

November 2004



# The Do-It-Yourself Cost-Study Guide

Assessing Public Costs Before and After  
Permanent Supportive Housing:  
A Guide for State and Local Jurisdictions

By Martha R. Burt

# Acknowledgements

This publication was made possible by the generous support of The Robert Wood Johnson Foundation.

CSH is grateful to Tia Martinez and Carol Wilkins for their substantial contributions to this publication, and to Dennis Culhane for providing background information regarding the cost avoidance analysis of the New York/New York Initiative.

Readers may direct any questions or comments about this guide to [info@csh.org](mailto:info@csh.org).

*We encourage non-profit organizations and government agencies to freely reproduce and share the information from CSH publications. The organizations must cite CSH as the source and include a statement that the full document is posted on our website, [www.csh.org](http://www.csh.org). Permissions requests from other types of organizations will be considered on a case-by-case basis; please forward these requests to [info@csh.org](mailto:info@csh.org).*

*Information provided in this publication is suggestive only and is not legal advice. Readers should consult their government program representative and legal counsel for specific issues of concern and to receive proper legal opinion regarding any course of action.*



# Table of Contents

Introduction.....	1
The NY/NY Initiative Analysis .....	3
Other Cost Avoidance Analyses .....	7
Doing Your Own Cost Avoidance Analysis .....	11
Understanding What It Takes to Get Use, Cost, and Savings Information.....	15
Issues, Variations, and Applications .....	40
Appendix: Data Privacy Issues .....	45

## INTRODUCTION

Most stakeholders interested in ending chronic homelessness are familiar with the recently published cost avoidance analysis conducted for the New York/New York Initiative (hereafter, “NY/NY analysis.”<sup>1</sup> The NY/NY analysis looked at the NY/NY Initiative, which placed homeless individuals with serious mental illness (SMI) in a variety of permanent supportive housing (PSH) arrangements, and helped them stay there.<sup>2</sup> It examined three issues—changed **use** of public services, changed **costs** resulting from the changed use, and **savings** (whether PSH cost the same or less than pre-PSH use of public services).

The Corporation for Supportive Housing (CSH) promoted the NY/NY analysis because it thought the results might provide powerful evidence of supportive housing’s ability to produce outcomes that policy makers want—reductions in inappropriate use of emergency public services, and their associated costs. The results fulfilled these hopes, showing significant savings. Together with evidence that PSH is able to attract and retain chronically homeless disabled people (Shern et al., 1997; Tsemberis and Eisenberg, 2000) and that the probable number of such people is small enough to be a feasible policy target (about 200,000 to 250,000 nationally), the NY/NY analysis has galvanized many cities and the country as a whole to adopt the goal of ending chronic homelessness.

### Intended Users of This Guide

Stakeholders in many cities and counties have expressed interest in conducting their own NY/NY analysis. This guide lays out the issues that any jurisdiction needs to think about before committing itself to that endeavor. Its intended readers include city and county agencies; individual homeless assistance providers; homeless advocacy groups; associations and coalitions of providers, advocates, and consumers; foundations and private homeless assistance funders; and planning agencies, commissions, task forces, and committees.

---

<sup>1</sup> Culhane, Dennis P., Stephen Metraux, and Trevor Hadley. 2002. Public service reductions associated with placement of homeless persons with severe mental illness in supportive housing. *Housing Policy Debate*, 13(1), 107-163.

<sup>2</sup> We use “PSH” as the generic term for the housing provided by the NY/NY Initiative, but in fact the housing was of two distinct types. About two-thirds of the 3,600 units were what we usually think of as permanent supportive housing. The emphasis is on the housing, and the right to remain in it is distinct from acceptance of or participation in mental health or other supportive services. The remaining third were more properly thought of as treatment programs—community residential settings more traditional to mental health systems, where participation in mental health services is a condition of remaining in housing. In addition to differences in philosophy and perhaps acceptability to chronically homeless people, the latter require more staffing and are consequently more expensive on a per-unit basis than the former.

Throughout this guide we use the terms “you” and “researchers” interchangeably. Most of the time it makes little difference whether “you,” in your research guise, are a provider, a funder, a city agency, or an advocacy group. However, in the section that describes how to access the service use and cost data needed for analysis, it makes a big difference whether “you” are the agency that already controls the data or someone hired by that agency to do the analysis, or are, instead, a stakeholder who has to gain the cooperation of the agencies that hold the data. In those sections we describe the process of data acquisition separately for stakeholders who need to gain data access and agencies that already have it.

It is unlikely that any state or local jurisdiction will completely replicate the NY/NY analysis, as it was very expensive and time-consuming. Luckily, it probably will not be necessary for a jurisdiction to do that much in order to make a case for PSH within its boundaries. A jurisdiction may choose to examine only changes in **use** of publicly funded services, if that information is sufficient to impress policy makers that investing in PSH is a good idea. The examination may even be limited to one or two services rather than the eight institutions included in the NY/NY analysis, again depending on what local decision makers think is most important. Or a jurisdiction could do **use** and **cost reduction**, without taking the final step to calculate savings.

This brief walks you through the thinking and data gathering needed to make a decision about whether to take on a cost avoidance analysis, and if yes, what scope and complexity your jurisdiction needs and is able to produce. It offers a simplified description of the methods used for the NY/NY analysis, to serve as a baseline against which a jurisdiction can compare its own information needs, data availability, and analytic capability. It also offers for comparison brief descriptions of two similar but less expensive analyses and two simple pilot efforts. These descriptions are followed by a detailed presentation of what conditions need to be met for a jurisdiction to pursue its own NY/NY-type of analysis.

The final section of this brief describes options for producing useful cost data without completely replicating the NY/NY analysis. The section discusses how to think about what data one needs to design a policy, and what data one needs to generate the political commitment to invest in PSH as a normalized government activity.

## THE NY/NY INITIATIVE ANALYSIS

The NY/NY analysis was expensive—\$450,000, not counting the unpaid efforts during the two years before the project officially began that researchers and their partners at CSH devoted to getting permission to use public agency data and other start-up activities. It was complicated and thorough. It used a sophisticated design, examining public costs for eight public agencies,<sup>3</sup> each with its own data system, and constructed unique comparison groups for most. The NY/NY analysis included three components:

1. Identifying service use for two years up to PSH placement compared to service use for two years following PSH placement, for all 4,679 people who received a NY/NY placement.
2. Identifying service use for two years up to a date equivalent to PSH placement compared to service use for two years following that date, for 3,336 pairs, each consisting of one person who received a NY/NY placement and one matched for similar characteristics who did not (data limitations constrained a number of subanalyses to much smaller subsamples).
3. Identifying the costs associated with services in each period and the costs associated with the NY/NY placements, and calculating cost savings by comparing the costs avoided to the cost of PSH.

### **Selecting the Intervention Group and Creating Matched Pairs for Analysis**

Through the NY/NY Initiative, 4,679 chronically homeless people with SMI were placed in PSH over the eight-year period from 1989 to 1997. These became the base sample for the intervention group. Unique sets of matched pairs were created for each public agency in the service use and cost analysis. NY/NY Initiative residents who had used the public agency were first selected. The public agency database was then searched to find someone who matched each of the NY/NY residents on pre-placement service use, and who also appeared in the shelter database of the Department of Homeless Services. Several factors affected the number of matched pairs in the final subsamples for different public agencies—the number of NY/NY residents who had used the public agency in the four-year period under investigation, the ability to find people among the public agency's other users who met the match criteria, and the years for which the public agency had data available. Some public agency subsamples were as small as several hundred matched pairs, despite the large number of people who received NY/NY placements.

---

<sup>3</sup> The eight public agencies included one federal agency (the Department of Veterans Affairs), three New York state agencies (Medicaid, mental health, and corrections), and four New York City agencies (human services, homeless services, locally funded health and hospital care, and city jails).

## **Accessing Public Databases**

NY/NY researchers spent several years gaining access to public agency databases; each agency required separate negotiations. Once a public agency granted access, researchers could begin to search its database for records of service use by each person with a NY/NY placement, and then for matching people. Because people had entered PSH over such a long period, the NY/NY researchers had to search eight years of databases (if they were available) to get data for two years before and after entry into PSH.

## **Analyzing Data**

The basic data extracted from public agency records were number of visits (emergency room and outpatient) or number of inpatient days (for prisons, days of incarceration). Dates of service use were limited to the study period (two years before and after NY/NY placement, or its equivalent for the matched controls). Diagnoses were also extracted, and permitted the researchers to determine what proportion of service use was associated with particular mental illness and substance abuse disorders, whether the visit was for that disorder or not.

Using the number of visits/days and the subsample size, researchers calculated several very simple statistics, separately for the pre-placement and post-placement periods and for all NY/NY placements, the placement members of the matched pairs, and the control members of the matched pairs:

- Number of people using the service even once
- Total number of visits/days
- Average number of visits/days for all subsample members
- Average number of visits/days for the people who had had any visits/days.

These raw statistics revealed (mostly) large differences between the PSH placement and control samples in their use of crisis/emergency, residential, inpatient, and institutional services in the post-placement period.

The goal of the NY/NY analysis was to understand how much change in service use was caused by a PSH placement. Unfortunately, because matching (and even random assignment) never creates completely equal treatment and control groups, one cannot be sure that some factors other than the presence or absence of a PSH placement accounted for the impressive differences in the raw statistics. The NY/NY researchers therefore used statistical techniques (regression analysis) to control for a variety of factors that could also have affected the results. These included gender, diagnosis (schizophrenia, affective disorders, and drug/alcohol dependency diagnosis), and whether a person had used any of the service in the pre-placement period.

## **Obtaining Cost Data**

Except for the costs of PSH, the NY/NY analysis used per diem (or per visit, for outpatient Medicaid costs) cost information provided by the public agencies supplying

the service use data. Cost data for shelter days was available through New York's Department of Homeless Services, which pays for 80 percent of the shelter bed nights available in the city through contracts with private providers. Cost data for hospitalizations and outpatient care was available from the various health agencies, or from reimbursement sources such as the Medicaid office. Corrections and Veterans Affairs costs were likewise available from those public agencies themselves. It seems relatively simple to get these per diem/per visit costs, but it is less simple to be sure that using them will give us "apples to apples" comparisons—that is, to be sure the public agency costs and the PSH costs include or exclude the same cost components. We'll say more about this after describing how the costs of PSH were calculated.

The cost data the NY/NY researchers spent the most effort to acquire were the costs of PSH. Because the PSH in the NY/NY analysis was all part of an organized state-city program, the NY/NY Initiative, the NY/NY researchers were able to identify cost data in consultation with the state and city agencies that provided the money (some of which was federal and under the control of the state and city agencies). Researchers accounted for construction and renovation, debt service, operating, and supportive services costs for PSH. If tenants contributed rent (usually through a housing subsidy such as Section 8), the amount paid by the tenant and the amount paid by Section 8 were both calculated. Even with the consolidation of all the PSH under study within one initiative, many different providers actually ran the different PSH sites, under contract with different public agencies for different components of costs (capital, operating, and supportive services). Contract terms differed, and different providers had different sources of funds. Thus many budgets had to be reviewed and key informants interviewed to be sure that all relevant costs were being accounted for properly.

For the final cost analyses, researchers first figured out how many fewer units of service PSH residents used, compared to the control groups. For example, suppose they determined that those in PSH used 15 fewer inpatient psychiatric hospital days during the study period than their matched pairs. Researchers would then multiply the per diem cost of an inpatient day by 15 to arrive at the reduction in inpatient psychiatric hospital costs due to PSH placement. These calculations were repeated for each type of public agency use.

Researchers then added up all public agency costs and cost savings, annualized them (they had initially been calculated for a period of two-years), and calculated costs *per placement* and *per unit*. This distinction is important because, on average, turnover resulted in more than one person being placed in some units in a year's time (4,679 people were placed into 3,600 units between 1989 and 1997). Researchers presented annualized per unit costs avoided for each of the eight public agencies in the analysis, and compared these with the annualized cost of housing someone in a PSH unit.

Note that in the NY/NY analysis the PSH cost figures include capital and debt service costs. Further, their operating and service components tended to be calculated by taking total program budgets and dividing by the number of units. It is highly unlikely that the per diems obtained from public agencies included as comprehensive an array of costs, as public agencies tend to handle their capital and debt service expenditures through budgets that are completely separate from their operating budgets. One can just imagine how the NY/NY analysis results might have differed if capital costs had



either been included in the public agency rates or excluded from the PSH costs, to see how important these complications may become. Later in this paper we discuss the issues involved in getting as close as possible to “apples to apples” cost comparisons.

## OTHER COST AVOIDANCE ANALYSES

We include brief descriptions of several other cost avoidance studies to show the range of what has been done to date and how much mileage can be achieved with relatively limited and focused efforts. Two studies—Connecticut’s Supportive Housing Demonstration Program and the cost analysis of San Francisco’s Health, Housing, and Integrated Services Network—were fairly extensive while three were fairly limited as to sample size, time period examined, and number of programs and agencies involved. But each may serve as an example of what a local jurisdiction can do with cost information.

### **Connecticut Supportive Housing Demonstration Program**

Starting in 1992, the State of Connecticut and the Corporation for Supportive Housing (CSH) joined forces to promote the Connecticut Supportive Housing Demonstration Program, which ultimately produced 281 units of PSH in nine projects located in six mid-sized Connecticut cities.<sup>4</sup> The first tenants took possession in June 1996; 444 individuals received placements between then and February 2001. The projects are intentionally mixed-use; 70 percent of the units are reserved for individuals who were formerly homeless or at risk of homelessness, and about half are reserved for people with identified special needs, defined as HIV/AIDS, mental illness, or a history of chronic substance abuse. According to enrollment data, only 23 percent of enrollees had been in situations that would qualify them as literally homeless (shelters or streets) just before moving into PSH. Most had been living independently, with family, or in congregate living situations. Arthur Andersen LLP, together with several consultants and subcontractors, conducted an evaluation of the program, including a cost analysis. Key components of this study’s cost methodology, including which aspects of the NY/NY analysis it did not do, were:

- Of the 430 people who were placed in PSH and completed enrollment forms during the study period, the sample for the cost analyses included only the 126 people who also were Medicaid beneficiaries and signed a consent for access to their Medicaid data.
- Comparisons were made for residents’ public agency service use for two years before PSH placement and three years after. No control or comparison groups were used.
- The service use and cost analyses used only one public database, for Medicaid-reimbursed services. Results were analyzed separately for eleven types of care, grouped into four major categories: 1) substance abuse and medical inpatient

---

<sup>4</sup> Arthur Andersen, LLP; Center for Mental Health Policy and Services Research, University of Pennsylvania; Kay E. Sherwood; and TWP Consulting. 2000. *Connecticut Supportive Housing Demonstration Program: Final Program Evaluation Report*, New Haven, CT: Corporation for Supportive Housing.

treatment; 2) pharmacy, home health care, and medical equipment; 3) ambulatory substance abuse, medical, psychiatric, and case management assessment and monitoring; and 4) physician care, care by other providers, and dental care.

- To guard against false positive results because no control group was used, researchers divided the “before” and “after” periods into 6-month increments, four occurring prior to PSH placement and six following placements, to create a time series. Results mostly showed that service use was consistent in each period before PSH placement, and increasingly different in each period after a person entered PSH. Documenting consistent patterns of pre-PSH service use and post-placement changes lends more confidence to the findings in the absence of a control group.
- No calculations were made of the cost of the PSH itself, including on-site medical, mental health, substance abuse, or other supportive services. Nor were use or cost data obtained for use of emergency shelter or other residential options within the homeless assistance network either before or after PSH placement.

Results showed that tenants “decreased their use of acute and expensive health services, mostly medical inpatient services” and increased “utilization of necessary on-going health care and support...such as home healthcare and outpatient substance abuse treatment services, that enabled tenants to remain in the community, rather than be hospitalized.”<sup>5</sup> More tenants used appropriate outpatient services and fewer used inpatient services, increasing the efficiency of health care use.

### **San Francisco’s Health, Housing, and Integrated Services Network**

San Francisco is home to many PSH programs, several of which are linked to each other through a PSH network—the Health, Housing, and Integrated Services Network (HHISN). A CSH study<sup>6</sup> of supportive housing programs in San Francisco used limited components of the NY/NY methodology to document cost savings to the public health and mental health systems attributable to PSH. Key components of these studies’ methodology, including which aspects of the NY/NY analysis it did not do, were:

- The sample included only the residents of two PSH buildings that were part of the HHISN, and was small compared to the NY/NY analysis—253 people in the base sample, and as few as 95 for selected analyses.
- Comparisons were made only for residents’ public agency service use for two years before PSH placement and two years after. No control or comparison groups were used.

---

<sup>5</sup> Andersen, et al., p. 33.

<sup>6</sup> Early results are summarized in Proscio, Tony. 2000. *Supportive Housing and Its Impact on the Public Health Crisis of Homelessness*, New York: Corporation for Supportive Housing. Later follow-up results described here are available in Martinez, Tia and Martha R. Burt. 2003. *Changes in Service Use Patterns for Chronically Homeless People Placed in Permanent Supportive Housing*. Oakland, CA: Corporation for Supportive Housing.

- The service use and cost analyses initially focused on only two public systems of care. The computerized data systems for these two public agencies contain records of care paid for through Medi-Cal (California’s Medicaid program) and “uncompensated care” paid for by local government resources:
  1. Inpatient and emergency room care from San Francisco General Hospital and outpatient care received from the primary care clinics operated by the San Francisco Department of Public Health (together, the hospital and clinics are referred to as the Community Health Network, which also includes Laguna Honda Hospital, San Francisco’s public long term care facility)
  2. Mental health care from San Francisco Community Mental Health Services (the county mental health system, which also includes services run by nonprofit providers under contract with the county).
- Later analyses extended these examinations to include the county jail, jail medical and psychiatric services, and substance abuse services.
- To guard against false positive results because no control group was used, researchers divided the “before” and “after” periods into four one-year periods and showed that service use was basically similar in each of the years before PSH placement, and quite different in each year after a person entered PSH. The inference is that in the year immediately before entering PSH, sample members showed typical patterns of service use, not unusual patterns. Documenting the consistency of pre-PSH service use lends more confidence to the findings in the absence of a control group.
- No calculations were made of the cost of the PSH itself, including on-site medical, mental health, substance abuse, or other supportive services. Nor were use or cost data obtained for use of emergency shelter or other residential options within the homeless assistance network either before or after PSH placement.

Results were impressive, even with these significant deviations from an ideal evaluation design and calculation only of costs avoided in two public systems, not cost savings overall. As more cost avoidance/savings studies are done and obtain similar results, each one reinforces the others, increasing confidence in the results even of studies with limited designs. For residents of the two San Francisco PSH programs studied, emergency room visits fell by half (from approximately two to less than one per person per year). Inpatient days went from 441 in the year before PSH to 278 in the year after. For those in PSH for two full years, the second year of PSH saw an additional 15 percent reduction. For the year that began 24 months before PSH placement (i.e., months 13-24 pre-placement), days of residential mental health treatment outside of hospitals had been 465 for the 177 people for whom informed consent was obtained to search the relevant records, and for whom data were available. These days went to 415 during the year just before PSH placement, and to none during their first full year of residence. Costs avoided were comparable.

## **San Diego's Highly Focused Cost Studies**

San Diego has a Homeless Outreach Team (HOT) Program that helps chronically homeless street people with serious mental illness to leave the streets. When HOT was still a pilot (it is now county-wide, thanks in part to its financial analysis), the program asked the UCSD Medical Center to track care received at Scripps Mercy Hospital, UCSD Medical Center, and the San Diego emergency medical services system by 15 mentally ill chronically homeless individuals well-known to the HOT workers during an 18 month period from July 1, 1997 through December 31, 1998. The 18-month cost for these 15 individuals alone was just under \$1.6 million, including physician fees but not including care that might have been incurred at other regional hospitals. HOT's expenditures of \$69,820 a person a year, to get chronic street homeless people with multiple disabilities into stable housing situations represented a substantial savings from the alternative emergency service expenses.

San Diego also has a Serial Inebriate Program (SIP), which arrests chemically dependent homeless street people and, following a drug court model, offers them a choice of jail or treatment (California law stipulates up to six months in jail for public drunkenness). SIP staff did a cost analysis showing the very high levels of public resources being absorbed by serial inebriates, compared to the costs of treatment. They calculated a cost of \$977 for maintaining one serial inebriate in housing plus outpatient substance abuse treatment for one month, compared to \$1,470 for the cost of one police contact with an ambulance visit to the emergency room followed by a day in jail. Based in part on this analysis, local dollars were committed to make adequate treatment resources available. SIP participants are all people with very long histories of addiction and street homelessness, and all were treatment resisters. Nevertheless, over three years the program has an average annual substance abuse treatment completion rate of over 40 percent compared to average completion rates for voluntary substance abuse treatment programs for homeless people of between 20 and 33 percent, with most falling at the lower end.

## DOING YOUR OWN COST AVOIDANCE ANALYSIS

Jurisdictions thinking about doing their own cost avoidance analysis for their PSH programs can begin by asking themselves a number of questions. The first two questions are preliminary; the last three questions focus on the three major outcomes you might want to document—use, cost, and savings. We list the five questions first, then examine each in turn, discussing approaches for and challenges of getting the different types of data that are needed to answer each question as the questions arise.

1. Does the PSH we have meet minimal criteria for a cost avoidance analysis?
2. What types of people are in our PSH?
3. (Use) Does living in PSH change residents' use of public crisis services?
4. (Cost) How much less are public agencies spending for people living in PSH, compared to public cost in the year(s) before PSH?
5. (Savings) Does the money saved from reduced use of public crisis services equal or exceed the cost of PSH? (Is PSH cost-effective?)

### **Does Our PSH Meet Minimal Criteria for a Cost Avoidance Analysis?**

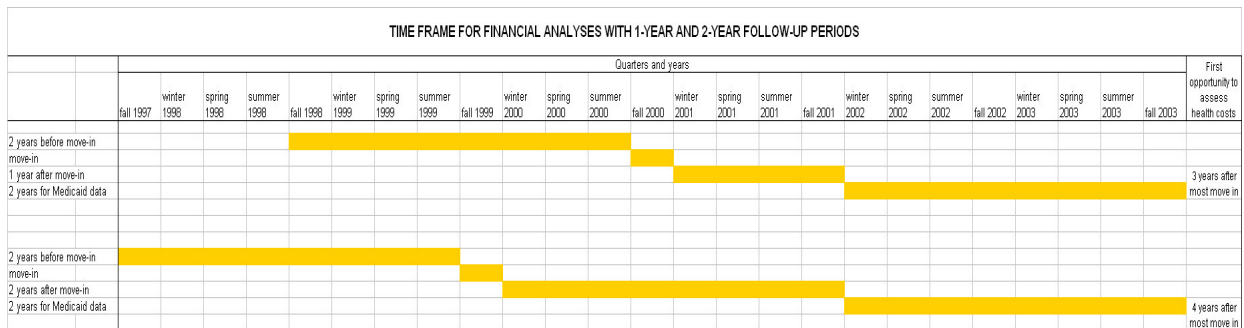
The basic criteria relate to how many units of PSH you have and how long you have had it.

- Your jurisdiction actually has to be providing some PSH to a reasonable number of people.
- The PSH, and the people in it, have to have been around for at least two to three years.

Whether you are only trying to measure changes in service use or are also trying to assess costs and savings, you need to have a reasonable number of people in PSH to include in the analysis. You should be trying for not fewer than 100 PSH units and occupants; two or three times that number is desirable, if possible. The more people who are in the intervention condition, the less likely you are to have a few people whose extreme behavior changes the results for the whole group. Sometimes smaller samples can be used to provide information meaningful to policy makers, but it is extremely difficult to get statistically significant results with small samples.

Further, the date that a reasonable number of people entered PSH generally has to be at least two to three years earlier than the date when you want to conduct your analyses. This is true for two reasons. First, you want people to have been in PSH at least one year, to give PSH a reasonable length of time to make a difference to its residents' service use. Some evidence indicates that two years is even better (see results for HHISN, summarized above). Second, it takes at least 12 months (and

sometimes up to 24 months) for service and cost data to become reasonably complete for Medicaid and other data systems that rely on billing information submitted by service providers.<sup>7</sup> Medicaid data, in particular, are necessary to assess use of medical, mental health, and some substance abuse services. The combination of waiting 12 to 24 months for Medicaid and other relevant data and preferring one or two full years of PSH residence as a requirement for inclusion in the financial data analysis dictates that the intervention group have entered PSH at least two to three and ideally up to four years ago. It is possible to do your analysis more quickly IF you are reasonably confident that billing information in the central databases is up to date and complete, but you will miss services and costs if the data do not meet those criteria.



The sample “count backward” timeline displayed above gives a quick picture of the time frames involved. It starts with allowing two years for Medicaid data to become available for the period after move-in. It then gives the period of tenancy after move-in—one year for the top set of bars, and two years for the bottom set of bars. It shows a three-month move-in period (which is actually pretty short for rent-up in many PSH projects). Finally, it shows a two-year period before move-in, which is what you will be using as the comparison to your tenancy period.

For a PSH project to be included in a financial analysis beginning in Fall 2003 that wants to cover one year after move-in, if you want to avoid missing data because of routine billing delays, the *last* tenant included in the analysis has to have moved into PSH no later than Fall 2001. (To assure that all Medicaid data is complete, the last tenant should have moved in by Fall 2000. If you want to cover two years after move-in, the last tenant has to have moved into PSH by Fall 1999.

<sup>7</sup> This is because hospitals or other service providers may be permitted to bill for services at any time during a fiscal year or contract period, so some service utilization information may not be submitted and recorded in central data bases until the end of the fiscal year. Also eligibility for Medicaid or other benefits may be pending at the time of service delivery. In these cases, when retroactive eligibility is established, service providers may bill for services provided up to 24 months earlier.

## What types of people are in our PSH?

The point of a NY/NY-type analysis is to show that homeless people cost the public a lot if they are left homeless—possibly nearing or equaling the cost of housing them and offering supportive services that help them keep their housing. The analysis shows this cost differential by looking at the year (or two years) before PSH move-in and the year (or two years) after move-in.

Several factors may affect whether you find big service use and cost reductions as a result of PSH, and also whether you will be describing changes attributable to *ending homelessness* or attributable to other effects of PSH. Chief among the factors to consider are whether the people receiving PSH used lots of very expensive services before moving into PSH, and whether they did so while being homeless. So it is very important to know where your PSH tenants were before they moved into PSH—were they homeless or in other publicly funded settings, and for how long—so you can characterize the comparison accurately.

Given the costs associated with PSH, the “before” period has to be one of high publicly funded service use for one to be able to demonstrate net savings. Generally, the people who will have high use levels are chronically homeless people with one or more major disabilities. The more your PSH tenants meet this description, the more likely you are to find that the costs of supportive housing are offset by savings that come from reduced service use and costs in other systems. Conversely, if your PSH tenants have fewer major disabilities (i.e., they do not have a severe mental illness, long-standing substance abuse disorder, various physical disabilities, HIV/AIDS, and/or other high-service-use conditions) they are less likely to have used expensive public services and your analysis is less likely to be able to show substantial cost savings.

It may be that significant numbers of your PSH tenants were never homeless or not homeless for long periods of time, but were in institutional or other high-cost service settings. Then the analysis may be able to show savings after PSH – but it will not be accurate to describe this as a comparison between the costs of homelessness and the costs of PSH.

If your goal is to compare the costs of PSH to the costs of homelessness, you will be able to do this best if your PSH tenants, and anyone you want to use as part of a comparison group, have been chronically homeless (so that they were homeless for all of the “before PSH” time period). Otherwise even if you are able to get data from several public agency databases for service use over the year or two before PSH placement, it will not be clear how you should interpret this information for people who were not homeless for most of that period. It should be relatively easy to tell whether people in a particular PSH program have a long history of homelessness. This should be enough if you want to do a program-specific analysis. If you want to include a comparison group of people who did not receive PSH, you will also need to have a way of knowing that these comparison people have equally long homeless histories, and that they continued to be homeless during the period after PSH placement for the tenant population.

Finally, the more heterogeneous the group of people who receive an intervention (in our case, PSH), the bigger the group needs to be to see reliable results because there will



be a lot of variety among people's histories before moving into PSH. If everyone in your PSH has serious illness with co-occurring substance abuse and a 10-year homeless history, you will need fewer people to see reliable patterns than if you have some with mental illness, some with only substance abuse, some with both, some with HIV/AIDS, some with chronic physical illnesses, and so on. Also, for interventions that have small impacts or different impacts for different types of people, you will need a big group to get measurable results. In PSH, because participating in supportive services is not a condition of tenancy, some tenants may be very involved in supportive services and may experience large changes shortly after moving in, while others may be harder to engage and may achieve changes more slowly and require greater service input before that happens.

### **A “Bonus” Outcome—Changes in Tenant Mix**

Although not part of the NY/NY analysis, communities seeking to conduct a cost avoidance analysis may also find themselves in the position to document an “extra” outcome—increasing use over time of PSH for (formerly) chronically homeless people with major disabilities. At the least, understanding who occupies PSH for the cost avoidance analysis will give you a baseline against which to track changes in tenant mix. In mid-2003, CSH's best estimates are that perhaps as few as 20 to 30 percent of PSH tenants are people with major disabilities who were once chronically homeless. Other PSH tenants include people with disabilities who were at risk of homelessness, people who were previously living in more restrictive settings, and people who have been homeless for shorter periods of time. Communities that start off with a higher percentage of tenants who were chronically homeless or those with the most severe disabilities who have a history of using high-cost services are more likely to be able to document cost advantages of PSH. It also follows that if communities are able to increase the percentage of PSH units occupied by people who have been chronically homeless over several years through targeting and selection criteria, and serve them effectively, they will see improvements in their savings of public dollars attributable to PSH.

## UNDERSTANDING WHAT IT TAKES TO GET USE, COST, AND SAVINGS INFORMATION

A community may decide to pursue one or more of the three components of the NY/NY analysis—use, cost, and savings. Cost and savings calculations depend on having service use data, so if you are only doing one component, that will be service use. Targeting, a fourth possibility not pursued in the NY/NY analysis, was already described—whether PSH is reaching its intended target population(s) of homeless people with disabilities, including those with the most long-term homelessness or those who face the greatest barriers to housing stability. Which you choose to do will depend on your assessment of (1) whether service use and cost data from public agencies exist in retrievable formats, (2) whether the data are or can become accessible for analysis, (3) whether you can obtain comparable unit cost data from public agencies and PSH, and (4) what types of data will make the most powerful impact on your community’s policy makers.

Before discussing any of these issues, however, it is essential to consider requirements for accessing any type of public data linked to individuals. This issue, often referred to as “informed consent” or “data privacy,” is so important that we discuss it first, as an inquiry’s total strategy will depend on whether or not the researchers need to obtain informed consent to access public data from the people in PSH whose use of public crisis services you want to track. In most situations, analysts will face this issue whether they work for the relevant public agencies as staff or contractors or are completely independent.

### **Understanding the Need for Tenant Identifiers, and Hence for Informed Consent**

Generally speaking, any economic analysis you undertake will involve using personal identifiers to match the people in PSH to their records of service use from one or more public agencies. At a minimum these identifiers will include each individual’s first and last names, gender, and date of birth. Either a full social security number (SSN) or the last four digits of the SSN are also desirable. The NY/NY researchers used the first four to identify homeless people in the various public agency databases they searched, and used SSN as a backup match element in case the first four did not yield a definite match. At first glance one would think that SSN would be sufficient, but SSNs may leave a lot of room for error, both in the way people give them and the way they are recorded. One can always try to use the SSN first, but have the other information available as backup—you will need it.

Because researchers will be dealing with data that can identify individuals and put them at some potential risk, the project will have to develop appropriate procedures to protect the privacy of PSH tenants (and comparison group members, if any) and assure the confidentiality of all data pertaining to them. These procedures fall under the general rubric of getting ***informed consent*** and ***protecting data privacy***.

Many public agencies have legal or regulatory obligations that limit their ability to provide data with identifiers to people who are not agency employees or agents. Further, recent federal regulations known as HIPAA (after the Health Insurance

Portability and Accountability Act of 1996, which required them) place serious constraints on access to certain types of health-related information. Restrictions also affect the ability of the public agencies involved to share each others' data at the individual level.

That being said, there are many legitimate reasons why public agencies and external evaluators want to access public administrative data, and many circumstances under which access is allowable. Allowable situations (described in more detail in the Appendix) include use of de-identified data sets and those with limited identifiers, and procedures that have received the approval of an Institutional Review Board (IRB) established for the purpose of protecting data privacy. Standards for informed consent are particularly high with respect to mental health and substance abuse service data.

Analysts (including those who work for public agencies) should be sure they know what these standards are and develop plans to meet them, or they may find themselves unable to access data about important aspects of care. Your approach for obtaining informed consent should be reviewed by an appropriate authority, usually and Institutional Review Board (IRB). All universities, many research firms, and some public agencies have their own IRBs. Most, but not all, research situations require researchers to obtain the informed consent of the people whose records are to be examined.

Under certain limited circumstances, such as when *all* of the records to be used, including the records of shelter use and PSH tenancy, are maintained by the same public agency or agencies that are conducting or sponsoring the data analysis and the actual analysts are working as agents of these agencies, the law does not require informed consent from each person in the analysis. This was the situation in New York, where the Department of Homeless Services maintains a database of emergency shelter users going back more than a decade, and tenancy in NY/NY housing was a matter of public record. Thus researchers were able to identify both “treatment” and case control individuals from public databases, without having to go to homeless assistance or PSH providers. Basing their approach on provisions of city and state regulations that allow public administrators to use their own data for purposes of managing and improving public agency performance and efficiency, which was an avowed purpose of the NY/NY analysis, the NY/NY researchers could match people within and across the different public agencies included in the study without having to obtain informed consent from the thousands of program participants and case controls for whom data were examined.<sup>8</sup> Even so, it took many months to work out the agreements for data access.

---

<sup>8</sup> Having direct conversations about requirements and procedures with the agencies whose data are desired for the study is the best way for researchers to learn what restrictions and requirements apply in their state. These agencies should be cognizant of the state and federal regulations that govern their behavior. These conversations are essential if the researchers are not agency employees or agents, but they are also worthwhile if the analysts work for a public agency, as there will still be issues of cross-agency data matching to be handled. The important thing for researchers to know going into these discussions is that *ways exist to do this* – you just have to find the ones that will work where you are.

Very few communities in the country will be in the NY/NY analysts' situation. That is, they will not have any public administrative data that can tell them who is in the "treatment" condition of living in PSH. Nor will any public administrative data be able to help them identify a comparison group of people who were homeless in the same pre-treatment period. The vast majority of emergency shelter and transitional and PSH programs are run by private nonprofit providers, with only the provider being knowledgeable about who has been in residence and when. These providers do not have the right to convey individual identifiers of consumers to public agencies without those consumers' consent.

As new HMISs are implemented, increasing numbers of communities will maintain information about the people who use residential homeless assistance programs in a community-wide database. Different communities will have different rules governing who can access these HMIS data; if yours is one that gives access to one or more public agencies, your requirements for informed consent may change in future years. But it will be several years before HMIS data are maintained consistently enough to provide comparable data for the years before and after homeless people move into PSH, and most data systems being developed will never be able to capture homeless assistance system use for the years before they began to collect data. So, while an HMIS may make the job of conducting an economic analysis easier in the future in some communities, each situation will still have to be examined carefully to see what types of privacy protections may still be needed.

There may be "in-between" situations that would take review by an IRB to determine whether or not informed consent is needed. For instance, say a public agency such as a health department funds services at a PSH program. Could the PSH program, without getting informed consent, supply the health department with tenant identifiers and move-in dates (but no further information about tenant characteristics) so the health department could analyze the cost offsets in other parts of its health system using its own staff or consultants? One might argue that this does not pose great risk to the tenants, that the health department would not be revealing sensitive health data to the PSH program, and that the PSH tenants are already known to the health department as they are clients and receive department-supported services. We are not an IRB, so we can't say whether this scenario would be acceptable. An IRB will need to review the specifics of this approach, but we believe in many cases it may be acceptable.

For better or worse, the need to obtain informed consent from PSH tenants (and if applicable, from a comparison group of homeless people) to access public records will be a fact of life for most jurisdictions interested in pursuing economic analyses of use, costs, and savings. *This will be true regardless* of who the analysts are, what type of access they ultimately get, and what data they get access to—because the issue is as much one of revealing to public agencies the identities and current circumstances of PSH tenants (and possibly comparison group members) as it is revealing to researchers what public agencies know about those tenants. As long as public agencies have no way of knowing who is or is not in PSH (or in a comparison group) unless the researchers or the PSH programs themselves reveal this information, the study will need informed consent to convey identities to public agencies as well as to search their records.

## Who Might the Analysts Be, and What Data Formats Might They Obtain?

The people doing the analysis of public administrative records for a NY/NY-type cost analysis are likely to fall into one of three categories, each with advantages and limitations:

- They are regular employees of the public agencies whose data will be analyzed;
  - *Advantages:* easy access to the full universe of agency data, and experience manipulating it.
  - *Limitations:* does not know which people are in PSH (or comparison group) or when they moved in, and must receive this information and relevant identifiers from outside source. Does not have other PSH tenant information that can help to explain observed service use patterns (e.g., length of homelessness, nature of disability) unless this too is received from an outside source.
- They are in a business relationship with the public agency. For example, the analyst could be in the budget office that provides oversight of the agency's spending, or work for one public agency that has an agreement with the other public agency(s) to share data for program planning and administration, or be a consultant engaged by the public agency and working "as an employee" to do this analysis, or have entered into a formal partnership with the agency;
  - *Advantages:* potentially as easy access to the full universe of agency data as regular agency employees; may or may not have as much experience manipulating it. Once matches are made, may be able to make a data extract and take a mostly de-identified data set outside the agency for further analysis, with appropriate data security safeguards.
  - *Limitations:* could be same limitations as for regular employees, but a contractor might also be able to fulfill both "inside" and "outside" roles, by working within the agency on administrative data sets *and also* being the ones to obtain informed consent, identifying information, and other personal characteristics from PSH tenants (and comparison group members).
- They are completely independent of the agency or agencies involved, such as being at a university or private research organization.
  - *Advantages:* Would be the ones to obtain informed consent, identifying information, and personal characteristics from PSH tenants and comparison group members.
  - *Limitations:* would have to negotiate agreements with public agencies for data access. The result could be a memorandum of understanding granting access as free as a "business relationship," including permission to take a de-identified data extract outside the agency for further analysis. At the most restrictive extreme, however, it might be limited to receiving only aggregated data from public agencies. In such a case, the agency would not allow the analysts access to its data, but would be willing to take

tenant identifiers and move-in dates (the critical pieces of information for reviewing agency files for this analysis) and have its own analysts search its databases, conduct the analyses, and assemble the results in aggregate form. The research project analysts would receive only these aggregated data—they would know what proportion of PSH tenants (and comparison group members) had emergency room visits, hospitalizations, and so on, but would not know which specific tenants used these services. They therefore would not be able to conduct analyses linking client characteristics to changes in service use.

Obviously the best of all worlds would be to have agency analysts working closely with the people who will obtain consent and information on individuals. Then the public administrative data can be manipulated by the people who know it best, and other analysts can take it and pursue more complex analyses to figure out what it all means and how it might relate to public policy. We'll return to some of these issues below, when we discuss the practicalities of obtaining agency data. But for now the important thing to note is that these issues will have to be resolved before you can write your procedures for obtaining consent and for assuring data privacy, which you will need to take to an IRB for approval.

### **Actually Getting Consent**

Consent can only come from the individuals themselves (unless they are incompetent to give it, in which case it must come from a conservator or guardian). It cannot come from the PSH program, as the program does not have the right to give it. Strict expectations apply for what people must be told about why the researcher needs the data, what potential risk there might be to the person if he or she gives permission, how that risk will be minimized, what benefits might accrue to the person or to others, and so on. We include a sample consent form in the Appendix that covers all of these points, from a project involving families in PSH in San Francisco.

Further, “normal” informed consent, such as one would get to conduct an interview with someone, is not enough when you need the individual’s permission to search public records. The requirements escalate in this case—which is our case in this guide. Releases must be very specifically worded, must name each agency for which permission is desired and must get explicit permission for each agency, must get explicit permission for accessing mental health and substance abuse diagnoses and treatment, and so on. Again, the Appendix contains an example of such a release form being used in the San Francisco family PSH project.

Any informed consent and release forms you develop should be reviewed by the agencies from which data will be obtained and by an IRB for compliance to relevant rules and regulations. In addition, it is a good idea to have the PSH program staff review these forms also, so they know what their tenants are being asked to sign and can provide feedback to the researchers about phrasing, issues likely to be red flags, and so on.

Many practical issues will need to be resolved before you are ready to begin asking tenants for their consent. These include who will actually do the asking, when, where, what else will be going on at the time, and so on.

*Who will do the asking* could be PSH program staff, case managers, researchers, and possibly others. Each has advantages and disadvantages, including who has the time to do it, who has the best access to and rapport with tenants, what might happen to staff-tenant rapport if tenants might see staff asking for consent as compromising trust or confidentiality, and how staff time is to be compensated if staff do the asking. Research project staff will have to work out these issues with PSH programs.

*When and where* tenants are asked for their consent depends in part on who is doing the asking. Case managers can ask during case conferences, but that time and place is not open to research staff.

One approach to seeking consent if the one doing the asking is not a staff member of the PSH program is for a research team member to ask tenants to complete a brief survey, for which they would be compensated for their time. The survey would seek information of direct value to the economic analysis, such as their history of homelessness, service use while homeless, prior experience with PSH and other homeless assistance programs, and so on. Consent for that survey would be built into the approach to the tenant, and the tenant proceeding with the survey would be sufficient informed consent. At the end of the survey, the interviewer would thank the tenant for the completing the survey, give the tenant the compensation agreed to, and then ask the tenant whether he or she would also give permission for the public records search, using the permission/release form in the Appendix, modified to fit the specific jurisdiction and agency data needs. The tenant would be free to refuse this consent, but would still receive compensation for completing the survey.

Researchers also need to consider whether the people who give consent will be representative of all of the PSH tenants (or of a comparable group of homeless people if you are using a comparison group). Culturally competent outreach strategies and modest financial incentives for participation in the initial survey can help get a good participation at that initial stage, which would then increase the odds of getting a reasonably representative sample. Even if seeking consent is the responsibility of program staff, care still needs to be taken to monitor representativeness—it is always easier for staff to ask the cooperative tenants or the tenants who have been there longest or whom they know the best, rather than all tenants.

Just as the information conveyed to PSH tenants in asking for their consent must be clear, understandable, and honest about the risks involved to the tenant, the procedures must be open and most especially non-coercive. The tenants must be free to refuse, and must feel free to do so if they so choose. You will need to describe these procedures clearly to an IRB and get approval before approaching anyone to ask for consent.

### **USE – Does living in PSH change residents’ use of crisis public services?**

Because the ability to calculate cost and savings depends on obtaining data on service use levels, we first discuss *use*, raising many practical issues as we proceed. You will want to describe service use by PSH tenants during the period before they moved into PSH and during their period of residence, and then compare the two to get changes in service use. The assumption is that while homeless, before PSH, the target population

will have used expensive crisis services such as emergency rooms, emergency psychiatric hospitals, other health services, and detoxification centers; the shelter system; the resources of law enforcement and corrections; and, if families, child protective services and foster care. In contrast, we assume that once in PSH, use of these expensive emergency services will decrease, even if use of (less expensive and more appropriate) public outpatient services and services attached to the PSH increases. Many communities could get considerable political mileage just out of demonstrating significant changes in crisis service use patterns, even without being able to attach costs to them.

If you do not have a comparison or control group, you will not be able to attribute all changes in use to PSH placement, but simple data on changes in service use patterns can still be powerful to most policy makers. Further, it may be sufficient to document changes in service use for only one or a few services or institutional types (e.g., emergency rooms and inpatient hospitalization); you need not think that if you cannot get use information from all the different types of agencies included in the NY/NY analysis, you cannot do anything. Two of the examples described earlier (HHISN in San Francisco and the Connecticut projects looked (initially in HHISN's case) only at limited types of data, but still produced results with policy relevance and policy effects.

### **Determining Which Public Agencies Have Usable Data**

Data on service use will have to come from the public agencies that deliver or pay other agencies to deliver the services—the health, mental health, substance abuse, shelter/housing, and/or corrections agencies that homeless people use. The person or organization that will be gathering and analyzing data may work for one of these public agencies, or have established a partnership or “business relationship” with the public and nonprofit agencies that finance and deliver a range of services. Because it is unlikely that the same agency will control all of the data needed to complete the analysis, most communities will find it necessary to establish agreements to access and share information across agencies that will link tenant data from supportive housing programs with the same people's data in larger data systems that track service use in a range of settings.

To include a public agency in the analyses, it must be able to track and retrieve units of service use. This means days or nights of stay for hospital inpatient, shelter, and corrections agencies, and appointments/ sessions/treatments for other agencies. The public agency has to be able to do this tracking with sufficient detail to determine how many days or units of services each individual has received over the time periods included in the analysis. So the public agency *must* be able to track these service units:

- ***ELECTRONICALLY, and***
- ***IN A WAY THAT CAN BE LINKED TO EACH PSH TENANT, and***
- ***IN A WAY THAT CAN BE LINKED BY DATES OF SERVICE.***



## Gaining Access to Public Agency Data

During our discussion of informed consent we began to describe the various forms in which economic analysis researchers might have access to public agency data. These include:

- The researcher has access to the entire service/billing database. The researchers then must find records belonging to the study sample within the larger database. Having the entire database means the researcher also may have the opportunity to compare results from the study sample to those from a larger group, or possibly even select a matched comparison group from the larger database.
- The researcher only has service use data for each person in the study sample, but identifiers are still attached at the individual level. In this case, agency analysts will have taken the identifiers supplied by the researcher and made an extract from the entire service/billing database to create the analytic file given to the researcher. As the researchers still have identified data, they can attach other explanatory variables to individual cases and conduct some analyses beyond simple counts of service use and changes.
- The agency creates a file of individual-level records, using individual identifiers to match service use information with personal characteristics (which it has received from the same source that supplies the identifiers). The agency then strips the data of individual identifiers and gives the researchers individual-level records with all the information necessary to conduct both descriptive and explanatory analyses, but without exposing individuals to risk of disclosure of personal health or other sensitive data because identifying information is no longer on the file.
- Agency provides aggregated data only, giving proportions of the study sample who experienced particular types of service, but not providing data at the individual level.

Researchers could have access to the entire service/billing database at the public agency itself, either because they are agency employees or have a business relationship with the agency. Alternatively, after appropriate safeguards are in place for assuring the security of the data,<sup>9</sup> a public agency may give a researcher a copy of the entire service/billing database to analyze outside the agency. The same options are available for access to a data extract containing only data on the study sample, which would still be at the individual level.

If no arrangements can be made for the analysts to use public agency data complete with identifiers at the individual level, a public agency may still be willing to do one of two things, both of which involve removing identifiers from the records. The first is a de-identified data set still at the individual level. For this option, the analysts supplying the individual identifiers and move-in dates would also supply personal characteristics in addition to PSH tenancy that might help explain service use patterns. The public agency

---

<sup>9</sup> Including such procedures as locked file cabinets, secure password-protected servers, dedicated computers with limited access, and designation of one or two people as the only ones to work with the data, after they have signed a confidentiality pledge.

would extract service use records for a specific person and link them with that person's own characteristics by means of individual identifiers. Once linked, the identifiers could be removed and the data set, now de-identified and thus cleansed of (most) risk, would be made available to researchers for further analysis.

The second possibility, which is least satisfactory from a research standpoint but better than nothing, is for the agency to conduct its own analyses and give the researcher aggregated data. The public agency will still need individual identifiers and dates of entry into PSH to be able to extract from its own database the records pertaining to the people in the PSH sample or comparison groups for the appropriate pre- and post-PSH periods. The researchers will have to supply these identifiers and dates, after obtaining consent from the people involved. Then agreement will be needed on procedures for resolving record-matching issues. Another point on which agreement is essential is which types of service units to search for; we discuss this issue in the next section. The public agency then does the search, combines the results into pre- and post-placement periods, and conveys the aggregated data to the researchers.

Which format is used will depend in part on who is doing the analyses and in part on the type of informed consent and data releases that your study uses. If you can only get aggregate data, it will limit your analysis options. You will not be able to do regression analyses to help you understand what factors affect any changes you see, but you will still be able to compare mean, median, or total levels of service use before and after entry into PSH.

If the analyst is the public agency "owning" the data, then of course it will have the whole file. However, even a public agency analyzing its own data will need a list of identifiers and dates of PSH placement from the homeless assistance network or PSH program(s) to guide it in selecting records for its clients who are homeless, chronically homeless, or in PSH. It is expected that the informed consent of the clients involved is required before these identifiers can be released to a public agency

### **Selecting the Services to Track**

So far, we have glibly been talking about "service use" as if that is a unitary or simple thing. It is not. Not only do homeless people and PSH tenants use all kinds of services from all kinds of public and private agencies, but the agencies themselves may have complex ways of recording service units. Several factors must be considered in selecting services to track, some of which are simple and some of which can be incredibly complicated. At the simple end is learning from PSH tenants which services have been important to them. At the complex end are coming to understand what databases exist and what they do and do not cover, and how to simplify your quest while still focusing on valuable information.

**Learning from PSH tenants.** A focus group discussion with tenants may explore the question of where they turned for help and which systems they used most often while they were homeless. When tenants share their perspectives about what has changed in their own lives since they moved into supportive housing, this information can help you focus on those systems where changes are most likely to have occurred.

**What changes in service use are policymakers and the public most likely to care about?** For example, if policymakers, budget officials, or the media are particularly focused on emergency room overcrowding, or the costs of psychiatric hospitalizations, then demonstrating reductions in use or costs in those systems may have more impact than a study that examines services across other systems.

**What databases are available and what do they cover?** In general, it will probably be easiest to get good data from systems where funding is allocated based on billing for specific services provided to individuals, and hardest to get good data from service systems or programs where funding is not contingent upon documenting specific services at the individual client level.

A big part of the decision about which services to track will depend on which systems cover the people and services you want to learn about. Because some of the best data will be tied to reimbursement for services, you will need to understand how services are funded and who pays for them. For example, depending on how funding flows in your jurisdiction, a public health agency may maintain one or more hospitals or clinics that serve most homeless and very low-income people, and document services provided through them in a single database with data about all health services the agency provides to people who are uninsured as well as those covered by Medicaid or other health insurance. This would be ideal for your purposes. On the other hand, a number of different private hospitals may serve low-income and homeless people, and each may use a different data system. The state Medicaid program would have data regarding services covered by Medicaid, but the private hospitals may not be reporting individual-level data for services provided to people who are indigent and uninsured. If study participants are enrolled in managed care health services, the public agency may provide funding to the managed care plan based on months of enrollment, and may not maintain data regarding actual service use for these consumers.

Mental health and substance abuse service and payment arrangements can be similarly complex. Most of the agencies that provide mental health or substance abuse services may receive funding from a single state or local government agency that maintains data on the individuals served (including dates of services) across multiple treatment settings, in which case you are in luck. But it is also possible that these services are funded through grants or contracts to agencies that each maintain client data separately. And it is quite possible that some treatment services are provided as part of a state or county-financed system with a central data base while other programs receive funding through federal grants and don't participate in the same data systems. So it will be important to know which services are included and which are left out of the data systems you have access to.

Jails and other criminal justice programs are likely to have data systems that are very different from those in health care systems that were designed for billing purposes. A data system that is designed to track inmates in custody may not contain information about costs of health services or psychiatric care those inmates receive.

**What services or service clusters do you really need to cover?** Tracking *all* of the costs of *all* of the services used before and after PSH placement will be close to impossible. Your task will be a lot simpler if you choose just one or a few service

systems/settings that are most meaningful and measure impacts in those systems. With this approach you can't compare total savings to the costs of PSH but you can demonstrate impacts that are a good sign of the magnitude and direction of changes you would get if you looked at a larger number of systems.

Even in a single setting such as a hospital it may be too difficult to capture all of the services used by each individual. It will probably be easier, yet still reasonable, to get simpler data such as days of hospitalization, and work with the hospital administrators to estimate average costs for a day of services. This approach loses some detail, and may miss changes in the intensity of service use – if for example, hospital visits before PSH were for very complex conditions that required unusually complicated care but post-PSH hospitalizations were simpler. But it will probably be worthwhile to simplify the data request and analysis.

**Which public agencies are interested in helping?** Which public agencies are active partners in the data collection effort, either because they are the ones that want the analysis or because they are willing to work with stakeholders who do? If a jurisdiction is ultimately interested in a fairly extensive economic analysis, analysts may want to begin by demonstrating the ability to manage data responsibly, and show the potential for measuring changes in service use in one system where data managers are enthusiastic partners, before other systems will make additional data available.

Once it is clear that a public agency is going to provide access to its databases, you will have to sit down with agency staff and come to a mutual understanding of how the agency records services. This includes what units they use, how they label the various services a client could receive, how many different places the data could be coming from, and what activities that cost money are likely to be missing. For instance, in a hospital a two-day inpatient stay could involve expenses related to the emergency room, an inpatient ward, and several specialty departments such as radiology, the pharmacy, or the lab. With all that, the *doctors'* charges probably *will not* show up, as doctors in many hospitals bill patients directly. Other charges and costs may be missing as well, such as overhead/hospital administrative costs and the capital costs of the hospital building itself.

**Developing matching procedures.** Another consideration is how you will find the “right” people in public agency databases. It is not uncommon when using client identifiers in different systems for more than one potential match to show up. Mistakes are common when recording identifying information in public databases, and people may give their information differently as well (e.g., using a different form of their first name, or giving their age or SSN incorrectly). Different people may have the same name. After everything electronic has been done to select the right records, human judgment *will* be required to make final decisions about which records to include and which to exclude. Whoever extracts the data for the relevant individuals from public agency databases will need a set of decision rules for resolving which records are the right ones. Those decision rules will depend, in part on your knowledge of your PSH tenants (for instance, if you have no one older than 55 in your PSH, you can eliminate anyone older than 60 as a potential match for your tenants).

Many public agencies have very complicated sets of service codes that they have to use for a variety of purposes. For instance, they could have ten different “therapy” codes, or five different ways of recording treatment for injuries or skin diseases. Given this likely complexity, two things are certain. First, you **must know what you want!!!**—because if you don’t, you surely will not get it. You will have to learn the names of the data fields that contain the data you want, and master the public agency’s service codes enough to be able to discuss which codes you want them to search for and how you want them to combine codes into larger, meaningful, categories for your analysis. Many projects trying to use public agency data founder on the shoals of “the data request.” Most agencies have one or more forms you have to fill out when asking them to extract data for you from their databases. Many people waste a lot of time because they don’t ask the right questions and don’t put the right things on the data request form. They use up a lot of public agency data managers’ good will with repeated requests, or else don’t get the data they need.

Further, these data must be recorded in a consistent manner over the years before and after study participants moved into PSH. This alone can be a formidable challenge since information systems and billing/service categories in public agencies change far too often, and in ways that frequently result in non-comparable data from year to year. You will need to talk with people in each public agency you want to include in your analysis, to get the full picture of how service units have been defined and recorded during all the years your study wants to cover.

### **Deciding Whether to Include Emergency Shelter/Transitional Housing Nights**

The final type of service for which one might want to know use levels is shelter and transitional housing days/nights of use. If you are going to include pre- and post-placement shelter/housing costs in your analysis, you will need to be able to tell how many nights each person has spent in shelter/housing. You could guess at this, or try various ways to get the person to give you a rough idea, but most communities will not be able to get an accurate read on shelter use at the individual level, unduplicated across agencies and over time. Community-wide HMIS data may help with this data vacuum in the future, but will not be able to supply information on service use from before the HMIS was implemented. The NY/NY analysis did include these service units, and they are a significant part of the service mix in that study. However, in most communities shelter and transitional housing programs are not public sector agencies nor do the communities have a homeless management information system or other source of data that tracks service use at the individual level across programs and over time (New York City, Philadelphia, and a few others are the exceptions). Therefore getting data on the number of nights your PSH tenants used these services before and after moving into PSH may be impossible.

### **Checklist for Feasibility of Conducting Analyses of Service Use Before and After PSH**

These are the key things you will need to determine before you can tell whether it will be feasible for you to conduct an analysis of changes in crisis/public service use for PSH tenants, and to determine the scope of your analysis:

- Which public agencies have usable data?
  - Are the data available electronically?
  - Are the data at the individual level, do they include service units and dates of services, and do they have identifiers you can use to match PSH tenants with units and dates of service delivery?
  - How comprehensive are the available service data? Do they include all similar services (e.g. mental health services) provided by all public and private agencies in the jurisdiction, or only those services directly provided by a single agency – or under contract with that agency? Which services / providers are included / excluded? Has inclusion changed during the period of time you are looking at?<sup>10</sup>
  - Is the same information available for all individuals who have received service or only for some (e.g. only for those covered by Medicaid or other 3<sup>rd</sup> party payer, or a specific program)?
  - Are definitions of services units and/or cost and reimbursement practices consistent across relevant years? Is complete or comparable information available for all of the years included in the study period or have there been significant changes in what is included / excluded?
  - What is the usual lag time between service delivery and data entry, and how long do you need to wait before you can be confident that all service data are present for a particular time period?
- For each agency with usable data, what do you have to do to get access to the data?
  - Informed consent from each PSH tenant?
  - Memorandum of understanding or other formal agreement?
  - Agency runs data itself and gives you:
    - Individual-level data?
    - Aggregate results?
  - Agency allows you to run the data at the individual level:
    - On-site as an official or semi-official staff person?
    - Off-site?

---

<sup>10</sup> If some service providers are excluded, you will not have a full picture of service use. This might be a fact of life, but when you report your results you should describe which providers are missing and what the implications might be for understanding your results. Likewise, if there have been significant changes in where homeless people and PSH tenants receive services during the study period, you will only capture the full picture if you include all the places they receive services. Otherwise, you will have to check to see whether a sharp increase or decrease in use of a particular type of service is due to people switching providers rather than to a real change in level of service.

- Which service units do you want to extract?
  - Which data fields?
  - Which service codes? How defined? Consistent across years?
  - How to make the data request to get what you want.
- What about service units (nights of shelter) in emergency shelter and/or transitional housing programs?
- Who will help you understand the service unit categories for each agency?
- Who will help you understand the service databases for each agency and how to query them to get what you want?

**Analysis—Now That You Have Data, What Will You Do With It?**

This section is written as if you, the reader, are going to be doing your own data analysis, using data supplied by one or more public agencies. But don't skip the section just because public agencies are going to be the ones doing the analysis. You will need to be familiar with the issues and challenges describe here in order to be able to discuss your data needs coherently with public agency data analysts, and to be sure that what you get is what you asked for and need.

Once you have gathered the raw data from participating agencies, the next step is to transform them into a form that can readily be analyzed. Unless you are in the relatively rare situation in which all participating PSH residents moved into housing on the same day, you cannot evaluate the impact of PSH by simply adding up all the service use before a set date and comparing it to all service use after that date. Because individual residents move into and out of supportive housing over the course of months or years, each participant will have his own unique pre and post periods, calculated by adding or subtracting a set number of days from his or her individual move-in date. Agency data on service history must then be classified according to these individual pre and post periods. Each individual record of service receipt must be linked with information about when the recipient entered and exited housing and then coded according to the pre/post period in which it falls.

The vast majority of large administrative data systems that track services are encounter-based. In such systems every visit, procedure, or hospital stay has its own separate record that is stored electronically in the system as a row of data. For instance, one visit to the emergency room usually will look like this:

Patient Name	Address	SSN	Pt Number	Hospital Service	Visit Date	Charges	Diagnosis
DOE, JOHN	1440 Jones St.	123-45-6789	254	EMR	7/29/2000	165.22	789.01

For even a relatively small number of participants, the sheer volume of data created by such a system over a four-year period can be daunting. For instance, information on service use for 70 residents of PSH from a countywide behavioral health service provider amounted to nearly 18,000 rows of data. Hospital and emergency room

providers typically provide two to four times this amount. While some of the transformations described below can be done using advanced functions in Excel and Access, we strongly recommend using a more powerful statistical software package such as SAS, SPSS, or STATA.

The key to managing all this information is to use your statistical software program to work methodically through the following five basic steps: 1) clean and code encounter-based service records; 2) attach housing entry dates to service records; 3) categorize each service record by the number of days before or after receipt of housing the service was delivered; 4) aggregate and summarize service encounters by individual resident; and 5) produce summary statistics documenting change over time.

### **Step One: Cleaning and Coding the Data**

The first step involves checking the data for any inconsistencies (referred to as “cleaning” your data) and then coding and collapsing relevant service variables.

- The cleaning process should begin by running frequencies on the variable describing the date of service receipt. The program will produce a long list of dates, in chronological order. You should check the earliest and latest date to confirm that you have received data for the complete period.
- Next, look at the series of dates and see if you can identify any significant gaps in coverage over the years. If the service sector covered in the data is one that residents use infrequently, then month or even year long periods without any service receipt may not indicate any problem. However, if the service system is more heavily used, then these types of gaps should be discussed and resolved with representatives from the relevant agency. They may mean whole months of service use were not captured by the system, or even worse, were never entered. You would not want to interpret such gaps as real drops in service if what they really mean is failures to record service receipt.
- Repeat this process separately for each service type included in the data set.
- Once the data have been cleaned and all problems resolved, the service types included in the dataset must be defined, collapsed, and coded. Usually, there is one variable in the dataset that describes the type of service delivered to the resident. In our example above, the variable is called “HOSPITAL SERVICE.” You should print out frequency runs for this service variable. Data systems usually use a series of three- to eight-letter codes to describe the service type. It is essential that you obtain from the agency a summary sheet defining the meaning of these codes (often referred to as the “code book”). The number of service types included in an administrative data system can range anywhere from 5 or 6 to over 100. When you run the frequencies you can see how many of these service types are used by your population and which are the most prevalent. You must then make a decision about how you will categorize and collapse these service types bearing in mind that the smaller the number of cases in a particular service type, the more difficult it will be to show statistically significant changes over time. Once you’ve made these decisions, create a new variable that reflects your revised groupings and recode all cases into it.



Other things to watch out for are new services or programs entering into the data system you are using, or changes in service availability from providers who are not in the data system. These types of discrepancies should be reviewed with the MIS coordinator at the relevant agency, and be reflected in any report you do using these data.

- **New things in your data system:** Typically, agency administrative data systems record a core group of services when they are installed. Then over time they expand to include a wider variety of service types that in earlier times were either recorded manually or entered into a separate information system. It is crucial to understand when different services were “folded into” the system you are getting data from, otherwise you risk producing findings that do not reflect residents’ actual use but instead simply track the gradual expansion of an agency’s data system. For instance, what looks like a sudden peak in use of psychiatric care in the past year may simply reflect the fact that psychiatric outpatient visits once recorded in a community clinic’s separate data system have recently been folded into the health department’s centralized information system. Check your frequency runs for service types that abruptly begin (or end) months or years after the study’s start date.
- **New or changed levels of service from agencies NOT in your data system.** Suppose you are using data from your community’s public health system, which used to serve everyone because no other providers existed. But some time during your study period another agency (e.g., the VA) began an assertive campaign of outreach and expanded health and psychiatric services to veterans, and many veterans switched from using the public health system to using the VA system. The VA data system is totally separate from the public health data system, and in this example you have no way to access it. You will, however, see a distinct drop in use of certain public health system services—not because fewer such services are actually being used, but because the change has made the continued use invisible through the lens of the public health data system.

### Step Two: Attaching Housing Entry Dates to Service Records

Once you’ve cleaned and coded your data, the next step is to attach housing entry dates to all service records. To complete this step your dataset with housing tenure information must use the same unique identifier as your dataset with service use information. This is essential because you are going to rely on this variable to link the two datasets and to import the housing entry dates into the service use data. Using the “MATCH” or “MERGE” function in your statistical analysis software, attach each resident’s move-in date to every record of service received by that individual. Each row of data should now have an additional column holding the variable “MOVE IN DATE.” If your software allows you to view the data, your row should now look like this:

Patient Name	Address	SSN	Pt Number	Hospital Service	Visit Date	Charges	Diagnosis	Move-in Date
DOE, JOHN	1440 Jones St.	123-45-6789	254	EMR	7/16/2000	165.22	789.01	03/30/2000

### Step Three: Labeling Service Use Records as Pre or Post Placement in Housing

The next step involves categorizing services by whether they occurred before or after an individual's receipt of housing. In order to do this, a new variable must be created that measures the time between the date the resident received a particular service and the date the resident moved into supportive housing. To create this new variable, which we'll call TIME, take the service receipt date variable and subtract the move-in date variable. The resulting variable represents the number of days before or after move-in that the service episode occurred. The values for the "pre-housing" service episodes will all be negative numbers while the post period values will all be positive numbers. Based on these values, the service episodes can each be individually categorized into the appropriate pre and post period groups. For instance, all services with TIME values between -1 and -365 would be recoded into a new variable, which we'll call PREPOST, as one year prior. Any time interval can be chosen, including months or quarters. Your data should now look like this (we show options for one year, one quarter, and one-month time periods:

Patient Name	Address	SSN	Hospital Service	Visit Date	Move-in Date	Time	Prepost
DOE, JOHN	1440 Jones St.	123-45-6789	EMR	7/16/2000	03/30/2000	108	1 Year Post/ 2 <sup>nd</sup> Quarter Post/ 4 <sup>th</sup> Month Post

At this point, you should be able to use your data to do some very basic runs comparing total and average use of services before and after receipt of housing. First, return to the service use variable created in Step One. Use your software programs SELECT function to limit the analysis to only those cases that fall into one of your categories. Now do a frequency run of the PREPOST variable. The resulting output will describe the total number of service episodes in that particular category consumed in each pre/post period for the entire sample across the entire study period. Re-run this analysis for each of the service type categories. In addition to providing a description of the pattern of total use, patterns of average use can also be computed by dividing the totals in each time period by the total number of residents.

### Step Four: Aggregating and Summarizing Data by Individual Resident

For some purposes the very basic summary statistics described in Step Three may suffice. However, for other purposes you will need to perform more complex analyses, including calculating the probability of service use across time periods, exploring difference in use by demographic or diagnostic characteristics of residents, and performing advanced statistical tests. In order to do this, you will have to perform a final data transformation.

Up until now you have been working with a dataset organized around service visits (i.e., each row is a service unit). To perform more complex analyses the data must be re-organized around individual study participants. This involves collapsing the thousands of rows of service encounters into a row of data for each resident that represents his or her entire service history plus PSH experience. Service encounters must be aggregated

and transformed into a series of summary variables describing the number of visits a particular resident made in a particular pre/post period. Then, using unique identifiers, these summary variables must be linked with demographic and diagnostic variables to create a comprehensive case record for each resident.

Creating the service use summary variables is the most challenging aspect of this final step in the data transformation process. Most statistical software programs have an “aggregation” function that will allow you to collapse multiple cases into a summary variable. The first step is to create a new variable for each service category by pre-post period combination. For instance, separate variables should be created for “ER visit two years prior,” “ER visit one year prior,” “ER visit one year post,” and “ER visit two years post.” Each of these variables will have only two possible values, the number 1 if the service delivered was an ER visit and the visit took place in the stated time period and the number 0 if the service delivered did not meet these specifications. These so-called “dummy” variables can be created through using the “if/then” syntax available in all statistical software programs. The code would go something like this: IF (SERVICE = ER) and (PREPOST=PRE1YEAR) then ER1YEARPRE =1; ELSE=0. The specific syntax will depend on the program you are using. The resulting row should look like this:

Patient Name	Address	SSN	Hospital Service	Visit Date	Move-in Date	Time	Prepost	ER visits 2 years prior	ER visits 1 year prior	ER visits 1 year post	ER visits 2 years post
DOE, JOHN	1440 Jones St.	123-45-6789	EMR	7/16/2000	03/30/2000	108	1 Year Post	0	0	1	0

Once you have created the dummy variables, you can then use the SUMMARY or AGGREGATE functions to collapse the multiple rows of data for each resident into one summary case. You will want to instruct the system to aggregate the data by each resident’s unique identifier. When the program groups and collapses records by, say, social security number, it creates summary variables based on the values of the original variables for all the cases in that group. Summary variables can be based on a count of the cases, on the first or last value, on the highest or lowest value, or on the sum or average of values. In our case, we want to take each dummy variable representing a certain service used during a certain time period, for instance the variable “ER VISITS TWO YEARS PRIOR,” and create a summary variable that totals the values across all cases with similar social security numbers. We can then name this variable “SUM OF ER VISITS TWO YEARS PRIOR.” The result is a new dataset made up of variables that summarize the total number of service type encounters an individual resident received during a particular pre or post period. Finally, you should use the MATCH or MERGE function to import demographic and diagnostic variables into this new dataset.

The total number of observations in your final dataset should equal the total number of participating residents and the final case record should look something like this:

Resident Name	SSN	Sum of ER visits 2 years prior	Sum of ER Visits 1 year prior	Sum of ER Visits 1 year Post	Sum of ER Visits 2 years Post	Race	Gender	Age	Substance Abuse	Mental Illness	Move-in Date
DOE, JOHN	123-45-6789	10	13	7	5	W	M	48	1	1	03/30/2000

## Step Five: Produce Summary Statistics Documenting Change Over Time

Using the fully transformed database, you can now produce probabilities, means, medians, and distributions for all service use variables across time. Change between pre and post periods can be assessed using your software's Paired-Samples T Test procedure. The Paired-Samples T Test procedure computes the differences between values of the two variables for each case and tests whether the average differs from 0. For instance, you could use the procedure to assess the difference between the mean value of the variable "ER VISITS ONE YEAR POST" and the variable "ER VISITS ONE YEAR PRIOR." If this difference in means is significantly different than zero, then this provides some evidence that supportive housing may impact levels of service use. You can further support this inference by assessing whether there are any significant changes between successive pre-housing periods. For example, you could assess the difference between the mean value of the variable "SUM OF ER VISITS TWO YEARS PRIOR" and the variable "SUM OF ER VISITS ONE YEAR PRIOR." If you find no significant change between the pre-housing periods, then this lends support to your hypothesis that permanent supportive housing is an important factor in facilitating the drop in service use.

You can also easily examine the difference in the degree of change in service use by residents' race, gender, family composition, homeless history, age, income source, or disability. In order to do this type of analysis, you must create a new variable, formed by taking the post period service use variable and subtracting the pre-period service use variable. For instance, you could create a new variable called CHANGE that equals SUM OF ER VISITS ONE YEAR POST minus SUM OF ER VISITS ONE YEAR PRIOR. This change variable will be a negative number if use decreased after entry into supportive housing and a positive number if use increased after entry into housing. If there was no change, the value of this variable will be zero. You can then use your software's Independent-Samples T Test procedure to assess differences in the average level of change in service use between different demographic groups. Regression analyses can also be used to examine the independent impact of resident characteristics on changing levels of use.

You may also be interested in exploring changes in the probability of use of a particular service across time periods. To do this you will have to create a new variable with values of either one or zero, indicating whether the resident did or did not use a particular service during a particular time period. The easiest way to do this is to recode the variable representing an individual's sum of service visits during a certain time period into a new binomial variable. We will refer to this new variable as PROB VISIT. All residents with one or more service encounters during a particular time period are assigned the value one for PROB VISIT while all residents with no service encounters are assigned the value zero. The mean value of PROB VISIT equals the probability of use of this service type during the relevant time period. Changes in probability across time periods can be assessed using McNemar's test.

While exploring change over time is the most important component of your evaluation, you may also want to look at levels of pre-housing service use in order to assess whether your PSH programs are effectively **capturing high users of public services**. We suggest categorizing your residents into at least two groups representing high and

low users. You will have to create criteria for defining what level of use qualifies an individual as a high user. For some services, such as emergency department visits, there are standard measures in the literature defining high use (greater than 3 visits a year). For other service systems you may want to ask the agency analyst for description of the level of use among their top quintile of users. The last option is to just label the top 10 percent of users among your sample as the high users. You should then compare their demographic characteristics, including referral source, time homeless, as well as their rate of change in service use (CHANGE variable) over time. You should also assess the your PSH programs' rate of retention of high users versus low users. This data can then be used to help your programs to effectively target outreach to those groups with the highest need.

### **COST – How much less are public agencies spending for people living in PSH, compared to public cost in the year(s) before PSH?**

Once you have service use data, you may choose to take the next step and try to identify the cost of a unit of each type of service. With unit cost data and service use data, it is pretty simple to multiply and get final cost figures for before and after people move into PSH. These data will tell you whether there have been changes (hopefully reductions) in expenditures for the types of services you have included in your service use analysis. As with the service use data, if you do not have a comparison or control group you will not be able to attribute all of the cost changes to PSH placement, but your data may still convince many policy makers. Many challenges exist in getting good unit cost data, as discussed below.

### **Obtaining Cost Data**

Some information about costs may seem relatively easy to get; but appearances may be deceiving. Health agencies should be able to put a unit price on a hospital day, an emergency room visit, and an outpatient appointment. In all likelihood they can put a range of prices on these units, depending on their length, complexity, and the nature of the illness. They have to do this to get reimbursed from Medicaid and any other insurers. Most public agencies that can produce the service data you will need for this analysis—by service units for different types of care—should also be able to include data fields displaying the cost of the service and possibly also who paid it. Corrections agencies commonly calculate per day and per year costs of keeping a person in jail or prison. You can use their cost calculations if you are including prison and jail days in your analysis.

However..... As usual, the devil is in the details. You will need to probe at each agency until you have at least a basic understanding of what different agencies (and their data systems) mean when they attach a “cost” to a unit of service, and what components of cost are included in the calculations of these unit prices. If the components vary greatly, you will need to determine whether you can “unpack” the components somehow, and if not, how you are going to handle cross-agency comparisons to get “apples to apples” comparisons of pre- and post-PSH costs. As you are likely to be able to unpack the PSH costs better than public agencies will be able to unpack theirs, the solution may be to leave out of your PSH cost estimates the components that public agencies tend to

leave out of their unit cost estimates. For instance, corrections agencies will probably be able to give you a per diem cost, but be sure you understand what they do with their capital costs—usually they won't be included. They may also exclude some activities that are done in the jail or prison under contract to outside agencies, such as medical care, mental health or substance abuse treatment, or even employment services.

It is important to understand clearly the information that is being provided and any limitations that may impact how you can interpret it. For example, actual costs may differ from approved budgets (because of savings, over-spending, or other unanticipated changes) or from amounts charged to other funding sources such as private insurance or Medicaid. Further, reducing use of services in some settings may not automatically translate into financial savings to the agencies that provide or fund those services. A large part of the average cost for services in any setting is fixed cost that does not disappear when service use changes at the margin. But you can make the case that (1) scarce resources or limited capacity in existing systems can be used more effectively for others who need it and/or (2) over time, as a significant number of people move into PSH, it will be easier to achieve savings that can be re-directed.

The per diem rates that are easy for agencies to give are often the amounts they can bill to third-party payors such as Medicaid or insurance companies. Actual costs may differ substantially from these rates, being either higher or lower. Getting closer to actual costs in public agencies is often an arduous task, but one you have to try to accomplish if you are seeking to compare the costs of PSH to the costs of other services. You can be less rigorous if you want to use cost information to demonstrate the overall direction and magnitude of changes attributable to PSH – especially if you will describe those changes in terms of percentage reductions rather than actual dollar amounts. At a minimum, you need to know what types of costs they include or exclude when they calculate these rates.

Sometimes all you will be able to get will be Medicaid reimbursement rates—that is, you won't be able to “unpack” the full cost of a service. As such rates usually do not cover the real cost of providing services, it may be important to determine how much public money is being used to cover the uncompensated portion, or at least get an approximate percentage to apply across the board. For each provider/system it will be necessary to get some information about the ratio of costs to charges reflected in the data system. Given these complexities, it can be very helpful to have the agency budget staff engaged as advisors at some time during the study process, whether for a one-time consultation or as part of an evaluation advisory group.

Then there are the situations in which the public agency does not deliver services itself, but contracts with other entities to do the actual services. For these situations in which funding is provided through an overall contract budget for the program and reimbursement is not provided on a fee for service or charge per unit basis, it will be necessary to calculate average costs per day or unit of service. These contractual arrangements often occur with mental health and substance abuse services, and may also be true for a residential treatment program or jail medical or psychiatric care. The average costs obtainable from contracted service providers may not reflect the level of intensity of services that some participants will have used while homeless, but it is often the only feasible approach for including their data. The same situation may occur even

in public agencies, or anywhere that costs are structured at the program level rather than at the individual client level.

The final type of agency for which one might want to calculate costs are shelter and housing programs. If you were able to include nights in emergency shelter and/or transitional housing programs in your service use analysis, you will be in a position to think about whether you want to include their costs in your cost analysis. As a variation on the NY/NY analysis, a jurisdiction could do without shelter costs, although that is really leaving out an important element of service costs and potential cost savings (11 percent of costs in the NY/NY analysis).

A city such as New York, which pays on a per-night basis for 80 percent of the emergency shelter available in the city through contracts with private providers, has a very good handle on shelter costs. However, most cities do not have this sort of central public payment mechanism. In cities without it, analysts would have to gather program budget data from a high proportion of providers offering emergency shelter, as the NY/NY researchers had to do for PSH. Getting estimates of their per/night or total costs is likely to be more complicated than getting public agency costs. The approach is the same, but you will probably have to work with a larger number of programs to get all the cost information you need.

### **SAVINGS – To what extent does the money saved from reduced use of public crisis services help to pay the cost of PSH? (Is PSH cost-effective?)**

Savings is the final component of the NY/NY analysis. You don't have to calculate savings, of course. Policy makers can be very persuaded by data showing PSH impact on service use and costs, even if only for a few systems of care and without assessing the costs of PSH and comparing them to determine whether total savings outweigh the costs of PSH. Your choice of whether to include a savings calculation may depend on whose perspective one takes. County policy makers trying to decide whether to commit locally controlled service dollars may be concerned about PSH's impact on use and costs of county health and criminal justice services—as long as the savings outweigh the county's costs for services in PSH they will be happy to get federal funding for capital and operating costs of PSH.

If you are going to assess public savings from PSH, you will need the cost data from your analysis of service use and costs. You will need to compare these to the costs of the PSH itself, including the services that are attached to it. Savings are the difference between the cost of PSH and the reductions in costs to the other agencies that formerly handled the crisis services for chronically homeless people (to the extent these reductions are attributable to PSH and not to other factors). As with the unit cost data for service use, there are many issues to consider in getting good cost data for PSH. There are also issues in the comparability of what is included in the unit cost estimates for PSH versus the service units delivered by public agencies. We discuss these issues below.

If a provider offers PSH in addition to emergency shelter and other types of shelter and housing, the PSH part of the budget would have to be separated from the shelter part. Further, some parts of the expense of operating these shelter and housing programs

might not show up in the provider's regular budget. These include expenses such as capital costs for the building, debt service, operating costs, staff costs for supportive services and supervision, staff supports and resources from the central office of an umbrella agency,<sup>11</sup> and the costs of any supportive services offered on-site by other agencies and therefore not appearing in the provider's own budget. All of this entails some pretty serious investments in examining the budgets and off-budget sources of support for PSH, on a provider-by-provider basis.

Finally, many PSH providers have private as well as public sources of support. As you are after *public* costs, you cannot just take an entire PSH program budget and divide by the number of tenants to get per-tenant costs, because doing so will factor in some proportion of private resources as well. Analysts should consider whether they should simply remove these resources before calculating PSH costs, or report them separately to document the extent to which public funds are part of public-private partnerships to provide this type of housing.

### **Checklist for Feasibility of Conducting Analyses of Service Costs Before and After PSH and Savings from PSH**

These are the key things you will need to determine before you can tell whether it will be feasible for you to conduct an analysis of changes in crisis/public service cost for PSH tenants, and whether you will be able to compare them to PSH costs for the same time periods. (The checklist for the feasibility of assessing changes in service use appears above in the section on "USE.")

Cost queries:

- What cost data can you get from the agencies (public and/or private) that can supply service use data?
  - For which services?
  - What cost components are included (these don't have to be broken out separately, but you need to know if they were included in the calculation of a unit cost)?
    - Personnel
      - Employees
      - Other staffing including independent professional practitioners (e.g., doctors or specialists in hospitals) or staff from other agencies (e.g., health services on-site in shelters,

---

<sup>11</sup> For instance, the local headquarters of the Salvation Army, Catholic Charities, and similar umbrella agencies offer significant administrative, training, and other supports to the staff of their various shelters and PSH projects, and may also maintain certain services centrally (e.g., employment readiness, mental health counseling) that shelter residents or PSH tenants may use. Of course the same thing could be true when measuring the costs of other service systems. The hospital or jail costs used for comparison may or may not include similar support from the large public agencies of which they are a part. This should be something you check out in your assessment.



or other out-stationed staff who are part of the array of services)

- Contracted Services
- Facilities
  - Operating/maintenance
  - Capital costs
- Procedures and treatments
  - Radiology
  - Lab work
  - Medications
  - Actual treatment or therapy
  - Etc.
- Overhead/administrative costs
- Supplementation from an umbrella agency whose costs are not included in the public agency's service unit cost calculations
- Other
- How are costs organized?
  - Are costs calculated based on units of service? *per visit? or per day?*
  - Are costs aggregated (or aggregatable) into meaningful categories? For example, could all the costs associated with a single emergency room visit (facilities, doctors, procedures, lab work, radiology, etc.) be assembled together by searching on the date of the visit?
  - If organized by contract, so you have to do the total cost of the contract divided by the number of service units reported to calculate an average cost per unit of service within a specified timeframe, can you tell when the services were delivered, to see if they fall within the time frames you are after?
  - Are the same cost components included or excluded across relevant years – or has the agency made significant changes in how costs are calculated?
- Who will help you understand the cost structures for each agency?
- Who will help you understand the cost data system for each agency and how to query it?

Savings queries:

- Can you get PSH cost data?
  - How many programs would need to be involved?
  - What cost components are included in the PSH program budgets (see same list of components as above)?
  - What additional cost components help to augment the program or services offered through the PSH programs (see same list)?
  - How will you handle private resources that contribute to PSH budgets?

## ISSUES, VARIATIONS, AND APPLICATIONS

### **Comparison Group(s) – Pro and Con**

Comparison groups, case-control matching, or true control groups through random assignment reduce the likelihood that you will misinterpret the results from a treatment group alone. This is true because some changes in the desired direction may have taken place in the absence of the treatment. The only good way to examine this possibility is to have a no-treatment group and see what happened to its members.

You can check out why this is important using the published results of the NY/NY analysis. For instance, the group receiving PSH reduced its use of shelter by 163 days, on average, from the pre- to the post-placement period. But regression analyses comparing the people who moved into PSH to the matched controls found that after holding a number of other factors constant, the reduced number of days attributable strictly to NY/NY placement was 115, or only 70 percent of the raw difference for the group that received PSH. The difference is attributable (in part) to the fact that the comparison group used somewhat fewer shelter days also. Without the comparison group, researchers would have overestimated the effects of PSH and the costs it helped to avoid. As a noted evaluation expert, Peter Rossi, was fond of saying, “the less rigorous the methods, the better the [apparent] results.” Researchers conducting analyses of service use and cost changes and savings accrued must be careful not to oversell the results they get, especially when in response to real-world difficulties they have compromised some important requirements of good evaluation design.

The NY/NY analysis used public records of shelter use, matched to records of other public agencies, to select comparison samples. Most jurisdictions will not have these shelter use records, so they will not be able to follow the NY/NY procedures. However, they may have some alternative sources of comparison groups, depending on how they allocate their PSH. If there are large wait lists for your PSH and tenants are selected by lottery so each person on the list has as good a chance as any other to get PSH, you have *de facto* random assignment. The people remaining on the wait list are likely to be very similar to those who received PSH, and could serve as a comparison group. This approach will not work, however, if PSH providers select waitlisted people according to some nonrandom characteristics, as that would introduce selection bias and the PSH and comparison groups would no longer be reasonably similar.

Much of the expense of the NY/NY analysis lay in creating and then using the matched pairs. Most jurisdictions will not be able to afford this expense. We discuss some alternatives to creating comparison groups in the next section, on compromises, along with the limits of interpretation that may go along with them. Even with limits and caveats, at least one of these alternatives may be feasible for jurisdictions to use in creating momentum for policy changes.

### **Compromises, or, Doing “Just Enough”**

Most jurisdictions will not be able to invest the time or money to replicate the NY/NY analysis. Nor will most jurisdictions need to do so to make the case that investing in

PSH is a cost-effective as well as a humane thing to do. What most jurisdictions will face are the questions of what they *have the ability* to do, and what they *need* to do to convince key decision makers that PSH is a good idea.

The ability question focuses on which data sets you can readily access, whether you have enough PSH that has been around long enough to justify any analysis, whether you can tell who is chronically homeless, and so on. Answering the ability question means examining your own public agencies and data sources to see what you can actually do.

The necessity question focuses on political will, and what it will take to move it in the direction you want. Because key decision makers may respond differently to different types of information, every jurisdiction will probably need to put together a somewhat different set of information to carry the day. As you will certainly need to pick and choose which elements of the NY/NY analysis you will pursue, it is wise to make them the ones that have the best selling power—always assuming you have the data to do those particular analyses.

The choices you may have to make include:

- **Using fewer databases, at least to start.** The NY/NY analysis used databases from eight public agencies and every level of government. The Connecticut study used one database (state Medicaid agency data); the initial HHISN analyses used two databases from a county public health agency, although ultimately it included the databases of five public agencies. Obviously a study can choose from among available databases which ones it will use. The decision will rest in part on three factors—which ones have the data you need, which ones public agencies will let you use, and which ones will supply the most convincing evidence.
- **Covering more or fewer payment sources.** The Connecticut study included only Medicaid-reimbursed services, and had to cut its sample quite a bit to do that. The NY/NY and HHISN studies used city and county public agency databases that included care not covered by Medicaid or other insurance, which was paid for by local government revenues. By doing this they picked up more care, and care that people were more likely to use while homeless. Because Connecticut is a state-administered state, counties are essentially irrelevant, and cities play a fairly reduced role, (plus having an intervention that operated in six cities), researchers there probably did not have the option to examine databases revealing uncompensated care. Each jurisdiction will have to fit what it wants to do to the realities of how care is administered locally.
- **Working without a comparison group.** A good deal of the NY/NY analysis resources went into creating matched pairs to serve as controls. The Connecticut and HHISN studies did not do this, and in most communities researchers probably will not have the resources to do it either. A number of options exist that can help reduce potential misinterpretation of results when you have only an intervention group. We list four:
  - *Divide pre and post periods* into annual or 6-month periods to create a time series. If you then see distinct patterns that are persistent over time periods

before PSH placement, and you see those patterns shift in consistent ways for the time periods after placement, your case for PSH impact is stronger.

- *Use regression analysis with people as their own controls* pre and post PSH placement. The NY/NY analysis ran regressions with PSH placement as a key predictor variable and change in service use (e.g., days of shelter use before minus days of shelter use after placement) as the dependent variable, comparing those who got PSH with those who did not. The HHISN analysis ran essentially the same regressions (change was the dependent variable) but without a predictor of PSH placement, since everyone in the analysis was in PSH. See suggestions for analysis, above, for a bit more detail, and published reports of analyses referred to throughout this guide if you want a lot more detail.
- *Adjust your regression results, or interpret within broad confidence bands.* One can use the published results of the NY/NY analysis to show the difference in PSH impact from regressions compared to just looking at the results for the group that got PSH. With one exception, regression results are different, and not always in the same direction. Two analyses show smaller decreases in post-placement service use, two show larger decreases, and one (outpatient care) shows a smaller increase. The differences are not trivial—for one of the “smaller decrease” figures, the regression results show only 30 percent of the service decrease that one would infer from the raw data. The differences may arise from a number of causes. Some stem from differences in the personal characteristics and behaviors of the treatment and control samples, regardless of how much care was taken in matching. Some may arise because service systems change in ways that impact all consumers, whether or not they are PSH tenants. For example, new hospital cost containment procedures or simple budget reductions may limit access to care or shorten hospital stays for all consumers during the study period. The implication—use broad confidence intervals, calculate savings at both ends as well as the average, and interpret with caution.
- *Covering shorter time periods before and after placement.* We do not recommend this, as it greatly reduces the confidence you can have that you are observing a *real pattern* of behavior change. Better to wait until people have been in PSH two years and the public agency data are in hand (remember that it takes one to two years after treatment for Medicaid databases to become reasonably complete).
- *Using Different Cohorts of PSH Tenants.* One might be able to use those PSH tenants who entered housing in later years as a control group for comparisons with tenants who entered in early years (assuming that they were homeless during the period of time when the earlier tenants were living in PSH). Or, in jurisdictions that select PSH tenants by lottery from a wait list and the screening criteria for housing do not produce bias, those who remain on the wait list are likely to be comparable to those who become PSH tenants. One might be able to use them as a comparison group and thus get a comparison more easily than by using the matching process of the NYNY analysis.

## **Analyses of Special Population Subgroups**

Costs avoided or saved will be greatest for people who have been homeless the longest, and for people who have the most disabling conditions requiring health, mental health, and other types of care. If your PSH interventions are targeted to people who have been homeless for many years and have several serious disabling conditions, you are most likely to identify cost savings from PSH.

However, most PSH projects include tenants with a range of disability levels, and a variety of homeless experiences. Given what we know and considering the policy interests in PSH that serves a more diverse population, a community might consider it valuable to identify different subgroups in its PSH and look at service use, cost, and savings separately for each subgroup as well as for the general mix of tenants. Some characteristics that might help define meaningful subgroups are

- Disability (e.g., severe and persistent mental illness, or HIV/AIDS, or chronic substance abuse, or chronic physical illnesses),
- Homeless history (e.g., a very lengthy spell—say five years or more, repeated spells covering a long time such as five years or more, and, given the recent increase in family PSH,
- Household structure—adults by themselves, and adults with at least one child with them.

There may be not only differences in total costs/savings but the costs/savings may come in different systems. For example for very seriously mentally ill people who were not chronically homeless, the big savings may come from reduced use of residential and institutional mental health services, while for chronically homeless people the savings may come from emergency room, jail, detoxification, psychiatric emergency services, and, for families, foster care. However, one would need a fairly large sample of PSH tenants to have subgroups that are large enough to produce meaningful information that can be attributed to subgroup characteristics.

## **Applications of Use/Cost/Savings Results**

If your analyses cover only a small number of the PSH-eligible people in your jurisdiction (i.e., if you only do the analyses for one or two programs, as the HHISN study did), you may want to project your results out to identify potential savings for the entire chronically homeless population. To do this you will need data on length of homelessness, patterns of homelessness and movement in and out of institutions, and disabling conditions for most or all of the people who are homeless during a year, two years, or several years' time in your jurisdiction.

The best way to get this information is from a tracking database such as an HMIS, preferably one that includes both outreach and shelter contacts. The tracking database should be able to give you an idea of how the people you tracked compare to all chronically homeless people in the system. It should also be able to give you a sense of how many disabled people who might be PSH candidates—perhaps with a history of homelessness or perhaps becoming homeless for the first time—have entered

homelessness during a year or some other time period. If you don't have data from a tracking database, you will have a very hard time making projections, as point-in-time information does not help very much for chronically homeless people. If you do have relevant data from an HMIS, you can use it to calculate how many people will need PSH. Ultimately, the number of units of PSH you commit to build will depend on this type of projection.

## APPENDIX: DATA PRIVACY ISSUES

Data privacy issues include legitimate uses of personal data in public administrative databases, who may use the data and whose permissions are required, and data security. This appendix cannot cover any of these topics in depth, but we did want to give some indication of the issues and some ways that different researchers have handled them.

### **Legitimate Uses of Public Administrative Databases**

Public administrators may analyze the databases maintained by their own agencies for administrative and management purposes. The types of economic analyses covered in this guide can usually be interpreted as falling within one or more of these purposes, such as assessing cost-effectiveness of services, improving service delivery and targeting, or simple accountability—tracking how money is being spent.

Thus if *all* of the databases one needs to search for an economic analysis are public agency databases, most state laws would allow the jurisdiction to conduct the analysis—including matching across the databases of different agencies—without having to obtain informed consent directly from PSH tenants. But this situation is rare, as it would mean that all users of homeless assistance programs, including PSH tenants, are tracked as part of a public agency database. For the NY/NY analysis, this situation actually *did* prevail for the people receiving NY/NY housing, and people selected as the matched controls were already in New York City’s emergency shelter tracking database. So the researchers never had to go directly to NY/NY tenants or the matched sample to obtain their informed consent.

Even without this ideal situation (from a research perspective), and even with newly formulated and more extensive data privacy regulations from HIPAA, there are many circumstances under which it is legitimate for researchers to access and analyze public administrative databases, including those covering health, mental health, and other sensitive topics. Exhibit A-1 describes some of these (text taken directly from DHHS Office of Human Subjects materials).

### **Who May Use the Data and Whose Permissions Are Required**

Researchers, including people employed by public agencies and others who have appropriate permissions, may use public administrative data for legitimate purposes. To conduct the analyses described in this guide, researchers in most jurisdictions will need to get informed consent from PSH tenants to search administrative databases and get the identifying information that will allow them to select the right records from those databases. Even if the researchers are public agency employees, they will not be able to identify tenants residing in privately run PSH programs and get their identifying information without the tenants’ direct consent.

Researchers should check with each agency whose databases are required for the analysis, to learn about agency-specific requirements for data release. There may be specific forms that each agency requires, or specific wording that a release must contain. We include a sample release form (Exhibit A-2) from a recent San Francisco



project, which represents the best information currently available to the author of what is required. As it took us about two weeks of staff time to be as sure as anyone can be that we have all the right language, at least for California, readers may as well benefit from that work. Researchers should take this form, revised for local use, to each agency that you want to be involved in the analysis, and get agreement that the form will satisfy that agency's requirements for data release.

### **Data Security**

Every jurisdiction will have to develop procedures to assure data security. These include procedures to assure that no unauthorized person has access to the data; that once data extracts are made from administrative records and a research database constructed, all identifiers are removed; that data, including de-identified data, are stored in a secure manner; and various provisions for the ultimate destruction of the data once analyses are complete. Every university and research organization, and many public agencies, has Institutional Review Boards that oversee data security and research compliance to these procedures. Check around locally to find one that will provide specific procedures to follow.

## Exhibit A-1: Special Provisions Related to Protected Health Information

The HIPAA Privacy Rule applies to the release of Protected Health Information (PHI)<sup>1</sup> by covered entities (health plans, health care clearinghouses, and certain health care providers who electronically transmit health information). Researchers should be aware of the Privacy Rule because it establishes the conditions under which covered entities can disclose PHI for research purposes. To gain access for research purposes to PHI, researchers may be required by covered entities to provide supporting documentation which those entities may need to meet their requirements under the Privacy Rule. Under the Privacy Rule, the use or disclosure of PHI for research purposes by a covered entity requires that at least one of the following conditions apply:

1. **Subject authorization for research.** This provision requires written authorization by the subject permitting the covered entity to use and disclose PHI for research purposes. (See Authorization Core Elements at page 12 <http://privacyruleandresearch.nih.gov> )
2. **Review preparatory to research.** Under this provision access is limited to a review of the data at the covered entity's site for specific tasks related to the development of a research project.
3. **Research on decedents.** Research on decedents is not subject to the Privacy Rule; however, documentation of death may be required by the covered entity.
4. **De-identified data set.** The Privacy Rule will not apply to data sets that exclude **all** of the 18 elements mentioned in footnote 1 below *or* to data sets for which an expert in statistical and scientific methods determines and documents that the methods used to de-identify the data set results in a very small risk that an individual could be identified.
5. **Limited data set and data use agreement.** The Privacy Rule permits the use of a "limited" data set for research purposes. That data set excludes the categories of direct identifiers in footnote 1 except the data may include city; state; ZIP Code; elements of date; and other numbers, characteristics, or codes not listed as direct identifiers. A data use agreement between the covered entity and the researcher's recipient organization will document that a "limited" data set may be used or disclosed only for purposes of research, public health, or health care operations and describe how the data will be protected
6. **IRB approval of waiver of HIPAA authorization.** The covered entity may grant (or require of the recipient researcher's organization) an IRB-approved waiver or alteration of the authorization requirement. The IRB must base its decision on the presence of the following elements:
  - (a) The use or disclosure of the information involves no more than minimal risk to individual privacy based on an adequate plan to protect information identifiers from improper use and disclosure; an adequate plan to destroy identifiers at the earliest opportunity; and adequate assurances that the information will not be reused or disclosed to any other person or entity, except as required by law for authorized oversight of the research.
  - (b) The research could not practicably be conducted without the waiver or alteration.
  - (c) The research could not practicably be conducted without access to and use of the information.

<sup>1</sup> Protected Health Information under HIPAA is health data that contain one or more of the following identifiers: (1) names, (2) all geographic subdivisions smaller than a State – see special rules for zip codes, (3) telephone numbers, (4) fax numbers, (5) e-mail addresses, (6) Social Security numbers, (7) medical record numbers, (8) health plan beneficiary numbers, (9) account numbers, (10) all elements of dates - except year – relating to any individual, (11) certificate/license numbers, (12) vehicle identifiers, (13) device identifiers, (14) URLs, (15) IP address numbers, (16) biometric identifiers, (17) full face photographic images, (18) any other unique identifying numbers, characteristics, or codes.

---

## Exhibit A-2: Informed Consent B – FPSHI Evaluation

---

### I. NATURE, PURPOSE AND DESCRIPTION OF STUDY

As part of the study we are conducting for the Schwab Foundation, we would like to obtain your permission to access information about different types of services that you and your children may have used at the following agencies: San Francisco City and County Department of Public Health, Department of Human Services, Veteran's Administration, Forensic Services, and the Sheriff's Office. The research team will only access information about the types of services your family has received and their cost. To help find your records, I will request your social security number and date of birth, and that of your children. You will be asked to give us any other names or aliases you used when you received services in the past. I will also review a list of all the agencies and the type of information they collect. Your participation in this research is voluntary. You are free to choose not to participate in this study. If you choose not to participate, it will have no effect on you, your housing, or the services you can use. If you do choose to participate, you have the right to revoke this consent at any time, except to the extent that action has already been taken based on it.

### II. BENEFITS

By granting permission to obtain information related to your use of services, you can help organizations decide how best to provide supportive housing services to families like yours.

### III. RISKS, PRIVACY AND CONFIDENTIALITY

It is possible that you may feel uncomfortable providing researchers with access to records of the services you used. In addition, there is a risk of loss of confidentiality. However, your records will be assigned an ID number, and whenever possible this number will be used on records instead of your name or social security number. All confidential information will be stored in a locked cabinet that only the research team can access. Confidential information in electronic format (such as computer hard drives, magnetic tapes or disks) will be stored so that only the research team can access it. None of the information about you will be shared with building management, service providers, or other tenants. No names will be used in any reports or publications that result from this study.

### IV. MORE INFORMATION

If you would like more information about the study, you can call Harder+Company Community Research at (415) 522-5400 and ask for Michelle Magee or Clare Nolan. Do you have any questions for me right now? Would you like to participate in this study? If you agree, I will give you a copy of this Consent form to keep so that you can be sure about what you have agreed to.

**My rights as a study participant have been explained to me and my initials below indicate that I understand my rights. I have been given a copy of this form to keep.**

Signature: \_\_\_\_\_ Date: \_\_\_\_\_

Printed Name: \_\_\_\_\_

## Authorization for Release of Personal Information for Research Family Permanent Supportive Housing Initiative Evaluation

### 1. Study Information

Study Title: Family Permanent Supportive Housing Initiative

Principal Investigators: Michelle Magee (Harder+Company Community Research, San Francisco, CA) and Martha Burt (The Urban Institute, Washington, DC)

### 2. Purpose of Authorization

The attached Informed Consent form requests your participation in a research study. This Authorization for Release of Personal Information is a required in addition to the Consent form. It does not change any of the information described in the Consent Form. Rather, it is a separate authorization form specific to obtaining personal health and service use information about your family for the research study named above. It provides information about the use and disclosure of this information and it requests permission to disclose and use personal information. If you sign this form, you agree to the release of Personal Information to the research team. Personal Information contains specific information about you and your family, including personally identifiable information such as name, address, phone number and social security number.

**3. Individual Information to be Used or Disclosed.** By signing this document, you will authorize the parties listed in item 5 (below) to provide the principal investigator(s) and members of the research team with access to the following information about you and your family members identified on this form.

- **Housing** – Use of subsidized housing, receipt of housing vouchers
- **Health** – inpatient stays, outpatient visits, emergency room use at San Francisco General Hospital, and use of public health clinics (including Veterans Affairs clinic), health-related case management services, receipt of jail medical services, and Medi-Cal
- **Mental Health** – Use of psychiatric services, psychiatric hospitalization
- **Substance Abuse Services** – use of alcohol and drug treatment services
- **Public Assistance** – CalWorks, General Assistance, and other income supports
- **Child Welfare** – out-of-home placements and kinship care
- **Criminal Justice** – convictions, time served in jail
- **Education** – student attendance, discipline, referrals, grades, health and immunization records

**4. Specific Authorizations.** The following information will not be released unless you specifically authorize its disclosure by initialing the relevant line(s) below:

\_\_\_\_\_ I specifically authorize the release of information pertaining to drug and alcohol abuse diagnosis or treatment. (42 C.F.R. §§2.34 and 2.35)

\_\_\_\_\_ I specifically authorize the release of information pertaining to mental health diagnosis or treatment. (California Welfare and Institutions Code §§5328, *et seq.*)

**5. Parties Who May Disclose Your Individual Information.** The researchers named above and their research staff may obtain individual information about you and your family members identified on this form from the following entities:

- San Francisco City and County Department of Public Health – including Community Health Network, Community Mental Health Services, Community Substance Abuse Services and Forensic Services
- San Francisco Department of Human Services – including Division of Housing and Homeless

- Programs and Family and Children's Services
- San Francisco Veteran's Administration
- San Francisco Sheriff's Office
- San Francisco Unified School District
- California Department of Health Services

**6. Duration of Investigator Access and Use of your Individual Information.** This Authorization will expire at the end of the research study or December 31, 2006, whichever occurs earliest.

**7. Right to Revoke or Cancel Your Authorization.**

You can cancel your Authorization to allow use of your health information at any time by:

- a. Writing to the principal investigator identified in the Informed Consent form, or
- b. Informing a member of the research team that you wish to revoke your authorization. They may ask you to fill out a form.

If you cancel the Authorization, the research team may continue to use information about you and your family members identified on this form that has already been collected. No new information will be collected about you after you cancel the Authorization.

**9. Authorization.**

By signing this Authorization, you agree that 1) you have read and understood this form, 2) you have been given the opportunity to ask questions, and 3) you agree to the release of your and your family's Personal Information as described in this form. If you do not sign this Authorization, the researchers will not have access to your and your family's information. This will have no effect on the housing and services you currently receive. You will receive a signed copy of this Authorization for your records.

Participant Signature	Printed Name
-----------------------	--------------

Social Security Number	Date of Birth
------------------------	---------------

Other Names, Aliases Used (including maiden name)

This consent also applies to the following members of my family for whom I have the legal authority to consent (if applicable):

Printed Name	SSN	Signature Release
Printed Name	SSN	Signature Release
Printed Name	SSN	Signature Release
Printed Name	SSN	Signature Release
Printed Name	SSN	Signature Release
Printed Name	SSN	Signature Release

<i>Interviewer</i>	<i>Site</i>	<i>Date</i>
--------------------	-------------	-------------