only available to properties that entered into a cooperative agreement with HUD before October 1, 2017.

Exhibit 1-1. Distribution of Study Properties at Time of Random Assignment

Group	Properties	Notes
Treatment	43	Three of the 43 properties were designated as waitlist properties for the treatment group.
Active control	40	
Passive control	48	10 of the 48 properties were designated as waitlist properties for the active control group.
Total	131	

SOURCE: Deb 2017, p. 4.

At the time of random assignment, the treatment and control properties were found to be well balanced across the key characteristics that were available to HUD at the time and of significance to the evaluation:

- Presence of an onsite service coordinator (self-reported in the application, and not necessarily funded through HUD's Service Coordinator Program).
- Properties with a Project Rental Assistance Contract (PRAC) (PRAC properties may have different characteristics than non-PRAC properties).
- Property size (number of units at the time of application).
- County-level FFS penetration rate.

After testing that the sample was balanced across these four property characteristics, HUD determined that random assignment was complete and announced the SSD grantees in January 2017. Between January 2017 and July 2018, HUD negotiated cooperative agreements with the owners of the 40 treatment group properties and 40 active control group properties. Five properties originally assigned to the treatment and active control groups declined to participate. HUD replaced them as follows:

- **Treatment group:** Two of the 40 properties originally assigned to the treatment group declined to participate. HUD contacted the first two properties on the treatment group waitlist to replace these two properties. The first waitlist property also declined, so HUD contacted the third property on the wait list. Thus, the final treatment group consists of 38 properties originally assigned to the treatment group, plus the second and third properties on the treatment group waitlist.
- **Active control group:** Four of the 40 properties originally assigned to the active control group declined to participate. HUD entered into a cooperative agreements with the 36 properties from the original list and the first property from the active control waitlist before October 1, 2017 and was able to offer these 37 properties the \$5,000 research incentive. Recruiting the remaining three properties took longer, and HUD was not able to offer the \$5,000 because the funds were only available for agreements signed prior to October 1, 2017. Between October 2017 and July 2018, HUD entered into cooperative agreements with three properties from the waitlist. Thus, the final treatment group consists of 36 properties originally assigned to the active control group, plus the first, second, third, and fifth property

on the waitlist. (The fourth waitlist property was contacted but declined to be part of the active control group and thus remained in the passive control group.)

- **Passive control group:** Of the 48 properties originally assigned to the passive control group, four became active control sites as described above. Thus, the final passive control group consists of 44 properties.

Exhibit 1-2 shows the final distribution of properties across the three groups.

Exhibit 1-2. Final Distribution of Study Properties

Group	Properties	Notes
Treatment	40	Includes two properties from the treatment group waitlist.
Active control	40	Includes four properties from the active control group waitlist.
Passive control	44	The original 48 properties minus the four properties that became active control properties.
Total	124	

SOURCE: HUD Supportive Services Demonstration Site List, December 4, 2017.

1.2 Overview of IWISH Model

The IWISH model has six foundational components:

- **Resident engagement to maximize participation:** RWDs and WNs begin conducting outreach to and building relationships with residents before the formal start of IWISH enrollment. During enrollment, the RWDs and WNs meet residents one on one and in groups to make sure they understand what the program has to offer and are motivated to participate.
- **Standardized assessment:** The RWD and WN offer all IWISH participants a standardized health and wellness assessment when they enroll and at least annually until the end of the program. In addition to the standardized assessment, RWDs and WNs conduct person-centered interviews with residents to understand their needs and priorities more comprehensively.
- **Individualized plans:** Each participant is offered the opportunity to work with the RWD and WN to develop an Individualized Healthy Aging Plan (IHAP) that reflects their needs and preferences. The IHAP identifies actionable goals as well as the coordination participants will receive from the resident wellness staff. The RWD and WN also create a Community Health Aging Plan (CHAP) for the property based on the most common needs of residents.
- **Centralized data platform:** IWISH sites are required to use the web-based Population Health Logistics (PHL) system to track information about IWISH participants, including assessments, IHAPs, and the programming and service coordination usage. PHL is designed to allow wellness staff to track the needs, priorities, and progress of each IWISH participant. PHL builds on an existing platform that has been tailored to the IWISH demonstration. None of the IWISH sites used PHL prior to the demonstration. None of the control properties will use the PHL system.
- **Local partnerships:** IWISH staff are encouraged to form partnerships with local health and social service providers. IWISH staff are encouraged to coordinate services with these partners in order to enhance opportunities and resources for their residents.

-
- **Evidence-based programming:** With the assistance of the implementation team, resident wellness staff at each property are encouraged to identify and implement evidence-based health and wellness programs to address the needs of their residents. Evidence-based means the program has been found to be effective based on rigorous evaluation.

Primary responsibility for implementing these six foundational components lies with the two IWISH staff—the RWD and the WN. The RWD and WN share many responsibilities, including resident engagement, assessments, and individualized plans. However, RWDs and WNs each take primary responsibility for certain tasks. The RWD oversees all activities, coordinates with property management, and is responsible for developing community partnerships. The WN is more narrowly focused on residents’ health needs and goals.

1.3 Evaluation Components

The evaluation of the HUD SSD began in October 2017 and will continue through March 2022. The evaluation has a qualitative component—the process study—designed to learn how treatment group properties implemented the IWISH model and how property staff and residents responded to it. The main data sources for the process study are interviews with staff at the IWISH and active control properties and focus groups with residents.

The study also has a quantitative component—the impact study—designed to measure the effect of the intervention on key outcomes related to residents’ use of healthcare services as well as housing stability. *This document focuses on the impact study.* The impact study examines early evidence of the impact of the IWISH model on healthcare utilization, housing stability, and transfers to institutional settings as residents’ age. The analysis will measure key outcomes using HUD administrative data on housing and residents, linked to data on healthcare utilization and spending from the Centers for Medicare and Medicaid Services (CMS) and State Medicaid agencies. The research team will seek to obtain these data for all Medicare and Medicaid beneficiaries residing at IWISH, active control, and passive control properties.

The main analyses will use the cluster-randomized design of the demonstration and estimate the impact of the IWISH model as the difference between the average outcomes among residents at IWISH properties and the average outcomes among similar residents in both control groups. The research team will use multivariate regressions to control for differences in resident and property characteristics that may remain after the properties were randomized.

2. Research Questions and Outcome Measures

The goal of the impact study is to determine whether there is evidence that the IWISH model supports aging in place and affects the use of Medicare and Medicaid covered healthcare services among low-income, elderly residents receiving HUD assistance. For this reason, residents receiving HUD assistance are the unit of analysis, and the property characteristics or other variables measured at a higher level are attached to individual residents (because random assignment was at the cluster level, cluster-robust standard errors are used throughout to account for any possible correlations across individuals). Four primary research questions guide the impact study:

1. What is the impact of IWISH on utilization of Medicare and Medicaid covered unplanned hospitalizations and other acute care?
2. What is the impact of IWISH on utilization of Medicare and Medicaid covered primary care and other non-acute healthcare services?
3. What is the impact of IWISH on housing exits and resident tenure?
4. What is the impact of IWISH on transitions to long-term institutional care?³

Each of these research questions can be evaluated using multiple outcome measures. Testing the impact of an intervention on a large number of outcomes introduces an issue commonly referred to as the “multiple testing problem.” Some outcomes will have statistically significant estimated impacts only by chance, and this is more likely to happen across a large number of estimated outcomes. To avoid the appearance of “cherry-picking” results from among a large number of outcomes to support claims of IWISH’s success, the study team specified four key outcome measures as “confirmatory” when drawing conclusions about IWISH’s impact. The team, in consultation with the expert panel for the evaluation, determined that the four confirmatory outcomes selected were most important for answering the four research questions and for determining whether the goals of IWISH are met. In addition to the confirmatory outcomes, each research question also has a set of “secondary” outcomes. Secondary outcomes are additional important indicators tied to the logic of how the IWISH model is expected to influence outcomes. We will include both confirmatory and secondary measures in our impact analyses.⁴

Exhibit 2-1 shows the full set of outcome measures to be analyzed for each research question, with the confirmatory measure shown in the second column and the secondary measures in the third column. If we find an impact on any of the confirmatory measures, we will be able to provide a definitive answer to that research question. For example, if we find that IWISH has a negative impact

³ The evaluation originally intended to investigate a fifth primary research question: “How does IWISH impact mortality rates?” We do not think we are likely to detect impacts on mortality within the time period of the study. Therefore, we designated mortality as an exploratory outcome rather than a primary outcome for the impact analysis.

⁴ For the purposes of this research design, we reserve the term “exploratory” to differentiate descriptive, non-experimental analyses from the main, “experimental” analyses. Within the experimental analyses, we use the term “secondary” to describe outcomes that are not “confirmatory.”

on the total days of unplanned hospitalization, the confirmatory outcome for the first research question shown in the exhibit, we will be able to confidently state in our report that IWISH reduces utilization of Medicare and Medicaid covered unplanned hospitalizations and other acute care. If we do not find an impact on the confirmatory outcome but find an impact on one or more of the secondary outcomes – for example, positive impacts on ambulance trips and all-cause emergency department (ED) visits not resulting in admission – we will be able to draw some inferences about the potential of IWISH to reduce utilization of high-cost acute care but will be less confident in our conclusions about IWISH’s overall impact.

In the sections that follow, we describe in more detail our approach to answering each research question, as well as how we will analyze the cross-cutting measures shown in the last panel of Exhibit 2-1.

Exhibit 2-1. Impact Study Outcome Measures

Research Question	Confirmatory Outcome Measure	Secondary Outcome Measures
1. What is the impact of IWISH on utilization of Medicare and Medicaid covered unplanned hospitalizations and other acute care?	Total days of unplanned hospitalization	<ul style="list-style-type: none"> • Unplanned hospital admissions • Unplanned hospital readmissions within 30 days of previous hospital discharge • All-cause ED visits not resulting in hospital admission • Ambulance events
2. What is the impact of IWISH on utilization of Medicare and Medicaid covered primary care and other non-acute healthcare services?	Number of days with a primary care visit (i.e., for services related to evaluation & management)	<ul style="list-style-type: none"> • Use of Home and Community Based Services (HCBS) • New use of specialty care services
3. What is the impact of IWISH on housing exits and resident tenure?	Residency ended, for any reason	<ul style="list-style-type: none"> • Residency ended due to death
4. What is the impact of IWISH on transitions to long-term institutional care?	Residency ended due to transition to a long-term care facility	<ul style="list-style-type: none"> • Days in a nursing facility for long-term care • Days in other long-term care facilities
Cross-cutting measures that address multiple research questions	N/A	<ul style="list-style-type: none"> • Days in the community (i.e., not in inpatient care or an institution) • Total medical and pharmacy payments • Total payments inpatient care • Total payments outpatient care • Total pharmacy payments

2.1 Research Question 1: Impact of IWISH on Utilization of Unplanned Hospitalizations and Other Acute Care

An important goal of IWISH is to reduce the cost burden to Medicare and Medicaid of low-income elderly residents of HUD-assisted properties. Unplanned acute-care hospital admissions are a major cost driver, especially among dual-eligible beneficiaries (Erdem, 2013; MedPAC, 2016; Komisar and Feder, 2011). Furthermore, transfers of elderly or disabled people to acute care settings can increase stress and the risk of infection, medical errors, or other trauma or complications. Studies have shown that a substantial portion of hospital admissions and related expenditures are avoidable or preventable (De Brantes et al. 2010; Segal et al., 2014).

A key driver of costs and other negative consequences of unplanned hospitalizations is the amount of time spent in a hospital or emergency care. Thus, the confirmatory measure for research question 1 is the total number of days that residents are admitted to a hospital for unplanned, acute care services during the demonstration, shortened to *total days of unplanned hospitalization*. Findings for this measure might be conveyed as: “During the three-year demonstration period, residents of IWISH properties spent an average of 10 fewer (more) days per year in hospital for unplanned, acute care services than residents of non-IWISH properties,” or “During the three-year demonstration period, IWISH had no impact on the number of days that residents spent in a hospital for unplanned, acute care services.”

For this confirmatory measure and other outcome measures that use the “unplanned” terminology, we will use version 4.0 of the Planned Readmissions Algorithm used in CMS’s Hospital-Wide All-Cause Readmissions measure to classify hospitalizations as planned or unplanned.⁵ We will not attempt to categorize hospital stays (or ED visits) as “unnecessary” or avoidable. Categorizing an inpatient stay or ED visit as unnecessary or avoidable depends not only on patient acuity but also the appropriateness of the site of service and availability of alternate sources of acute, unscheduled care (Honigman et al., 2013; Mistry et al., 2008), information that is not readily available through the data sources for this study.

In addition to the confirmatory outcome of total days of unplanned hospitalization, we will examine four other measures related to the use of acute care services. Exhibit 2-2 presents the outcome measures for research question 1 and the data sources to be used.

⁵ The algorithm (<https://www.cms.gov/Medicare/Quality-Initiatives-Patient-Assessment-Instruments/HospitalQualityInits/Measure-Methodology.html>) uses the principal discharge diagnosis category and all procedure codes on each inpatient claim, coded using the AHRQ ICD-10 Clinical Classification System software (<https://www.ahrq.gov/research/data/hcup/icd10usrgd.html>). Planned admissions are those in which one of a pre-specified list of procedures took place. Admissions to psychiatric hospitals or units are also classified as planned admissions. Admissions for acute illness or for complications of care are never classified as planned. If a typically planned procedure was performed during an admission for an acute illness, the admission is not classified as planned since the procedure would not likely have been planned.

Exhibit 2-2. Outcome Measures for Research Question 1 (● = confirmatory outcome)

Outcome Measure ^a	Data Source
● Total days of unplanned hospitalization	Medicare Inpatient File
Unplanned hospital admissions	
Unplanned hospital readmissions within 30 days of previous hospital discharge	
All-cause ED visits not resulting in hospital admission	Medicare Outpatient File
Ambulance events	Medicare Carrier and Outpatient Files

^a All utilization measures (such as days of unplanned admissions and hospital readmissions) will be specified as the number of days or events per resident per quarter (three months) to accommodate longitudinal analysis, and because we expect some loss to follow up due to resident turnover and mortality.

The first two secondary measures are “unplanned hospital admissions” and “unplanned hospital readmissions within 30 days of previous hospital discharge.” Unplanned hospital admissions measures the overall frequency of hospitalizations, while 30-day hospital readmissions are an important measure of the overall quality of care for hospitalized patients. Fewer hospital readmissions would reflect improvements in the quality of discharge planning and post-acute care received by the residents at IWISH properties after they were discharged from a hospital.

The third and fourth secondary measures are “all-cause ED visits not resulting in hospital admission” and “ambulance events.” Reduced numbers of ED visits and emergency ambulance trips not only represent less utilization of unplanned acute care services but also serve as important intermediate outcomes that could signal the potential for the IWISH model’s long-term success. Not all ED visits and ambulance trips warrant inpatient care, and they are likely to occur more frequently than hospital admissions. Therefore, it is likely that the impact of the IWISH model on these services will be detected before the impact on the utilization of unplanned inpatient services.

2.2 Research Question 2: Impact of IWISH on Utilization of Primary Care and Other Non-Acute Healthcare Services

One of the purposes of IWISH is to identify residents’ unmet needs and connect them to appropriate health and social services in the community, such as primary care, home and community-based services (HCBS), or specialty care services. Early increases in the utilization of these services could lead to higher costs in the short-term but could also indicate the long-run potential of the IWISH model to reduce unnecessary utilization of high-cost services and thus decrease overall healthcare spending. The expert panel recommended that we track utilization of primary care and other non-acute care services as an important intermediate outcome in the event that we do not see an impact on hospitalization rates within the three-year demonstration period.

The confirmatory outcome for Research Question 2, will be the total number of days that the resident had at least one physician office visit for primary care, that is, services related to the evaluation and

management of the patient, shortened to *number of days with a primary care (E&M) visit*.⁶ These findings might be phrased as: “During the three-year demonstration period, residents of IWISH properties made an average of two more (fewer) visits per year to a primary care physician than residents of non-IWISH properties,” or “During the three-year demonstration period, IWISH had no impact on the number of visits that residents made to primary care physicians.”

Exhibit 2-3. Outcome Measures for Research Question 2 (● = confirmatory outcome)

Outcome Measure	Data Source
● Number of days with a primary care (E&M) visit	Medicare Carrier and Durable Medical Equipment (DME) Files
Use of HCBS services	State Medicaid claims and encounter data
New use of specialty care services	Medicare Outpatient, Carrier, DME Files

As shown in Exhibit 2-3, we will also examine two secondary measures related to the use of two types of non-acute care services: HCBS and specialty services. Defining the use of HCBS services is not straightforward. States can apply to CMS for waivers (i.e., 1915(c) waivers) to provide HCBS that meet the needs of people who prefer to get long-term care services and supports in their home and community rather than in an institutional setting. While there are broad federal guidelines, states vary in the covered services and rates approved for the waivers, which could confound comparisons of HCBS utilization in our pooled sample of residents. Furthermore, based on our previous experience with state Medicaid data, it is likely that we will not be able to identify 1915(c) waiver services or other Medicaid-funded HCBS reliably in every state’s data. To address these issues, we will proxy for HCBS in the state Medicaid data by measuring the utilization of a personal care assistant (PCA), a fundamental component of many HCBS programs.⁷

The type of physician specialists included in the definition of specialty care services will be determined after we receive the baseline Medicare data, because the definition should be based on the prevalence of specific underlying conditions observed in the residents’ baseline data. For example, endocrinologists would be included if diagnoses for diabetes mellitus is highly prevalent in the

⁶ Members of the expert panel commented that some successful models of care have been linked to a decline in the use of evaluation and management (E&M) procedure codes in medical claims. It is possible that IWISH residents will actually receive more primary care than before but through non-billable services facilitated by a care coordinator or wellness nurse. Panelists did not recommend an alternative measure of primary care utilization. However, they recommended that when reporting on findings for this measure we include statements that caution the reader around this issue and that we qualify our findings based on what we might learn from the process study.

⁷ We also considered using case management services as a proxy for use of any HCBS. However, case management services may be offered to a broader population, for reasons unrelated to community-based services for people with disabilities. Therefore, we believe PCA use is a better proxy.

baseline sample of claims data.⁸ We will measure *new* visits to a specialist since an increase in the rate of new visits is likely to result from increased access to care due to the IWISH model, while an increase in the rate of existing visits could indicate a worsening of the health condition unrelated to IWISH. A specialist visit will be defined as a “new” visit if the resident did not visit the same type of specialist in the previous year.

2.3 Research Questions 3 and 4: Impact of IWISH on Housing Exits and Impact on Transfers to Long-Term Care Facilities

An important goal of IWISH is to promote aging in place for residents of HUD-assisted properties, especially by delaying transfers to a nursing facility for long-term care. Long-term care is costly and, like transfers to acute-care settings, can have negative consequences for the well-being of an older adult.

Research question 3 focuses on the impact of IWISH on housing exits and tenure. Longer rates of tenure and lower rates of exits are generally desirable outcomes for multifamily housing owners and managers, as there are substantial costs associated with turning units over and residents tend to prefer to age in place. Research question 4 focuses specifically on transitions from living in the community to residing in an institution offering long-term care. Exhibit 2-4 presents the outcome measures for research question 3, and Exhibit 2-5 presents the outcome measures for research question 4.

Exhibit 2-4. Outcome Measures for Research Question 3 (● = confirmatory outcome)

Outcome Measure	Data Source
● Residency ended, for any reason	HUD data
Residency ended due to death	HUD data, Medicare Beneficiary Summary File

The confirmatory outcome for research question 3 is the likelihood that residency ended, for any reason, during the demonstration, shortened to *residency ended for any reason*. To analyze this measure (and the confirmatory measure for research question 4), we will use a discrete-time logistic hazard (or survival) model to measure how often residents of treatment properties moved out of their units compared to how often residents of control group properties moved out of their units (Jenkins 1995).⁹ With this type of analysis, the impact of IWISH will be expressed in terms of the risk of a resident of a treatment property moving out at any time during the demonstration relative to a resident of a control group property. For example, “Residents of IWISH properties were half (twice) as likely

⁸ However, we will not focus on visits to certain specialists by a “prevalence-based” cohort because the study is not powered to estimate the impact of IWISH on utilization among specific subgroups. For example, a prevalence-based cohort would include only individuals previously diagnosed with diabetes if the measure were to include endocrinologists as a type of specialist, or heart failure patients if it were to include cardiologists.

⁹ Hazard models are often used to estimate hazard ratios. In our case, the hazard ratio will measure the rate at which residents still resided at the treatment properties at any point in time compared to the control group. A hazard ratio of one means that there is no difference in the rates that treatment and control group residents remained residing at their properties over the course of the demonstration. A hazard ratio that is greater than one or less than one means that the rate was better in one of the two groups.

to end their residency during the three-year demonstration period than residents of non-IWISH properties” or “Residents of IWISH properties and non-IWISH properties were equally likely to end their residency during the three-year demonstration period.”¹⁰

We will model one secondary outcome measure for research question 3: the likelihood of residency ending due to death. Comparison of the impact of IWISH on the confirmatory outcome and mortality will help assess whether the impact of IWISH on housing exits are driven by the model’s impact on mortality or its impact on all other reasons for housing exits (e.g., evictions, or moves to different types of housing).

Research Question 4 focuses specifically on transitions from living in the community to residing in an institution offering long-term care. To answer this question, we will test whether residents of the treatment properties were more or less likely to end their residency because they were admitted to a long-term care facility (e.g., skilled nursing facility, intermediate care facility, or psychiatric residential treatment facility). The confirmatory outcome is the likelihood that *residency ended due to transition to a long-term care facility* of any type (see Exhibit 2-5). As with the confirmatory measure for research question 3, the impact of IWISH will be expressed in terms of likelihood or odds. For example, “Residents of IWISH properties were half (twice) as likely to transfer to a long-term care facility during the three-year demonstration period than residents of non-IWISH properties” or “Residents of IWISH properties and non-IWISH properties were equally likely to transfer to a long-term care facility during the three-year demonstration period.” The two secondary outcome measures for research question 4 examine the number of days spent in a long-term facility, distinguishing between nursing facilities and other types of facilities. The time spent in a long-term care facility during the study period would be a key driver of the total costs of care for residents who transitioned to long-term care.

Exhibit 2-5. Outcome Measures for Research Question 4 (● = confirmatory outcome)

Outcome Measure	Primary Data Source
● Residency ended due to transition to a long-term care facility	HUD data, Minimum Data Set 3.0 (MDS), State Medicaid claims and encounter data
Days in a nursing facility for long-term care	Minimum Data Set 3.0 (MDS), State Medicaid claims and encounter data
Days in other long-term care facilities	State Medicaid claims and encounter data

In addition to estimating impacts on the confirmatory and secondary outcomes for research questions 3 and 4, we will also examine the reasons that residency may have ended by categorizing them into five categories: transfer to a nursing facility for long-term care; transfer to any other type of long-term care facility; eviction; mortality; and all other reasons (e.g., moved in with family, moved to a different HUD-assisted property, moved to other housing type). We do not have administrative data source that would provide further detail on where residents moved if they did not move to a nursing home or other type of long-term care facility and did not die. We are exploring with the

¹⁰ Estimates from hazard models can also be interpreted to determine the difference in the average length of time that the residents of IWISH properties and control properties would be expected to reside at the property after the demonstration begins, controlling for baseline factors like age and housing tenure.

implementation team whether it will be possible to collect more detailed data on housing exits for at least those residents enrolled in IWISH through the PHL system, assuming that RWDs know where their clients move, and possibly for residents of control properties. Ideally, we would like to break out the other reasons that residents left the property and identify how many moved in with family or friends, how many moved into another HUD-assisted property, and how many moved to another housing type.

Since housing exits, age of the resident, and overall tenure at the property are inextricably linked, our main analyses will focus on housing exits, but we will also use non-experimental, descriptive analyses to examine the relationship between housing exits, age at exit, and tenure and how the relationships might differ between the treatment and control groups. We will also conduct non-experimental, descriptive analyses of changes over time in the rate of evictions from the property, if feasible with the available HUD administrative data.¹¹ These exploratory analyses will support the confirmatory measures, providing context or giving further weight to the estimated impacts.

2.4 Cross-Cutting Measures

We identified five outcome measures that are of interest to HUD but that span multiple research questions. Findings based on these “cross-cutting” measures will support the findings related to more than one confirmatory measure. These measures are shown in Exhibit 2-6, and all will be analyzed as secondary outcome measures.

The first measure is days in the community, defined as the number of days that residents were not admitted to a facility for acute inpatient or long-term care. An increase in the number of days in the community is associated with a decrease in the number of days hospitalized and a decrease in the rate at which residents transition to a long-term care facility. This measure will be derived from a measure currently under development at CMS. We plan to have further discussion with CMS and the expert panel about the specification of this variable before we conduct the impact analysis for the study’s final report. If possible, we would like to use a measure that is consistent with CMS measures, but only if it is specified in a way that makes sense for the IWISH evaluation.

The other four measures are measures of total healthcare costs and three subcategories of healthcare costs: inpatient costs, outpatient costs, and pharmacy costs. A short-term increase in the use of non-acute healthcare services and a longer-term decrease in the use of acute care services will likely be associated with a short-term increase and longer-term decrease in total spending on medical and pharmacy services. Total medical and pharmacy costs will be measured as the combined amount of Medicare fee-for-service spending and Medicaid spending for persons dual enrolled in both Medicare and Medicaid.

Exhibit 2-6. Cross-Cutting Outcome Measures

¹¹ When multifamily property owners process a move-out, they are supposed to complete the form HUD-50059-A and enter a code that best describes the reason for the move-out. Two of the codes refer to owner-initiated move-outs (i.e., evictions). Discussion with HUD suggest these data may not be reliable—not all properties have well-populated data in this field, and evictions may be undercounted in the data as tenants frequently move out before facing formal eviction.

Outcome Measure	Primary Data Source
Days in the community (not in an inpatient or long-term setting)	Medicare Inpatient, Outpatient, MDS, and State Medicaid claims and encounter data
Total medical and pharmacy Costs ^a	All Medicare claims and drug events files, State Medicaid claims, encounter, and drug events data
Total inpatient costs	Medicare Inpatient File
Total outpatient costs	Medicare Outpatient, Carrier, and DME Files
Total pharmacy costs	Medicare Part D Events File, State Medicaid Prescription Drug Events (PDE) data

^a Total medical and pharmacy costs include combined FFS Medicare and Medicaid expenditures for dual-enrolled beneficiaries.

As described further in Chapter 3, the impact study will use non-experimental, descriptive analyses to examine certain components of healthcare utilization and spending that are not included in the lists of outcome measures in the preceding exhibits. For example, to help explain or provide context for certain impact estimates, we might also be interested in describing changes in the distributions of home healthcare visits, outpatient hospital visits, the use of durable medical equipment, or days in skilled nursing facility for post-acute care from baseline through the demonstration period.

2.5 Patient, Property, and Community Characteristics

In order to estimate the impact of IWISH, we need to examine whether residents' levels of healthcare utilization and costs and other measureable resident and community characteristics that could affect the impact of IWISH are well balanced across the treatment and control groups at the start of the IWISH demonstration. In other words, we need to confirm that the groups were sufficiently randomized by the experimental design. Unbalanced characteristics that are identified can be controlled for by including them as covariates in multivariate regressions used to estimate the risk-adjusted impact of IWISH on resident outcomes. Exhibit 2-7 describes the types of resident and community characteristics that could influence residents' health and housing outcomes and potential measures that might be included in the multivariate regressions to control for such factors.

Exhibit 2-7. Measurable Resident, Property, and Community Characteristics That Could Influence Healthcare Utilization and Costs or Housing Exits

Category	Potential Measures
Income Level	<ul style="list-style-type: none"> Set of binary indicators identifying the resident's income status as low-income, very-low income, or extremely low income (a proxy for the resident's socioeconomic disadvantage).
Housing Tenure, years	<ul style="list-style-type: none"> Number of contiguous years that the resident resided at the property (housing exits may be less likely among individuals that only recently moved in).
Family size	<ul style="list-style-type: none"> Number of family members also residing in the residential unit
Property conditions	<ul style="list-style-type: none"> REAC physical inspection score
Demographic characteristics	<ul style="list-style-type: none"> Continuous or binary indicators for age, gender, race/ethnicity, disabled (based on original reason for Medicare or Medicaid, TRACS data, or diagnosis codes identified in baseline claims data)
Insurance status	<ul style="list-style-type: none"> Set of binary indicators identifying the resident's source of insurance: Medicare only, Medicaid only, or Dual-enrolled in Medicare and Medicaid

Category	Potential Measures
Prior use of healthcare	<ul style="list-style-type: none"> Counts or binary indicators for any use of certain healthcare services in the prior 12 months, e.g., acute inpatient stays, ED visits, SNF stays, and home healthcare visits.
Prior healthcare spending	<ul style="list-style-type: none"> Total Medicare FFS payments in prior 12 months Total Medicaid FFS and Medicaid Managed Care payments in prior 12 months
Health status	<ul style="list-style-type: none"> Binary indicator for depression Chronic conditions (measured either as a count of 27 total chronic conditions predefined by CMS, or measured using the Charlson Comorbidity Index¹²) End-stage Renal Disease (based on original reason for Medicare)
State	<ul style="list-style-type: none"> A set of binary indicators of the seven study states, to adjust for geographic variation in Medicaid policies and Medicare and Medicaid payments.
Local healthcare market characteristics	<ul style="list-style-type: none"> Variables to characterize supply and demand in the local market for healthcare services (e.g., county, Hospital Referral Region, Core Based Statistical Area). For example, the number of primary care physicians or the number of hospital beds per 10,000 residents age 65 or older (<i>supply</i>), and the number of inpatient covered stays or SNF covered stays per 1000 Medicare beneficiaries (<i>demand</i>).

¹² The Charlson Comorbidity Index is a prognostic indicator based on 23 chronic diseases, and one medication. It is typically computed using aggregated diagnoses data found in one year of claims. It has been validated in a number of different settings and populations, particularly the Medicare and Medicaid populations. Studies have also found that the Charlson Comorbidity Index is a predictor of future utilization and costs by Medicaid managed care beneficiaries and private employer-provided insurance beneficiaries (Charlson et al., 2014a; Charlson et al., 2014b).

3. Data Sources

The impact study will use individual-level administrative data from HUD, CMS, and state Medicaid agencies for the seven states with properties in the demonstration. We will also use various public use data to risk-adjust study outcomes for community characteristics that might affect the impact estimates. Exhibit 3-1 presents the data sources, described in more detail below.

Exhibit 3-1. Impact Evaluation Data Sources and Types

Source	Type
HUD	<ul style="list-style-type: none">• TRACS and iREMS• REAC physical inspection data
CMS Medicare	<ul style="list-style-type: none">• Medicare Beneficiary Summary Files (Enrollment, Chronic conditions, Cost & Use)• FFS claims (Parts A, B, D)• Medicare Advantage (Part C) Encounter data for Managed Care Beneficiaries
State Medicaid	<ul style="list-style-type: none">• Enrollment• Fee-for-service claims• Managed Care Encounters• Prescription drug events
Public Use	<ul style="list-style-type: none">• CMS Geographic Variation Public Use File• Area Health Resource Files• AARP Livability Index• Distressed Communities Index• Others as warranted

3.1 HUD Data

TRACS and iREMS

HUD’s Tenant Rental Assistance Certification System (TRACS) is the main system that HUD uses to collect and store data on the individuals and families living in HUD multifamily housing. TRACS contains data for all individuals who reside in units with assistance from HUD’s multifamily programs. These data include age, race/ethnicity, household size, household income and assets, and household-level transactions such as move-ins, annual reexaminations, and exits.

HUD’s Office of Policy Development and Research (PD&R) maintains quarterly extracts of individual-level and household-level TRACS data that include the latest available transaction for every person served in HUD’s multifamily programs over the previous 18 months. HUD’s Integrated Real Estate Management System (iREMS) is the official source of data on Multifamily Housing’s portfolio of insured and assisted properties. We will draw on iREMS as needed to supplement the TRACS data. We will request the following data points for all HUD-assisted individuals living in the 124 properties:

- Date of birth
- Gender
- Race/ethnicity
- Disability status
- Household size
- Number of months receiving HUD assistance
- Annual household income
- Sources of income
- Rent burden
- Date of project move-in

-
- Number of dependents in the household
 - Date of project exit (if applicable)

We expect to request extracts of TRACS data at baseline (September 2017) and at the following points in time during the demonstration period: September 2018, December 2018, March 2019, September 2019, September 2020, and December 2020.

For the impact study, we will use TRACS data for the following:

- To describe the characteristics of residents that might be associated with healthcare utilization, cost, or housing turnover, such as demographics, income level, housing tenure, and size of the household.
- To test whether the resident characteristics were sufficiently balanced across properties under the experimental design. Multivariate regression models will control for imbalanced resident characteristics.
- To identify move-outs from the property by residents. TRACS has data fields on terminations and move-outs (including the reason for move-outs) that provide key data for the outcome measures related to housing turnover. If we find that these fields are not well-populated, another approach is to use the “record truncation” method, whereby we assume exit at the time of the last household record (Locke et al., 2011). We will also supplement the TRACS data with Medicare and Medicaid enrollment and claims data to determine whether residency was ended due to death, eviction, transfers to an institution, or some other reason, as well as to describe the resident’s age and tenure at the property when residency ended.
- To produce a finder file with personal identifying information for all residents of treatment, active control, and passive control properties to match to Medicare and Medicaid data.

REAC Physical Inspection Reports

HUD’s Real Estate Assessment Center (REAC) conducts physical property inspections of each multifamily assisted housing property every one to three years.¹³ We will collect summary physical inspection scores from HUD’s REAC for each of the demonstration properties. The inspection score could be used to explore how outcomes are related to variation in the physical condition of the properties where treatment and control group members live.

3.2 CMS Medicare Data

The Abt team will work with HUD to request Medicare data for all residents of treatment, active control, and passive control properties from the Research Data Assistance Center (ResDAC), CMS’s data contractor. We will request claims level data and summary data.

Claims-level Data

Claims-level data include Research Identifiable Files (RIF), which contain FFS claims for institutional (Part A) and non-institutional (Part B) providers. Claims include beneficiary identifiers, providers of service identifiers, dates of service, diagnosis codes, procedure codes, and

¹³ Properties with lower inspection scores have more frequent inspections.

reimbursement amount. RIFs are organized by type of claim and include records on Inpatient, Outpatient, Physician/Supplier, Skilled Nursing Facility, Home Health, Hospice, and Durable Medical Equipment cost and use.

RIFs also contain drug event records for all Medicare Part D beneficiaries. This file contains drug identifiers, quantities dispensed, dispense dates, and reimbursement amounts for all drugs dispensed outside of an inpatient or outpatient setting and covered by a beneficiary's prescription drug plan.

The claims and drug event records will serve as the primary source of data for constructing measures of healthcare service use and spending. The overwhelming majority of low-income elderly residents are eligible for Medicare only or are dually eligible for both Medicare and Medicaid coverage. Even for dual-eligible beneficiaries, Medicare is the primary payer for healthcare services. Medicaid pays for specific services not covered by Medicare and sometimes covers the cost of premiums, deductibles, co-pays or co-insurance (benefits vary across states).

We will also obtain the Long Term Care Minimum Data Set 3.0 (MDS) as part of the RIFs we collect. The MDS is a standardized tool for screening and assessment of health status for all admitted residents of long-term care facilities or skilled nursing facilities certified to participate in Medicare or Medicaid, regardless of payer. It contains a comprehensive set of items that measure physical, clinical, cognitive, and psychological status. These data will be primarily used to identify residents' admissions to a nursing home for long-term care regardless of whether they are covered by Medicaid when they were admitted.¹⁴

Summary Data

Summary data are contained in the Master Beneficiary Summary files (MBSF), which include the following "segments".

- The **Base A/B/D** segment contains data on Medicare beneficiaries' past and current enrollment status in Parts A, B, and D, dual eligibility status, reasons for entitlement (e.g., disability, end-stage renal disease, old age), and receipt of low-income subsidies. It also includes demographic characteristics of beneficiaries (e.g., date of birth, gender, race) and date of death.
- The **Chronic condition** segment provides beneficiary-level flags for 27 common and chronic conditions in the Medicare and Medicaid populations, including acute myocardial infarction, atrial fibrillation, chronic kidney disease, chronic obstructive pulmonary disorder, depression,

¹⁴ Medicare does not cover long-term care services. Medicare does cover skilled nursing facility (SNF) stays for rehabilitation after a hospital stay of three or more days. The SNF stay is fully covered for up to 21 days and partly covered for days 22 through 100. Medicare covered stays in a SNF for post-acute care can be identified using the MDS or Skilled Nursing Facility claims, therefore we only plan to use the MDS to identify all transfers to a nursing facility for long-term care. The MDS does not indicate whether admission to the facility is for post-acute care or long-term care, so we will assume that any MDS admission assessment not preceded by a Medicare claim or Medicaid claim or encounter record for an inpatient hospital stay is for long-term care and not short-term post-acute care.

diabetes, heart failure, hip/pelvic fracture, rheumatoid/osteoarthritis, stroke/TIA, and selected types of cancer.

- The **Other Chronic or Potentially Disabling Conditions** segment contains beneficiary-level flags for nine mental health and tobacco use conditions, 15 developmental disorder and disability-related conditions, and nine other chronic physical and behavioral health conditions which were developed by CMS specifically to enhance research of the Medicare-Medicaid dually enrolled population.
- The **Cost and Use** segment compiles institutional and non-institutional claims records to calculate summary measures of utilization and expenditures, aggregated by calendar year, for each beneficiary.

Information on enrollment status in the Base A/B/D segment will be important for identifying whether residents should be included in analyses that focus on Medicare utilization and spending or excluded because of gaps in Medicare FFS enrollment and therefore gaps in data on utilization and cost during the baseline and demonstration periods. The Base A/B/D segment will also be the primary source of information on dates of death in analyses of mortality. All other information provided in the MBSF will be primarily used for descriptive analyses of baseline characteristics or to construct baseline covariates included in multivariate analyses.

With the staggered availability of certain types of CMS data, we will make four requests to ResDAC for Medicare fee-for-service claims and summary data over the course of the evaluation (an initial request and three modifications of the request). The first two sets of Medicare data will provide two years of baseline data on residents' pre-demonstration utilization and costs of healthcare services (October 2015 through September 2017). We will use these data to conduct descriptive analyses of baseline characteristics of residents of IWISH and control properties. The second two sets of data will provide data for the entire demonstration period (October 2017 through September 2020).

Medicare Advantage (Part C) Encounter data for Managed Care Beneficiaries

Medicare encounter data include data submitted to CMS by Medicare Advantage (i.e., managed care) organizations that provide services to beneficiaries under the Medicare Part C benefit. Encounter data are RIFs that contain the same information about the service, treating provider, and other patient-level detail found in FFS claims, but do not include information on reimbursement to providers. RIFs are organized by type of claim and include records on Inpatient, Outpatient, Physician/Supplier, Skilled Nursing Facility, Home Health, and Durable Medical Equipment use.

Medicare encounter data was not available through ResDAC at the start of this evaluation and we had not planned to acquire it from other sources. However, Medicare encounter data is now available to researchers that request the data through ResDAC. Only preliminary files for services rendered in FY2015 have been made available as of December 31, 2018, but we anticipate that several years of data on managed care encounters will be available by summer 2021, when we will begin the analysis for the study's final report. We will request all available Medicare encounter data available from October 2015 through September 2020. If the data are not available for the entire study period, we

will determine in consultation with HUD and the expert panel how best to integrate the incomplete data into the analyses described below.¹⁵

3.3 State Medicaid Data

We are pursuing individual Data Use Agreements (DUAs) with each of the seven study states to access an excerpt of state Medicaid data. Our data requests will include enrollment, FFS claims, prescription drug events, and managed care encounter records for all residents of IWISH and control properties. These data will include the same fundamental beneficiary-level information as the RIFs discussed above. However, the total content and structure of these files will vary across states.

State Medicaid data will include encounter data for any beneficiaries enrolled in managed care. Medicaid agencies pay a monthly per member per month fee to managed care organizations (MCOs) to provide healthcare services required by enrollees. The organizations submit “encounter” claims with the same information about the service, treating provider, and other patient-level detail found in FFS claims but, for most (but not all) states, the encounter data will not include information on reimbursement to providers. For states that do not include reimbursement amounts on encounter records, we will impute the amount in one of three ways:

1. Medicare provides a public fee schedule of FFS payments for specific services and procedures. Some states use this fee schedule for their own reimbursement decisions while others have a modified or specific fee schedule. With fee schedules available, the amount that would have been paid for the same services under FFS can be imputed using procedural codes (i.e., the Healthcare Common Procedure Coding System, or HCPCS) listed on the encounter records.
2. FFS claims often provide information on both the amount charged by a provider and the amount paid. Encounter data often includes the amount charged but not the amount paid. The amount paid on encounter records can be imputed by using comparable FFS claims to compute the ratio of average amount charged to average amount paid for specific services and procedures, then calculate the amount that would have been paid for the same services under a FFS plan as a proportion of the amount charged to the MCO.
3. If the amount charged is not provided on the encounter claims, then the amount paid can be imputed as the average amount paid by comparable FFS claims for those services and procedures.

¹⁵ If we cannot obtain Medicare encounter data for any portion of the baseline and study periods, we will exclude Medicare beneficiaries residing at treatment and control group properties not continuously enrolled in Medicare Parts A and B (i.e., a Medicare FFS plan) from January 1, 2016 to December 31, 2017. If the cluster-randomized design was effective at randomizing Medicare managed care beneficiaries across the treatment and control properties then excluding Medicare managed care beneficiaries from the study sample will not violate the experimental conditions, and the impact estimates will be unbiased and generalizable to the larger population of low-income seniors receiving HUD assistance. If we find that Medicare managed care beneficiaries are not randomized across the treatment and control properties, we will determine in consultation with HUD and the expert panel how to address concerns about potential bias and the generalizability of the results.

State Medicaid data will serve as the primary source of data on healthcare service use and spending for residents eligible for Medicaid only (i.e., < 65 years of age and not qualified for disability or end-stage renal disease insurance). It will supplement Medicare claims data for dual-eligible beneficiaries since Medicare would be the primary payer for most healthcare services.

We will collect the Medicaid data from each of the seven states involved in the study only once, in April 2021. We expect that State Medicaid agencies will also require six months of runout to obtain complete adjudicated claims and encounter data. However, the required runout may vary across states, since more time might be needed to process the encounter data collected and submitted to states by managed care organizations. The expectation is that the April 2021 data request will provide Medicaid data across the seven states for the October 2015 through September 2020 timeframe, the same as the Medicare data.

3.4 Public Use Data

Various publicly available databases will be used to characterize residents' communities and local markets for healthcare services. These data will be used to risk-adjust resident outcome measures for contextual factors that might affect impact estimates. We expect to use the following public use data sets:

- ***CMS Geographic Variation Public Use File***: county level data on the local demand for Medicare covered healthcare services.¹⁶
- ***HRSA Area Health Resource File (AHRF)***: county level data on the local supply of healthcare services.¹⁷
- ***Distressed Communities Index***: ZIP code level index combining seven metrics to characterize the relative economic well-being of a community.¹⁸
- ***AARP Livability Index***: seven composite measures collected at the county, ZIP code, and census-tract levels and used to rank communities based on transportation, housing, neighborhood, environment, health, engagement, and opportunity.¹⁹
- ***American Community Survey (ACS)***: data collected by the U.S. Census Bureau on community characteristics (e.g., race, ethnicity, poverty rate, homeownership rate) at the census tract level (and other levels of geography).²⁰

¹⁶ See https://www.cms.gov/research-statistics-data-and-systems/statistics-trends-and-reports/medicare-geographic-variation/gv_puf.html.

¹⁷ See <https://data.hrsa.gov/topics/health-workforce/ahrf>.

¹⁸ These metrics are: adults without a high school diploma; poverty rate; prime-age adults not in work; housing vacancy rate; median income ratio; change in employment; and change in establishments. See <https://eig.org/dci>.

¹⁹ See <https://livabilityindex.aarp.org/livability-defined>.

²⁰ See <https://www.census.gov/programs-surveys/acs>.

We will assemble the latest data from each source at the start and end of the study.

3.5 Linking Data Sets

We will use TRACS data to create a finder file with personal identifying information (PII) for all residents of treatment, active control, and passive control properties. A unique, permanent Abt Study ID will be assigned to every resident in the study population and included in this finder file. The finder file will be provided to ResDAC and State Medicaid agencies to match the residents to Medicare and Medicaid beneficiaries' enrollment, claims and encounter data on the basis of the PII.

We will request that ResDAC and the State Medicaid agencies match the finder file to their data using SSN only. That provides the least restrictive match. If there is a match based on SSN but another identifier such as name, date of birth, or gender does not match we will flag the record for further analysis. We will instruct ResDAC and State Medicaid agencies to destroy the PII once the residents have been matched to Medicare and Medicaid beneficiaries, leaving only the Abt study ID to be used to link all three data sets at the individual level and over time.

After receiving the data from ResDAC and the State Medicaid agencies, we will run further tests to analyze the partial matches and to determine how the match rates change when we apply restrictions such as date of birth and gender. If there is a substantial difference in the match rate between the SSN-only approach and a more restrictive approach, we will analyze the data to see if we can identify any systematic discrepancies and work with HUD to determine which match criteria to use for the impact analysis dataset.

These processes will be done in accordance with regulations through CMS, HUD, and state agencies that administer state Medicaid programs for the states selected for study. The study team will also follow the detailed provisions and processes in the study's data security plan to protect the PII and protected health information collected as part of the evaluation.

4. Analysis Approach

The main approach to estimating the impact of IWISH on resident outcomes is an experimental approach that uses the cluster-randomized design of the IWISH demonstration. Residents of the active and passive control group properties are pooled into one comparison group, and the impact of IWISH is estimated generally as the difference between the average outcomes among residents of IWISH properties and the average outcomes among residents of control properties.

We will conduct additional analyses of key outcomes in order to examine: non-linear trends in the cumulative effect of IWISH on healthcare utilization and spending during the demonstration; potential heterogeneity of the treatment effect across important subgroups of individuals; and the extent to which sample attrition due to deaths might bias the estimated impact of IWISH on utilization and spending.

The main approach to estimating the impact of IWISH estimates the impact of *offering* housing-based supportive services to low-income elderly residents residing at HUD-assisted properties that were selected as treatment properties to implement the IWISH model. The impact of offering services often is referred to as the impact of the “intent-to-treat” (ITT). In contrast, the impact of the “treatment on the treated” (TOT) tells us the impact of the IWISH model on residents of the property who choose to enroll in the program. TOT estimates will help to assess whether the impact of IWISH on average outcomes are really driven by the outcomes of residents who participate in the IWISH model.

This section describes the experimental and quasi-experimental analyses. It also describes how we will use Instrumental Variable (IV) methods to estimate the effects of IWISH on those who enrolled in IWISH (the TOT estimates), adjusting for potential non-random differences in the characteristics of residents who did and did not enroll.

4.1 Experimental Sample

Residents of SSD Properties at Launch of Demonstration

The key advantage to cluster-randomized designs is that, like individual randomized controlled trials, they provide an unbiased estimate of a program’s impact on individual-level outcomes when randomization of individuals cannot be achieved. Unbiased impact estimates are only assured, however, if analyses are based on *all eligible residents of the properties at the time they were randomized* or a random sample of them. To ensure that the sample of residents is not contaminated by non-random entry of new residents across properties, the experimental method restricts the study sample to individuals residing at treatment and control properties at the launch of the demonstration (October 1, 2017).

If resident characteristics were sufficiently balanced across the treatment and control groups using the cluster randomized design, then we expect the measured baseline characteristics of IWISH residents and control residents to have similar distributions. Before proceeding with impact estimates, we will use a variety of appropriate statistical tests to assess whether the distributions of baseline covariates are indeed similar between the treatment and control groups. Such tests include comparing means and prevalence of characteristics using standardized differences, ratios comparing the variance of continuous covariates, comparisons of higher order moments and interactions, and graphical methods

such as quantile-quantile plots, side-by-side boxplots, and non-parametric density plots (Austin, 2009). If the distributions of some baseline characteristics are not similar, then we can still achieve an unbiased impact estimate by using multivariate regressions to control for potentially confounding factors, as described below.

Individuals who Move in to SSD Properties after October 1, 2017

The distributions of baseline resident characteristics could still be similar between the IWISH and control groups when including residents who moved in during the demonstration. To test this, we will define the baseline date for new residents as the date on which they moved into their unit, and we will use the same statistical tests to compare the treatment and control groups described above. If characteristics prove to be similar, we will relax the sample restriction and expand the experimental sample to include all individuals who were living in a treatment or control group property between October 1, 2017 and September 30, 2018. At this stage, we are not contemplating expanding the sample to include individuals who moved in after September 30, 2018 because of cost considerations.²¹

4.2 Experimental and Quasi-Experimental Analyses

IWISH was intended for residents of HUD-assisted properties aged 62 or older, or individuals in households headed by a person 62 or older. Although outreach activities by IWISH staff will be targeted to these residents, the program will be open to any individual willing to enroll. Moreover, residents who do not enroll will likely still be able to receive limited services upon request. Therefore, the impact study will include two types of analyses to examine the overall impact of the IWISH model: intent-to-treat (ITT), which estimates the impact of *offering* housing-based supportive services under the IWISH model, and treatment-on-the-treated (TOT), which estimates the impact of participating in the IWISH model. These analyses are described below.

Intent-to-Treat (ITT) Analysis

ITT analyses compare outcomes for *all residents* of the treatment and control group properties. These will be the primary set of analyses for the impact study. Specifically, the comparison is between the following two populations:

²¹ We collect the CMS Medicare data in four waves. The first two waves will be received by April 2019 (pending any delays in obtaining CMS approval of our data request) and will contain two years of baseline data and one quarter of data from the study period, for all residents of IWISH and control properties identified using the September 2018 TRACS extract. Data covering the rest of the study period will be collected for these residents by April 2021. It would be expensive to request the necessary two years of baseline data for individuals who moved in after September 2018, since that data would overlap with years of data already collected and paid for in the first two waves. Some portion of the baseline data for these new residents would have to be paid for once again in the third and fourth waves, at full cost.

- Individuals who were living in a treatment group property as of October 1, 2017,²² regardless of whether they meet the targeted criteria for outreach activities by IWISH staff.
- Individuals living in an active or passive control group property as of October 1, 2017, regardless of whether they meet the targeted criteria for outreach activities by IWISH staff.

Treatment-on-the-Treated (TOT) Analyses

TOT analyses estimate the effects of participating in the IWISH model (we define participation as enrollment in the program) using quasi-experimental, Instrumental Variable (IV) methods. The IV estimator uses the variation in participation that is induced by the cluster-level random assignment of properties to the treatment and control groups to estimate the impacts of participation on outcomes for those induced to participate (Angrist & Imbens, 1991, 1995). It is often expressed as a two-stage model, the first stage predicting participation based on the characteristics of all residents of the treatment and control properties, and the second stage using the predicted probability of participation for all residents as the variable used to show the IWISH model’s effectiveness (instead of a binary indicator for residing at a treatment property). The basic steps for this two-stage approach to produce the IV estimates are described below. However, the actual IV model is a one-step estimator. In the situation where we can be certain of the sign of the first-stage relationship between assignment to treatment and participation, an unbiased and efficient estimator is available for a linear model (Andrews and Armstrong 2017). For nonlinear models, an asymptotically efficient estimator uses the Generalized Method of Moments (Nichols 2007).

Multivariate Regressions

The experimental approach will estimate the effect of IWISH on healthcare utilization and costs using multivariate generalized linear models (GLM) of the type specified in Equation 1 below.

$$E(Y_{ij}|W_{ij}, X_i, Z_j) = f(\alpha + \delta W_{ij} + \Gamma_1 X_i + \Gamma_2 Z_j) \quad (1)$$

Y_{ij} represents the outcome of resident i at property j , and W_{ij} is an indicator for residing at an IWISH property. The estimated coefficient $\hat{\delta}$ captures the impact of IWISH on the outcome. Covariate matrices X_i and Z_j include baseline resident and property characteristics, which will reduce the variance of $\hat{\delta}$ and control for potential confounding factors that are imbalanced across the treatment and control group. The function $f(\cdot)$ is the inverse link function appropriate for the distribution of the outcome. For instance, the identity function, $f(x)=x$, might be appropriate for linear regression of a continuous outcome; the inverse logit function, $f(x)=\exp(x)/[1+\exp(x)]$, appropriate for a binary outcome, or an exponential function, $f(x)=\exp(x)$, for a log link, appropriate for count or nonnegative outcomes.

²² Once we have baseline Medicare claims data, we can reevaluate the sample restriction that all individuals in the experimental sample need to reside in the property on or before October 1, 2017. If we find that the distributions of baseline resident characteristics are statistically similar across the treatment and control groups when including residents who moved in during the demonstration, we can relax the sample restriction and include all individuals who were living in a treatment or control group property as of September 30, 2018 in the experimental sample.

With a cluster-randomized design, outcomes are assumed to be correlated among residents of property j but independent across properties. All estimates will use cluster-robust standard error estimators to assure valid inferences in the presence of intra-cluster correlation with an unknown distribution (Kezdi, 2004; Nichols and Schaffer, 2007).

To estimate the impact of IWISH on housing exits, mortality, and transfers to an institutional setting for long-term care we will use a semi-parametric discrete-time model estimated with logistic regression (Jenkins, 1995), which will model outcomes similar to the multivariate model expressed by the generic regression equation above.

Longitudinal Analyses

The main set of analyses for the impact study will focus on the cumulative impact of IWISH on healthcare utilization over the demonstration period. However, we will also conduct experimental analyses that examine the impact of IWISH over time to see whether there may be a multiyear relationship between supportive services provided through IWISH and residents' healthcare utilization and spending. It is possible that the cumulative effects of the IWISH program will not immediately impact residents' utilization of healthcare services or will have a greater impact over time. The equation above can be easily modified to test for a non-linear effect over two or more periods. Let T_t be a vector of indicators equal to one in time period t and zero otherwise. Then the GLM can be modified as in Equation 2 below.

$$E(Y_{ijt}|W_{ij}, T_t, X_i, Z_j) = f(\alpha + \beta_t T_t + \delta_t(W_{ij} \times T_t) + \Gamma_1 X_i + \Gamma_2 Z_j) \quad (2)$$

Here, Y_{ijt} is the outcome of resident i at property j in time period t . The vector of coefficients β_t captures any trend in the outcome over time that is common among all residents. The estimated vector of coefficients δ_t captures the impact of IWISH in each respective period. The cluster-robust standard error estimator still ensures valid inferences in the presence of both intra-cluster correlation and serial correlation of outcomes due to repeated measurements for each resident.

Subgroup Analyses

Equations 1 and 2 can be expanded to explore heterogeneity in IWISH's impact across subgroups of residents or properties by estimating coefficients on interactions of W_i and specific baseline covariates in vectors X_i or Z_j . For example, subgroups of residents who are less able to anticipate, cope with, or recover from acute care incidents, or who require special care from nurses, may be of particular interest to policymakers and stakeholders. These more vulnerable subgroups might include residents who are 75 years or older, or residents with certain physical or mental disabilities or conditions. Other subgroups of interest might include residents of treatment and control properties with one or more service coordinators employed before the demonstration began, or all residents of properties with a relatively high percentage of residents who are 75 or older.

Once the administrative data is collected, we will determine (in consultation with HUD and the Expert Panel) which subgroups are of greatest interest to stakeholders *and* sufficiently prevalent at baseline to facilitate hypothesis testing of heterogeneous treatment effects using experimental and IV analyses (minimum detectable effect sizes are discussed below). Experimental and IV analyses will only be used to test for heterogeneous treatment effects across subgroups of residents that can be identified at baseline (i.e., fixed or exogenously determined groups of residents).

Sample Attrition

Since IWISH is intended to address the unmet needs of low-income elderly residents, and motivate them to take steps to improve their health and well-being then it is reasonable to hypothesize that a disproportionately smaller share of residents at IWISH properties will exit the study because of death relative to residents at control properties. Estimates of IWISH's average impact on cumulative healthcare utilization and spending could be biased if the residents at IWISH properties are less likely than residents at control properties to die before the end of the demonstration, a common concern when estimating treatment effects in a population with a non-negligible mortality rate, such as an elderly population.^{23, 24}

We will conduct an additional set of analyses to examine the extent to which sample attrition due to mortality (i.e., truncation by death) may have biased the estimated impact of IWISH on healthcare utilization and costs. Multiple statistical approaches have been proposed to estimate treatment effects when outcomes are truncated by death (Kurland et al., 2009). Members of the expert panel recommend that our team conduct a rigorous review of the statistical literature on the topic closer to the time of the analysis in order to identify the best approach for this particular study.

Currently, the most appealing method is a straightforward approach proposed by Hayden et al (2005), which makes use of baseline characteristics of the residents to weight observations by the likelihood the person would have survived had they been randomized to the other arm of the experiment. Specifically, this approach uses logistic regression to estimate the probability of survival among residents in the treatment group and the control group, respectively. The outcomes of residents in the treatment group are weighted by the fitted probabilities of survival calculated using the model that was estimated for the control group. The outcomes of residents in the control group are weighted by the fitted probabilities of survival calculated using the model that was estimated for the treatment group. The sandwich estimator appropriately adjusts standard errors and confidence intervals.

Pre-Post Descriptive Analyses

The experimental approach measures the impact of IWISH on resident outcomes using regression analyses to compare the mean outcomes between the residents of IWISH and control properties, adjusting for potential confounding due to baseline resident or property characteristics. We will also use descriptive pre-post analyses to examine changes in the distributions of certain outcome

²³ We are not likely to have sufficient statistical power to rigorously test the hypothesis that IWISH impacts mortality, given that we expect a relatively low frequency of the event. Regardless, we cannot assume that mortality will occur randomly across the treatment and control groups given the goals of IWISH, and therefore any non-random attrition from the sample may potentially bias estimates of the treatment effect.

²⁴ A change in Medicare coverage is another potential source of attrition from the sample, as some residents might switch from Medicare FFS coverage to managed care plans offered under Medicare Part C. If the study team is unable to obtain Medicare encounter data for the full study period, these residents will be lost to follow up since we will not have complete information on their utilization of healthcare services (i.e., "administrative censoring"). Hazard models inherently address censoring of any kind, but generalized linear models do not. The technical expert panelists discounted concerns over the possibility of administrative censoring, since it is likely to be infrequent in this small sample of low-income elderly individuals and likely to occur randomly. They raised more concern over the more common issue of "truncation by death."

measures. For instance, using two years of baseline data, we might expect a disproportionate share of total healthcare spending attributed to some portion of high-acuity residents at IWISH and control properties. IWISH may result in larger increases in the use of some health care services among low-acuity residents, with newly recognized chronic conditions, and larger reductions in the use of other types of health care services among high-acuity residents. We will be interested in how the distribution of total healthcare spending among IWISH residents might change from before the demonstration to the end of the demonstration, relative to changes measured among the residents of control properties. We will use univariate analyses to describe the distributions of the outcome measures using means, standard deviations, inter-quartile ranges, deciles, or other appropriate summary statistics. Differences in distributions between the baseline and demonstration period or treatment and control groups will be tested using statistical tests most appropriate for the underlying distributions of the measures (t-tests, chi-squared tests, Wilcoxon Rank Sum test, Analysis of Variance, etc.).

Instrumental Variable (IV) Approach to TOT Analysis

The quasi-experimental, IV approach will be used to estimate the effect of IWISH on healthcare utilization and spending by residents who enrolled in the program. The IV estimator is produced by using multivariate regressions estimated in two-stages in order to control for baseline characteristics of the residents that could influence both their decision to enroll in IWISH and their outcome (Angrist & Imbens, 1995). Luca and Cole (2017) provide an excellent summary of a mathematical derivation of the IV estimator and the description below borrows heavily from their description of the two-stage approach to estimation.

Using the notation above, W_{ij} is an indicator for whether property j was randomly assigned to the treatment group. Let D_i be an indicator for whether resident i enrolled in IWISH. In an IV framework, the IV is W_{ij} which affects enrollment, D_i , which in turn affects the resident's outcome, Y_{ij} . Let covariate matrices X_i and Z_j include baseline resident and property characteristics that could influence enrollment and outcomes. The structural equation of interest is:

$$E(Y_{ij}|D_i, X_i, Z_j) = f(\alpha_3 + \lambda D_i + \Gamma_1 X_i + \Gamma_2 Z_j) \quad (3)$$

Stage 1: Logistic regression is used to model enrollment in IWISH as a function of the instrument (i.e., assignment to the treatment group) and other exogenous covariates. The regression can be written as:

$$E(D_i|W_{ij}, X_i, Z_j) = f(\alpha_4 + \gamma W_{ij} + \Pi_1 X_i + \Pi_2 Z_j) \quad (4)$$

The estimated coefficient $\hat{\gamma}$ is referred to as the “first-stage effect” of the instrument, and will approximately measure the proportion of the sample that enrolled in IWISH. Note that the covariate matrices X_i and Z_j must be exactly the same as the covariate matrices in Equation 3.

Stage 2: Fitted values from the first-stage (the predicted probability that resident i enrolled in IWISH) are plugged directly into Equation 3 in place of D_i , the indicator for whether the resident enrolled in IWISH. The estimated coefficient $\hat{\lambda}$ is the treatment effect of interest.

To ensure the correct standard errors are computed, all two-stage estimation procedures will be conducted in the Stata version 15 statistical package, using packaged routines such as *ivreg2*, *ivpoisson*, *ivprobit*, or *ivtobit*.

Minimum Detectable Effect Sizes

The power properties of analyses with a known sample size are summarized by the minimum detectable effect size (MDES). Sample clustering, captured by the intra-cluster correlation (ICC), and unequal-sized clusters both lower the effective sample size in cluster-randomized experiments, thereby increasing the MDES.²⁵ Exhibit 4-1 shows the smallest detectable differences²⁶ between residents in the treatment and control groups for the four confirmatory outcomes as a function of ICC. According to HUD TRACS data from September 2017 (extracted September 2018), there were 4,274 residents across 40 treatment properties and 9,934 residents pooled across the 84 active and passive control properties. To calculate the MDES for the ITT analyses, we assumed that all residents at the treatment and control properties were covered under a Medicare FFS plan or by Medicaid since we do not currently have data on Medicare and Medicaid enrollment. By rule of thumb, we inflate the ITT MDES by the “first-stage effect” to approximate the MDES for the IV TOT analyses.²⁷ We assume that 80 percent of the residents will have chosen to enroll in the IWISH program by September 30, 2018.

Exhibit 4-1. Minimum Detectable Differences in Key Outcome Measures Varies Strongly with Clustering¹

Days of unplanned hospitalization per 1000 resident months					
Control group average [SD] = 145.3 [388.1]					
<i>ICC</i>	<i>0</i>	<i>0.01</i>	<i>0.02</i>	<i>0.05</i>	<i>0.10</i>
ITT: Detectable Difference	19.9 (14%)	29.8 (21%)	36.5 (25%)	51.4 (35%)	69.3 (48%)
TOT: Detectable Difference	24.9 (17%)	37.3 (26%)	45.6 (31%)	64.3 (44%)	86.6 (60%)
Number of primary care (E&M) visits per 1000 resident months					
Control group average [SD] = 675.1 [760.5]					
<i>ICC</i>	<i>0</i>	<i>0.01</i>	<i>0.02</i>	<i>0.05</i>	<i>0.10</i>
ITT: Detectable Difference	39.0 (6%)	58.3 (9%)	71.6 (11%)	100.8 (15%)	135.9 (20%)
TOT: Detectable Difference	48.8 (7%)	72.9 (11%)	89.5 (13%)	126.0 (19%)	169.9 (25%)

²⁵ Sample attrition has small consequences for MDES, relative to clustering. In any case, we will measure utilization outcomes on a “per resident per month” basis and therefore we will not drop residents who move out of the property during the demonstration from the samples used to estimate the cumulative impact of IWISH on healthcare utilization and costs.

²⁶ Based on the minimum effect sizes detectable 80 percent of the time when rejecting the null hypothesis at the 95 percent significance level ($p < 0.05$).

²⁷ There is no agreed upon statistical approach to estimate the MDES of an IV estimator. The rule of thumb ignores that the IV estimator has a nonfinite mean and variance in exactly identified models, and will slightly underestimate the true MDES. Assuming 80% of residents enrolled and using the rule of thumb, the first-stage effect will be approximately 0.8 and the MDES for the IV TOT analyses is calculated as the ITT MDES divided by 0.8.

Probability of housing exit during the demonstration					
Control group probability = 0.14					
<i>ICC</i>	<i>0</i>	<i>0.01</i>	<i>0.02</i>	<i>0.05</i>	<i>0.10</i>
ITT: Detectable Difference	0.02	0.03	0.03	0.05	0.07
TOT: Detectable Difference	0.03	0.04	0.04	0.06	0.09
Probability of transfer to a nursing facility for long-term care					
Control group probability = 0.07					
<i>ICC</i>	<i>0</i>	<i>0.01</i>	<i>0.02</i>	<i>0.05</i>	<i>0.10</i>
ITT: Detectable Difference	0.01	0.02	0.03	0.04	0.05
TOT: Detectable Difference	0.01	0.03	0.04	0.05	0.06

Abbreviations: SD = Standard Deviation; ICC = Intra-cluster correlation; ITT = Intent-to-treat analysis; TOT = Treatment-on-the-treated analysis.

Sources for Control group assumptions: Total days of unplanned hospitalization and total primary care (E&M) events (The Lewin Group, 2014; US Department of Health and Human Services, 2012, 2017); Probability of housing exit during the demonstration (Pictures of Subsidized Households, 2016); the probability of transfer to a nursing facility for long-term care was assumed to account for 50 percent of housing exits during the demonstration.

¹ Detectable effect sizes for comparing two sample means or two proportions were computed in Stata 15 using the “power” command. Estimates assume two-sided hypothesis tests, 80 percent power, a 95 percent significance level, 105 (SD=54.1) observations per cluster, and the standard deviation of all outcomes are equal in the two groups.

Exhibit 4-1 presents the minimum detectable differences for each outcome with ICCs ranging from zero to 0.10 to show that the loss in precision due to clustering can be substantial. Since we do not currently have Medicare or Medicaid claims data for the baseline period we cannot estimate the baseline ICC for our confirmatory outcomes. However, we were able to use data from an individual-level extract of TRACS data from September 2017 to measure residents’ baseline ages and tenure, which are likely correlated to some extent with health care utilization and housing exits. The ICC for age was 0.10 (95 percent confidence interval: 0.07 - 0.13) and the ICC for tenure was 0.13 (95 percent confidence interval: 0.09 – 0.16). Gains in efficiency can be had from conditioning on baseline covariates, but the gains tend to be very small. The estimates in Exhibit 4-1 are for comparisons of two sample means or two proportions and do not account for risk-adjustment. We focus on discussion of standard experimental comparisons (the ITT analyses).

The baseline values for days of unplanned hospitalization per one thousand resident months and the number of primary care (E&M) visits per 1000 resident months are based on estimates from the 2014 study by The Lewin Group, “Picture of Housing and Health: Medicare and Medicaid Use Among Older Adults in HUD-Assisted Housing” (The Lewin Group, 2014) and statistics from the U.S. Department of Health and Human Services (2012, 2017). Based on these assumptions, the smallest difference that a standard experimental comparison will detect is 19.9 days of hospitalization per 1000 resident months (or a 14 percent difference) when assuming zero correlation of outcomes within a property. The minimum detectable difference increases to 69.3 days of hospitalization per 1000 resident months (or a 48 percent difference) when assuming an ICC of 0.10. For the number of primary care (E&M) visits, the minimum detectable difference is 39.0 (6 percent) per 1,000 resident months when the ICC is zero and 135.9 (20 percent) per 1000 resident months when the ICC is 0.10. Assuming an ICC similar to that of residents’ baseline age, 0.10, the minimum detectable differences for both these measures correspond to a MDES of 0.18. MDES less than 0.20 are often considered as small (Cohen, 1988).

The baseline probability that residents will move out of their unit during the demonstration was calculated as the weighted average of annual housing turnover rates for Section 202 housing in the seven study states (Pictures of Subsidized Households, 2016). The literature did not contain a reliable source of information on the rate of housing exits due to transfers to a long-term care facility by low-income elderly residents at HUD-assisted properties. For exposition, we assumed that half of all housing exits by low-income elderly residents are due to transfers to a nursing home facility for long-term care. Based on these assumptions, experimental comparisons can detect minimum differences in the probability of housing exit ranging from two to seven percentage points for ICCs between zero and 0.10, and minimum detectable differences in the probability of transferring to a nursing home ranges from one to five percentage points.

Descriptive Analyses of Contextual Effects

The impact study will also include a series of descriptive analyses to contextualize the estimated impact of IWISH on key outcomes, using data collected in the process study. We expect that the effectiveness of the IWISH model will vary across treatment properties according to their fidelity to the IWISH model and that organizational or environmental factors that explain or contribute to the variation in fidelity across treatment properties may also contribute to variation in resident outcomes. Differences in service coordination and health and wellness programming provided at treatment properties and active control properties during the demonstration could explain some of the differences in resident outcomes across the treatment and active control properties.

Estimates of the multivariate generalized linear models described above will be used to create risk-standardized measures to characterize each property's relative performance with respect to key outcome measures and other pertinent outcome measures on which the IWISH model had a significant impact. The risk-standardized measure of outcome Y for property j is calculated using the following formula.

$$\text{Risk standardized } Y_j = \left(\frac{\text{Observed } Y_j}{\text{Expected } Y_j} \right) \times \text{Observed } Y \text{ in the entire sample}$$

The observed outcome for property j is the average (unadjusted) outcome among all residents of the property that were included in the impact analysis. The expected outcome for property j is calculated as the average predicted outcome across all residents of the property that were included in the impact analysis. The observed outcome for the entire sample is the average (unadjusted) outcome among all individuals included in the impact analysis.

The risk-standardized measure is a normally distributed measure adjusted for differences in baseline resident and property characteristics across the properties, which can be used to rank a property's performance among its peers. The risk-standardized measures can be compared descriptively with a taxonomy of properties characterizing fidelity to the IWISH model, potential determinants of fidelity, different types or levels of service coordination and related programming available, or other property characteristics. Relationships between properties' performance and property characteristics can be statistically tested using the Pearson correlation coefficient or the Spearman rank correlation coefficient, for example. Simple linear regression will be used to examine the association between property performance and one contextual factor while controlling for other important factors, or between property performance and interactions of more than one contextual factor. Although, the

small number of treatment and control properties limits the number of covariates that can be included in a regression model while retaining enough degrees of freedom for hypothesis testing.

5. Impact Study Limitations

The findings from the impact study will provide important insights for Congress and stakeholders about the impact of IWISH on costly healthcare utilization, transitions to nursing homes, and housing stability. The cluster-randomized design provides a strong basis for measuring impacts, but the study will not be without limitations.

Medicare Managed Care

One limitation is that the data presented in the impact study will not reflect the experience of Medicare managed care beneficiaries if we cannot procure Medicare encounter data from ResDAC. Although HUD placed greater weight on areas with higher Medicare FFS penetration when selecting properties for randomization, we expect that some portion of the study sample will have enrolled in Medicare managed care plans over the study period. If the experimental design was effective in randomizing individuals across the treatment and control properties according to insurance status, then excluding all individuals who are enrolled in Medicare managed care plans (and ending follow-up of individuals who become enrolled during the study period) will not violate the experimental conditions. Excluding these individuals will underestimate the nominal levels of utilization and costs when describing the populations, but the estimated impact of IWISH on the outcomes will not be biased.

If Medicare managed care beneficiaries were not sufficiently balanced across treatment and control properties, then systematic differences in healthcare use patterns between FFS and managed care beneficiaries may result in biased impact estimates. In the latter case, we will explore the data further to determine whether a quasi-experimental approach could be used to address any imbalances in historical utilization patterns and resident characteristics between the treatment and control group. Otherwise, we will determine in consultation with HUD and the expert panel how to address concerns about potential bias and the generalizability of the results.

Because the sample size would decrease if Medicare managed care beneficiaries are excluded from the impact study, the minimum detectable effect sizes (MDES) pertaining to outcomes related to health care utilization would increase (i.e., the statistical power of our models will decrease) as the proportion of residents of IWISH or control properties who are enrolled in Medicare managed care increases. However, MDES is more sensitive to the number of Medicare managed care enrollees residing at IWISH properties (i.e., the size of the treatment group) than it is to the number of enrollees in the control group. We will also revisit the MDES calculations pertaining to our confirmatory outcomes as we explore the data further, then determine in consultation with HUD and the expert panel how to address potential concerns about reduced statistical power due to excluding managed care enrollees from the study sample.

Health Outcomes

A second potential study limitation is that Medicare and Medicaid administrative data do not provide a complete picture of health outcomes. The data do not include clinical information (e.g., lab values and patient logs), information on residents' lifestyles (e.g., drinking, smoking, and sexual activity), self-reported health status, or data on physical and cognitive functional status (e.g., assistance with activities of daily living or social functioning). Although higher levels of healthcare utilization can serve as a proxy for poorer health status, the study is still limited in its ability to directly measure the

extent to which the IWISH model impacts residents health status, health behaviors, and overall well-being. PHL data will capture some of these types of information but only for individuals participating in IWISH and only during the demonstration. Therefore, we will not be able to use the experimental design with PHL data, or even pre-post analyses, to estimate the causal effect of IWISH on functioning, health outcomes, or health behaviors.

Access to Healthcare Services

While the IWISH model may decrease healthcare utilization among participants, less use of healthcare services can also be attributed to constrained access to those services. However, the study population is a relatively homogenous socio-economic group and, given the experimental design that randomized properties within the same states, we expect that residents in both the treatment and control group properties will have similar rates of Medicare and Medicaid coverage, and face comparable market environments, with presumably equal chance at access to medical care and hospital services over the course of the demonstration. Thus, any reduction in healthcare utilization and spending attributed to IWISH will likely be due to better care coordination or improved health outcomes and behaviors rather than greater access to care for residents of the control group properties.

References

- Andrews I, Armstrong TB. [Unbiased Instrumental Variables Estimation Under Known First-Stage Sign](#). *Quantitative Economics*. 2017;8 (2) :479-503.
- Angrist, J. & Imbens, G. (1991). Sources of identifying information in evaluation models. <https://core.ac.uk/download/pdf/6919605.pdf>
- Angrist, J.D., & Imbens, G.W. (1995). Two-stage least squares estimation of average causal effects in models with variable treatment intensity. *Journal of the American Statistical Association*, 90(430), 431-442.
- Austin, P.C. (2009). Balance diagnostics for comparing the distribution of baseline covariates between treatment groups in propensity-score matched samples. *Statistics in Medicine*, 28(25), 3083-3107.
- Castle, N., & Resnick, N. (2016). Service-enriched housing: The Staying at Home program. *Journal of Applied Gerontology*, 35(8), 857-877.
- Charlson M., Wells, M.T., Ullman R., King, F., & Shmukler, C. (2014a). The Charlson comorbidity index can be used prospectively to identify patients who will incur high future costs. *PloS One*, 9(12), e112479.
- Charlson M.E., Wells, M.T., Kanna, B., Dunn, V., & Michelen, W. (2014b). Medicaid managed care: how to target efforts to reduce costs. *BMC Health Services Research*, 14(1), 461.
- Cohen, J. (1988). *Statistical Power Analysis for the Behavioral Sciences* (2nd ed.). Hillsdale, NJ: Erlbaum.
- Corporation for Supportive Housing. (2014). *Housing is the Best Medicine: Supportive Housing and the Social Determinants of Health*. New York, NY: Corporation for Supportive Housing. http://www.csh.org/wp-content/uploads/2014/07/SocialDeterminantsofHealth_2014.pdf
- De Brantes F., Rastogi A., & Painter M. (2010). Reducing potentially avoidable complications in patients with chronic diseases: the Prometheus Payment approach. *Health Services Research*, 45(6 Pt. 2), 1854-1871.
- Elsasser Watson, N., Steffen, B.L., Martin, Marge, & Vandembroucke, D.A. (2017). *Worst Case Housing Needs: 2017 Report to Congress*. Washington, DC: U.S. Department of Housing and Urban Development, Policy Development and Research. <https://www.huduser.gov/portal/sites/default/files/pdf/Worst-Case-Housing-Needs.pdf>
- Emsley, R. et al. (2008). Implementing Double-Robust Estimators of Causal Effects. *The Stata Journal*, 8(3), 334-353.
- Erdem E., Prada S.I., & Haffer S.C. (2013). Medicare payments: How much do chronic conditions matter? *Medicare & Medicaid Research Review*, 3(2).
- Ewen, H. H., et al. (2017). A model for aging in place in apartment communities. *Journal of Housing for the Elderly*, 31(1), 1-13.

Farber, N., & Shinkle, D. (2011). *Aging in Place: A State Survey of Livability Policies and Practices*. Washington, DC: AARP Public Policy Institute. <https://assets.aarp.org/rgcenter/ppi/liv-com/aging-in-place-2011-full.pdf>

Government Accountability Office. 2016. *Elderly Housing: HUD Should Do More to Oversee Efforts to Link Residents to Services*. Washington, DC: U.S. Government Accountability Office.

He, Wan & Larsen, L.J. (2014). *Older Americans with a Disability: 2008–2012*. U.S. Census Bureau, American Community Survey Reports, ACS-29. Washington, DC: U.S. Government Printing Office.

Ho DE, Imai K, King G, & Stuart EA. (2007). Matching as nonparametric preprocessing for reducing model dependence in parametric causal inference. *Political Analysis*, 15(3),199–236.

IWISH Implementation Team. (2017). *Integrated Wellness in Supportive Housing (IWISH) Operations Manual*, November 8, 2017.

Jenkins, S.P. (1995). Easy estimation methods for discrete-time duration models. *Oxford Bulletin of Economics and Statistics*, 57(1), 129-136.

Joint Center for Housing Studies of Harvard University. (2016). *Older Households 2015-2035: Projections and Implications for Housing a Growing Population*. Cambridge, MA: Joint Center for Housing Studies of Harvard University. <http://www.jchs.harvard.edu/housing-a-growing-population-older-adults>

Kandilov, A. et al. (2017). *Support and Services at Home (SASH) Evaluation: Evaluation of the First Four Years*. Washington, DC: U.S. Department of Health and Human Services, Assistant Secretary for Planning and Evaluation. <https://aspe.hhs.gov/pdf-report/support-and-services-home-sash-evaluation-evaluation-first-four-years>

Kandilov, A. et al. (2016). *Support and Services at Home (SASH) Evaluation: Second Annual Report*. Washington, DC: U.S. Department of Health and Human Services, Assistant Secretary for Planning and Evaluation. <https://aspe.hhs.gov/pdf-report/support-and-services-home-sash-evaluation-second-annual-report>

Keenan, T. A. (2010). *Home and Community Preferences of the 45+ Population*. Washington, DC: AARP Research & Strategic Analysis. <https://assets.aarp.org/rgcenter/general/home-community-services-10.pdf>

Kezdi, G. (2004). Robust standard error estimation in Fixed-Effects Panel Models. *Hungarian Statistical Review* (9), 96-116.

Kochera, A., Straight, A., & Guterbock, T. (2005). *Beyond 50.05: A Report to the Nation on Livable Communities: Creating Environments for Successful Aging*. Washington, DC: AARP Research Institute. https://assets.aarp.org/rgcenter/il/beyond_50_communities.pdf

Komisar, H.L. & Feder, J. (2011). *Transforming care for Medicare Beneficiaries with Chronic Conditions and Long-term Care Needs: Coordinating Care Across all Services*. Report from Georgetown University. Retrieved from: http://www.thescanfoundation.org/sites/default/files/Georgetown_Trnsfrming_Care.pdf

Kurland, B.F., Johnson, L.L., Egleston, B.L., & Diehr, P.H. (2009). Longitudinal Data with Follow-up Truncated by Death: Match the Analysis Method to Research Aims. *Statistical Science*, 24(2), 211-222.

LeadingAge and The Lewin Group. (2015). *Affordable Senior Housing Plus Services: What's the Value?* <http://www.leadingage.org/sites/default/files/Housing%20Services%20Value.pdf>

Levine, C. A., & Robinson Johns, A. (2008). *Multifamily Property Managers Satisfaction with Service Coordination*. Washington, DC: U.S. Department of Housing and Urban Development, Policy Development and Research.
https://www.huduser.gov/portal/publications/hsgspec/serv_coord.html

The Lewin Group. (2014). *Picture of Housing and Health: Medicare and Medicaid Use Among Older Adults in HUD-Assisted Housing*. Washington, DC: U.S. Department of Health and Human Services, Assistant Secretary for Planning and Evaluation. <https://aspe.hhs.gov/basic-report/picture-housing-and-health-medicare-and-medicaid-use-among-older-adults-hud-assisted-housing>

The Lewin Group and LeadingAge (2011). *Design of a Demonstration of Coordinated Housing, Health and Long-term Care Services and Supports for Low Income Older Adults*. Washington, DC: U.S. Department of Health and Human Services, Assistant Secretary for Planning and Evaluation. <https://aspe.hhs.gov/report/design-demonstration-coordinated-housing-health-and-long-term-care-services-and-supports-low-income-older-adults>

Locke, G., Lam, K., Henry, M., & Brown, S. (2011). *End of Participation in Assisted Housing: What Can We Learn About Aging in Place?* Assisted Housing Research Cadre Report. Washington, DC: U.S. Department of Housing and Urban Development, Policy Development and Research. https://www.huduser.gov/portal/publications/pdf/Locke_AgingInPlace_AssistedHousingRCR03.pdf

Luca D.L. & Cole R. (2017). *Estimating Program Effects on Program Participants*. Evaluation Technical Assistance Brief for OAH Teenage Pregnancy Prevention Grantees. Washington, D.C.: Health and Human Services Office of Adolescent Health. <https://www.hhs.gov/ash/oah/sites/default/files/estimating-program-effects-on-program-participants-brief.pdf>

McGunn, L., & Gifford, R. (2014). Accessibility and aging in place in subsidized housing. *Seniors Housing & Care Journal*, 22(1), 18-29.

MedPAC. (2016). *June 2016 Data Book, Section 4: Dual-eligible beneficiaries*. Retrieved from: <http://www.medpac.gov/docs/default-source/data-book/june-2016-data-book-section-4-dual-eligible-beneficiaries.pdf?sfvrsn=0>

Morse S. (2017, June 30). CMS cancels release of MA encounter data. *Healthcare Finance*. Retrieved from: <http://www.healthcarefinancenews.com/news/cms-cancels-release-ma-encounter-data>

Nichols A., & Schaffer, M. (2007). *Clustered Errors in Stata*. Retrieved from: <http://repec.org/usug2007/crse.pdf>

Nichols, A. (2007). Causal inference with observational data. *The Stata Journal* 7(4), 507–541.

Proctor, B.D., Semega, J.L., & Kollar, M.A. (2016). *Income and Poverty in the United States: 2015*. U.S. Census Bureau Report Number: P60-256.

<https://www.census.gov/library/publications/2016/demo/p60-256.html>.

Research Triangle Institute. (1996.) *Evaluation of the New Congregate Housing Services Program, Second Interim Report*. Washington, DC: U.S. Department of Housing and Urban Development, Policy Development and Research.

<https://www.huduser.gov/portal/publications/suppsvcs/interim.html>

RTI International. (2014). *Support and Services at Home (SASH) Evaluation: First Annual Report*. Washington, DC: U.S. Department of Health and Human Services, Assistant Secretary for Planning and Evaluation. <https://aspe.hhs.gov/pdf-report/support-and-services-home-sash-evaluation-first-annual-report>

Segal, M., Rollins, E., Hodges, K., & Roozeboom, M. (2014). Medicare-Medicaid eligible beneficiaries and potentially avoidable hospitalizations. *Medicare & Medicaid Research Review*, 4(1).

Sun, J. (2006). *The Statistical Analysis of Interval-Censored Failure Time Data*. New York: Springer

Sun, J., & Li, J. (2014). Interval censoring. In J.P. Klein et al. (Ed.) *Handbook of Survival Analysis*. Boca Raton, FL: CRC Press

U.S. Department of Health and Human Services, Office of Inspector General. (2012). *Coding Trends of Medicare Evaluation and Management Services*. Washington, DC.

U.S. Department of Health and Human Services, Administration for Community Living. (2016). *A Profile of Older Americans: 2016*. Washington, DC.

U.S. Department of Health and Human Services, Centers for Medicare and Medicaid Services, Office of Enterprise Data and Analytics. (2017). *2016 CMS Statistics*. Washington, DC.

U.S. Department of Housing and Urban Development. (2016). Picture of Subsidized Households. Database of HUD-assisted households available at <https://www.huduser.gov/portal/datasets/assthsg.html>.

What Works Clearinghouse. (2015). *WWC Standards Brief for Baseline Equivalence*. Washington, DC: U.S. Department of Education Institute of Education Sciences.

https://ies.ed.gov/ncee/wwc/Docs/referenceresources/wwc_brief_baseline_080715.pdf

U.S. Department of Housing and Urban Development
Office of Policy Development and Research
Washington, DC 20410-6000



January 2021