

Does One Size Fit All? What We Can and Can't Learn From a Meta-analysis of Housing Models for Persons With Mental Illness

H. Stephen Leff, Ph.D.

Clifton M. Chow, M.A.

Renee Pepin, M.A.

Jeremy Conley, B.Ph.

I. Elaine Allen, Ph.D.

Christopher A. Seaman, B.S.

Objective: Numerous studies have evaluated the impacts of community housing models on outcomes of persons with severe mental illness. The authors conducted a meta-analysis of 44 unique housing alternatives described in 30 studies, which they categorized as residential care and treatment, residential continuum, permanent supported housing, and nonmodel housing. Outcomes examined included housing stability, symptoms, hospitalization, and satisfaction. **Methods:** Outcome scores were converted to effect size measures appropriate to the data. Effect sizes were combined to estimate random effects for housing models, which were then compared. **Results:** All models achieved significantly greater housing stability than nonmodel housing. This effect was greatest for permanent supported housing (effect size=.63, $p<.05$). No differences between housing models were significant. For reduction of psychiatric symptoms, only residential care and treatment differed from nonmodel housing (effect size=.65, $p<.05$). For hospitalization reduction, both residential care and treatment and permanent supported housing differed from nonmodel housing ($p<.05$). Permanent supported housing achieved the highest effect size (.73) for satisfaction and differed from nonmodel housing and residential care and treatment ($p<.001$ and $p<.05$, respectively). **Conclusions:** The meta-analysis provides quantitative evidence that compared with nonmodel housing, housing models contribute to stable housing and other favorable outcomes. The findings also support the theory that different housing models achieve different outcomes for different subgroups. Data were not sufficient to fully answer questions designed to enable program planners and providers to better meet consumers' needs. It is important to answer these questions with research that uses common measures and adheres to scientific conventions. (*Psychiatric Services* 60:473–482, 2009)

Dr. Leff, Mr. Chow, and Ms. Pepin are affiliated with the Human Services Research Institute, 2336 Massachusetts Ave., Cambridge, MA 02140 (e-mail: sleff@hsri.org). Dr. Leff is also with the Department of Psychiatry, Cambridge Health Alliance, Harvard Medical School, Cambridge, Massachusetts. Mr. Chow is also with the Heller School for Social Policy and Management, Brandeis University, Waltham, Massachusetts. Mr. Conley is with the Goldman School of Public Policy, University of California, Berkeley. Dr. Allen is with the Statistics and Entrepreneurship Division, Babson College, Babson Park, Massachusetts. Mr. Seaman is with the Department of Mathematics, City University of New York.

A latent function of state mental hospitals was to provide housing for persons with serious mental illness (1). As state hospital use was reduced in favor of community care, the issue of where people with serious mental illness would reside gained prominence, especially as the phenomenon of homelessness emerged (2). Gradually, a consensus based on humanitarian values and research grew that housing is a keystone service for persons with serious mental illness, one that has direct and indirect effects on physical and mental health (3,4). For example, one study found that meeting the housing needs of persons with serious mental illness was more protective in regard to early mortality from natural and other causes, including suicide, than provision of any other needed service (5).

Influenced by the housing needs of persons with serious mental illness and the plight of homeless persons with mental illness, the mental health field has evolved different “housing-and-service bundles” (6) that we refer to here as housing models. These housing models evolved over time in response first to deinstitutionalization and then to concepts of community support and, finally, to reflect the recovery movement (7).

As housing models have emerged, a growing number of studies have evaluated the impacts of these models on consumer outcomes. This article reports the results of a quantitative syn-

thesis, or meta-analysis, of these housing interventions. For this synthesis we collected and reviewed individual studies and summaries of housing interventions. None of the summaries approximated a full-fledged meta-analysis. Four of the summaries were primarily narrative reviews (6,8–10). One included effect size data but computed only simple averages across studies without weighting by number of participants or adjusting for data variability (11). Newman (6), in particular, focused on the need for housing intervention research with design features that would provide unbiased, high-quality data.

Our goals in this meta-analysis were to describe as precisely as possible what we know about the effectiveness of housing models for persons with serious mental illness; assess the state of evaluation research in this area; and discuss next steps for users of housing services, housing providers, and evaluation researchers.

Before extracting data from the specific investigations, we reviewed studies and summaries to identify housing types, housing outcomes, sociodemographic and clinical characteristics, housing attributes (6), and theoretical perspectives employed in previous studies, using these to guide our data extraction, hypothesis formulation, and meta-analysis.

Background

Housing types

Our review of housing intervention evaluations indicated that the community settings where persons with mental illness reside have been referred to by various names. We propose that these settings can be classified into four groups that we describe below: three unique housing models and non-model housing. The different models emphasize different outcomes and in some cases serve different subpopulations of persons with serious mental illness. The fourth housing type—non-model housing—is for housing not categorizable according to the models above. “Socially complex services” (12) such as housing models can be classified in multiple ways, and any one scheme inevitably has elements of arbitrariness. We offer this scheme as a useful starting point.

Residential care and treatment model housing. The earliest models, developed in response to deinstitutionalization, provided room and board and supervision, and in some cases treatment services, to manage or reduce symptoms related to mental illness and in some cases co-occurring substance abuse. In the literature, these settings are referred to by names such as board-and-care homes, group homes, community residences, halfway houses, and cooperative apartments. For the purposes of this analysis, we classified these housing interventions together as residential care and treatment housing models.

Ideally, some residents in residential care and treatment model housing would progress to housing with less intensive services, but this was not a primary goal of programming in the evolution of this model. Services, typically support services but occasionally treatment, were routinely provided on site by “housing staff,” who sometimes lived in the program’s housing. House rules were often made and implemented solely by housing staff, although residents might play a role in establishing rules. Residential care and treatment housing programs, with the exception of board-and-care homes, are often “high demand” environments, where privileges and rules of conduct are well defined, participation in services (some delivered in the housing) is required, and abstinence from alcohol or other drugs is a prerequisite (13). As such, they tend to select residents who need continued treatment but are “ready” to follow house rules, and the programs may also be thought of as high-demand–high-readiness housing models.

Residential continuum model housing. As community care progressed and the community support movement emerged, it became apparent to stakeholders that persons with serious mental illness had rehabilitation goals—and, more recently, recovery goals—such as employment, illness self-management, and avoidance of adverse events, such as involvement with the criminal justice system (2). One mental health system response to this realization was to design housing continua that added to residential

treatment models certain housing interventions intended to foster consumer functioning and independence (9). These interventions were viewed as more “normalizing,” in that services are not routinely provided in the housing and staff presence in the housing is limited. Residents are expected to leave their housing during the day and attend treatment, day activities, or work. Embodied in the concept of a continuum is the idea that residents move from one housing model to another as they progress in their rehabilitation and recovery (8).

Residential continuum models overlap in their component programs with residential care and treatment models and permanent supported housing models (described below). However, these programs are unique because they have components of a continuum consisting of different housing models that residents are expected to move between depending on their course of recovery and rehabilitation. Residential continuum models may also be thought of as high-demand–high-readiness housing models.

Permanent supported housing model. Most recently, experience and the emergence of the recovery movement have suggested that residential continua are difficult to implement and perhaps counterproductive from the perspective of rehabilitation and recovery (9). Requiring residents to change their housing as they improve or regress in functioning is viewed as not “normalizing” and possibly harmful. Having to give up familiar living arrangements and in some cases important relationships because of improvements in functioning and advances in recovery is viewed as a disincentive to progress toward these goals. Moving persons out of autonomous settings and into more supervised settings when problems arise is also seen as unnecessarily adding to stress. Therefore, housing interventions emerged that assume that persons with serious mental illness can be provided with permanent housing into which staff support and treatment can be introduced as needed (wraparound services), allowing residents to remain in familiar settings and maintain relationships, regardless of their needs for support and treatment. We classified

housing interventions of this type as permanent supported housing models.

Permanent supported housing models generally provide low-demand or least restrictive environments. They have few rules of conduct, encourage service use in nonhousing settings, and view stable housing as a prerequisite for global lifestyle changes (the housing-first approach) (13). One type of permanent supported housing—housing ready—may require participation in services and may consider abstinence from alcohol or other drugs a prerequisite for program entry. Another type—housing first—does not require participation in services and follows harm reduction approaches to alcohol and other substance use. Given their less restrictive environments, permanent supported housing programs include persons who are in the early stages of readiness or motivation to make global lifestyle changes.

Nonmodel housing. In a subset of studies, housing outcomes are described for people living on the street, using shelter beds, or residing in housing that is described simply as a part of “treatment as usual.” For housing described as treatment as usual, no reference is made to any service provided specifically to find housing for individuals or to support them in it. Some of these studies merely describe outcomes under these living arrangements. However, treatment as usual is often compared with model housing, strongly implying the nonmodel nature of such housing. We refer to these living arrangements as naturally occurring nonmodel housing.

In discussing our meta-analysis, we review findings in the context of the logic and ethics of intervention testing developed for evaluating other medical and psychosocial interventions. Under this logic, the efficacy or effectiveness of an intervention is evaluated by comparing it with that of a placebo or of no treatment. If an intervention is identified that is more efficacious or effective than placebo or no treatment, then it is ethical to conduct studies that compare other interventions with that intervention to determine relative efficacy or effectiveness (14). This leads to the question: What constitutes an ethical and meaningful comparison of housing models?

Strayhorn (15) discussed the ethical and practical problems of approximating placebo comparisons in psychosocial intervention research. He proposed three alternatives to placebo and no-treatment controls that are relevant for this meta-analysis: the “minimal control,” the “dismantling control,” and the “alternative intervention control.” The minimal and dismantling controls consist of interventions that contain some but not all elements of the experimental intervention. We propose that naturally occurring nonmodel housing alternatives constitute minimal or dismantling controls for evaluating the effectiveness of model housing interventions. Therefore, for each outcome discussed below, we first compared model housing interventions to nonmodel housing to establish effectiveness. Then we compared alternative housing models to consider relative effectiveness.

Housing outcomes

The outcome most proximate to the intervention that model housing provides is stability in housing (9). Our review of previous housing studies found that other, more distal outcomes are also commonly investigated by use of diverse measures of similar underlying constructs and varying follow-up periods. The inclusion of these measures may reflect the theory that housing models enhance the effectiveness of other services by providing the preconditions for adherence to treatment plans. The outcome constructs that we attempted to code in our meta-analyses were stability in housing, psychiatric symptoms, psychiatric hospitalizations (hospital admissions and emergency room use), alcohol use, drug use, incarceration, employment, and satisfaction.

Sociodemographic and clinical characteristics

Housing intervention studies commonly report a number of sociodemographic and clinical characteristics, although they do not always use them in subgroup or covariance analyses. The sociodemographic characteristics that we sought to code for subgroup analyses were age (mean age of sample), gender (percentage female), race (percentage nonwhite), ethnicity (percent-

age Hispanic), and veteran status (percentage of veterans). The clinical variables we sought were diagnosis (percentage of persons with schizophrenia and bipolar disorder), level of functioning (Global Assessment of Functioning score <50 or an equivalent rating), and co-occurring disorders (percentage of persons with a mental illness and a substance use condition).

Housing attributes

Newman (6) has stressed the importance of studying the effects of specific housing attributes, such as the number of occupants in housing, neighborhood characteristics, and urban or rural setting. Our preliminary examination, like Newman's, showed that few studies provided data on these characteristics. Therefore, we did not extract data on housing attributes.

Study design characteristics

Evaluations of housing models, like evaluations of other psychosocial interventions, differ in their design quality. However, evaluating housing models poses more complicated challenges than evaluating many treatments (6). These challenges include the problem of defining meaningful comparison groups given that everyone resides somewhere, the ethical issues involved in random assignment of persons to housing that is perceived to be less than adequate, variability in how ostensibly similar housing interventions are implemented, and identification of resident- and housing-related covariates to control for in quasi-experimental designs.

Most studies that we were able to locate for use in our meta-analysis were single-group, pre-post, or post-only designs. Therefore, we used study arm as the unit of analysis in our meta-analysis, which prevented us from rating study quality in terms of the usual convention: randomized trial or observational study. However, in response to Newman's recommendations (6), we tried to code certain other indicators of data quality related to the psychometric properties of outcome measures and the way in which data were collected for use in exploring the correlates of data quality. The data quality indicators were whether measures were identified before the study,

reliability and validity of measures were reported or cited, fidelity measures were used for housing models, data collectors were blind to the intervention received, attrition or missing data were accounted for in analyses, and anomalous findings, if reported, were recognized and discussed.

Theory and hypotheses

Typically, the evaluations we reviewed were not theory guided. This lack of theory in housing research was noted by Newman (6) in 2001 and continues to be true. However, most studies reflected at least the implicit and general theory that housing outcomes are a function of type of housing or housing alternative (usually the independent variable), resident characteristics (usually moderating variables), and housing attributes (also usually moderating variables). We elaborated this theory with the following hypotheses, which we tested.

Hypothesis 1. Housing interventions are first and foremost intended to provide persons with acceptable places to continuously reside. Housing interventions will be most likely to achieve proximate outcomes reflective of stable housing.

Hypothesis 2. Housing interventions also can establish the preconditions for more distal outcomes, such as skill building, engaging in symptom treatment, exercising wellness behaviors, enacting normal role functioning, and developing self-esteem and other positive self-perceptions. Housing interventions should achieve these more distal outcomes to some degree but not to the degree that they achieve stable housing (6,9,14).

Hypothesis 3. Housing models will be differentially effective for different outcomes. Housing models with in-house services (residential care and treatment and residential continuum) will have stronger associations with reductions in psychiatric symptoms, hospitalizations, drug and alcohol use, and incarcerations than permanent supported housing. Permanent supported housing, because it is more normalized and flexible, will be associated with higher levels of employment and satisfaction (8,9,11).

Hypothesis 4. Housing interventions will be more effective when there is

person-environment fit between the characteristics of residents (viewed as proxy variables for their specific needs and preferences) and the characteristics of the interventions (5). Certain types of housing interventions are likely to be more effective than others for achieving certain outcomes with certain types of residents. Thus the question of the effectiveness of a housing intervention is a variation of the question posed for psychotherapy: What types of interventions work in what ways for whom?

Methods

To describe the status of evaluation research on housing for persons with serious mental illness and test the above hypotheses, we conducted a series of meta-analyses with the features described below.

Study selection

For this study, we searched MEDLINE and PsycINFO, combining the search terms “adults with severe and persistent mental illness” with key words related to housing models and homelessness, such as “housing,” “community residences,” “half-way houses,” “group homes,” “permanent supported housing,” and “homelessness.” Use of these search terms yielded a total of 3,986 unique studies. Attempts were made to obtain abstracts for all 3,986 studies. We also checked five narrative reviews (6,8–11) to identify any studies not discovered by our electronic search and queried housing experts cited in the literature, federal project officers for multisite evaluations conducted by the Substance Abuse and Mental Health Services Administration, and principal investigators for those evaluations about unpublished articles and reports that we might include in our analyses.

Abstracts and summaries were then reviewed to eliminate multiple reports of the same studies, narrative and qualitative reviews, studies that did not address adults with serious mental illness, studies conducted outside the United States (16–22), and studies that were judged as likely to lack necessary data. Each abstract was screened systematically by two raters using an electronic protocol. Inter-coder reliability was assessed and found acceptable ($\kappa=.99$). Using the criteria listed

above, we identified and were able to obtain 227 research and evaluation reports that met criteria for full review.

Once reports of all the studies identified were obtained, we rereviewed the studies before coding to ensure that they met the criteria (for example, a unique study, use of housing as an intervention, a majority of study participants with mental illness, and a high probability of synthesizable data). On the basis of these precoding reviews, we eliminated another 167 studies, leaving 60 studies for coding. Table 1 shows the reasons for exclusion and the numbers of studies excluded.

Actual coding indicated that 30 studies were either kin studies or contained insufficient data for synthesis. As a result, the final number of studies analyzed was 30; the studies examined 44 unique housing interventions, including nonmodel housing.

Data abstraction and coding

The studies identified for the meta-analysis were coded according to a coding guide available from the authors. Coding results were entered into an Excel spreadsheet. Information was extracted about study design, housing model or models, outcomes, and clinical and sociodemographic variables for the groups studied.

To evaluate the reliability of our coding, a subset of studies was coded by two coders. The average inter-coder agreement was approximately 87%, suggesting acceptable reliability. In addition, all studies coded were checked by a third person (one of the authors), who made final judgments in situations in which the coders disagreed.

Unit of analysis

Most of the studies we coded employed single-arm, pre-post, or post-only designs to evaluate the effectiveness of particular housing models. A small number of studies compared outcomes for persons in multiple housing models or compared outcomes for persons in housing models with outcomes for persons living on the street, using shelter beds, or residing in housing not identified as model housing. Given the predominance of single-arm designs, we made housing intervention the unit of analysis and, as study reports permit-

ted, we collected study information separately for each of the housing interventions in multigroup studies. Given the nature of the studies available, this method provided a reasonable alternative to a narrative summary (23–26). Because housing intervention was the unit of analysis for this investigation, all outcomes and resident characteristics are expressed as means, rates, or percentages for the housing interventions.

Synthesizing outcome measures

Studies measured outcomes in a variety of ways, using measures taken from diverse protocols or surveys applied over varying time periods. Follow-up periods, for example, ranged from six months to five years. This heterogeneity in the way outcome measures are reported in the mental health literature provides a unique challenge to meta-analysis.

To synthesize diverse measures of similar constructs for each housing intervention, we converted outcome scores to effect size measures appropriate for the data. In some cases this meant computing standardized change scores based on means. In others it required converting percentages to effect sizes. Typically, effect sizes from proportions are combined separately from those calculated from means. However, the two types of effect sizes can be combined if they are based on similar constructs (24,26,27).

Effect size scores were then combined, with each intervention weighted by number of participants and the inverse of variance to create random-effect estimates for all outcomes studied for each housing intervention. In studies that reported data for more than one follow-up, data for the last follow-up period was used for analysis. We then compared housing model effect sizes with the effect sizes estimated for persons in non-model housing, computing the standardized paired-difference effect size for each comparison. We used “comprehensive meta-analysis” (Biostat, version 2.3) to assist in the calculation. In this way, we were able to test whether persons in housing models had different outcomes from persons in nonmodel housing and to

Table 1

Reasons for exclusion of 167 studies from the meta-analysis and number excluded

Reason	Number excluded
Does not include a housing intervention	47
Is not based on a population in the United States	46
Focuses on predictive factors (of housing success and failure) and characteristics of housed and homeless individuals	20
Examines outcomes not associated with a housing condition, or outcomes not a focus of this meta-analysis	15
Lacks sufficient quantitative information to synthesize	18
Focuses primarily on treatment other than housing	12
Is not a unique study (data already extracted)	8
Lacks information to adequately determine whether the housing model is unique	1

compare the effectiveness of different housing models.

Subgroup analyses

When there were sufficient data, we also sought to determine for each outcome the housing interventions that worked best for specific subgroups (according to sociodemographic, clinical, and data quality variables) by categorizing outcomes as above or below the median for subgroup samples within the housing interventions. We then followed the meta-analytic procedure described above for each outcome, comparing effect sizes for housing intervention evaluations that differed in resident characteristics and data quality, when available data permitted.

Results

Interventions, designs, outcomes, and variables

Table 2 summarizes the data that we were able to extract for each of the housing interventions. Table footnotes include citations of the studies from which data were extracted (13,28–54,56,57). We found data for a total of 44 housing interventions. The largest proportion (41%) were permanent supported housing. The next most frequent type of housing model for which we were able to find data was residential care and treatment (29%). The number of interventions by decade suggests that evaluation studies paralleled the emergence of housing models: evaluations of residential care and treatment housing appeared most frequently in the first two decades, and residential continuum and permanent

supported housing interventions appeared most frequently in the most recent decade.

The evaluations represent data collected for 13,436 persons. The number of participants in the interventions studied ranged from 18 persons to more than 3,000 persons, which emphasizes the importance of weighting findings to reflect these differences.

A very small proportion of the intervention evaluations involved random assignment of participants. Because we did not compute effect sizes on the basis of within-study differences, randomization for our purposes does not necessarily imply that there were no differences between intervention groups. However, randomization may indicate a greater concern for avoiding study biases of all types. From this perspective it is heartening to see that the number of randomized studies increased by decade.

Possible data quality scores ranged from 0 to 13. We were able to rate data quality for only half of the evaluations. We were able to rate a higher proportion of the nonmodel housing interventions and the residential care and treatment interventions; these studies tended to have higher data quality scores.

Table 2 shows that the most frequently addressed outcomes were psychiatric symptoms, housing stability, and hospitalizations, in that order. Relatively few evaluations addressed individual outcomes. The distribution of outcomes also reduced the extent to which we could compare interventions for specific outcomes.

Table 2Characteristics of 30 housing intervention studies included in the meta-analysis, by housing category^a

Characteristic	Nonmodel housing (8 interventions) ^b		Residential care and treatment (13 interventions) ^c		Residential continuum (5 interventions) ^d		Permanent supported housing (18 interventions) ^e		Total (44 interventions)	
	N or %	% or range	N or %	% or range	N or %	% or range	N or %	% or range	N or %	% or range
Studies per time period (N)										
1983–1989	0		3		0		1		4	
1990–1999	3		6		2		3		14	
2000–2006	5		4		3		14		26	
Total participants (N and range)	389	18–91	1,636	26–580	3,457	119–3,338	7,954	21–3,338	13,436	
Follow-up periods (range of months)		6–24		3–12		12–24		6–52		
Randomized controlled trials (N)										
1983–1989	0		0		0		1		1	
1990–1999	1		3		1		2		7	
2000–2006	5		1		3		4		13	
Data quality score (M and range) ^f	9	3–13	10	3–12	—		5	0–10		
Outcomes measured (N of studies and %)										
Housing stability	6	75	3	23	5	100	3	17	17	41
Psychiatric symptoms	8	100	5	39	0	—	5	28	18	44
Hospitalizations	2	25	6	46	2	40	6	33	16	39
Alcohol abuse	3	38	2	15	2	40	4	22	11	27
Drug abuse	2	25	2	15	2	40	3	17	9	22
Incarceration	0	—	1	8	0	—	1	6	2	5
Employment	0	—	2	15	0	—	2	11	4	10
Satisfaction	4	50	5	38	0	—	5	28	14	34
Age of participants (M and range) ^g	39	32–44	35	28–50	40	37–40	41	34–57		
Female (M % and range) ^h	24	0–52	25	0–43	23	21–30	30	5–66		
Nonwhite (M % and range) ⁱ	58	23–90	64	15–90	73	50–73	63	15–90		
Hispanic (M % and range) ^j	17	10–24	18	5–24	15	15–15	9	3–15		
Veterans (M % and range) ^k	89	68–100	42	15–68	0	—	0	—		
Schizophrenia or bipolar disorder (M % and range) ^l	0	—	79	79	0	—	0	—		
Co-occurring substance use disorder (M % and range) ^m	59	35–100	100	67–100	81	62–100	73	34–100		

^a Studies typically examined more than one housing intervention model.^b Four studies included nonmodel housing (28,31,50,56).^c Ten studies included residential care and treatment model housing interventions (13,30,32–35,37,39,49, 57)^d Two studies included residential continuum housing model interventions (49,52)^e Fourteen studies included permanent supported housing interventions (40–49,51–54)^f Possible data quality scores ranged from 0 to 13, with higher scores indicating higher quality. Number of studies with data that could be rated: non-model, 8; residential care, 10; residential continuum, 0; permanent supported, 6^g Number of studies reporting age: nonmodel, 8; residential care, 13; residential continuum, 2; permanent supported, 16^h Number of studies reporting data on females: nonmodel, 8; residential care, 13; residential continuum, 4; permanent supported, 18ⁱ Number of studies reporting data on race: nonmodel, 4; residential care, 10; residential continuum, 3; permanent supported, 15^j Number of studies reporting data on Hispanics: nonmodel, 4; residential care, 8; residential continuum, 2; permanent supported, 12^k Number of studies reporting data on veterans: nonmodel, 3; residential care, 4; residential continuum, 0; permanent supported, 6^l Number of studies reporting data on these diagnoses: nonmodel, 0; residential care, 1; residential continuum, 0; permanent supported, 0^m Number of studies reporting data on co-occurring substance disorder: nonmodel, 7; residential care, 4; residential continuum, 3; permanent supported, 8

Table 2 also presents data on sociodemographic and clinical variables. One variable that we attempted to measure (persons with low functioning) is not shown in the table because of insufficient data. The data suggest relative similarity among housing types for the variables shown. Subgroup differences for outcome variables and interventions taken in combination are discussed below.

Housing model effectiveness

Table 3 presents the results of outcome comparisons for the four housing interventions. It shows the number of interventions synthesized and the random effect obtained for each housing intervention and outcome combination. For each outcome, if an effect size for a housing model differed significantly from the effect size for nonmodel housing, the significance level is

shown. If the effect sizes for any pair of housing models differed, these significance levels are also shown. The table does not include data on incarceration or employment because the number of evaluations that addressed these outcomes were too few for analysis.

Table 3 shows that all three housing models achieved significantly greater housing stability than nonmodel housing. This effect size was greatest for

permanent supported housing (.63, $p < .05$). However, none of the differences between the various housing models were statistically significant. It was suggested during the review process that we include findings from a specific review (49) that we had rejected because our inclusion criteria excluded reviews. We added the findings from the two relevant studies that were reported in that review (we did not have direct access to either study) to the studies shown in Table 3 and performed a post hoc analysis. This increased or did not change housing model effect size estimates for housing stability. In addition, it did not change the findings that all model housing options performed significantly better than nonmodel housing and did not differ significantly from each other.

With respect to psychiatric symptoms, only residential care and treatment differed significantly from nonmodel housing (effect size = .65, $p < .05$). However, this outcome was addressed by only one residential continuum housing evaluation and one permanent supported housing evaluation. The residential continuum evaluation yielded an effect size similar to that of residential care and treatment housing. The permanent supported housing evaluation yielded an effect size substantially lower.

For reduction in hospitalization, Table 3 shows that both residential care and treatment and permanent supported housing differed signifi-

cantly from nonmodel housing. Hospitalization reduction was not addressed by any evaluations of residential continuum interventions. Although the effect size for permanent supported housing was more than double that of residential care and treatment, this difference was not statistically significant, which suggests high variability within the groups.

With respect to alcohol abuse, none of the housing models differed from nonmodel housing. The effect size for residential care and treatment housing was high (.87) and many times greater than the effect sizes for the other two interventions, but these differences were not statistically significant, which again suggests high intramodel variability. Effect sizes for reduction in drug abuse were smaller than for other outcomes, and there were no significant differences between any of the interventions.

Permanent supported housing achieved the highest effect size for satisfaction (.73), and this effect size differed significantly from the ones achieved by nonmodel housing ($p < .001$) and residential care and treatment ($p < .05$), which were substantially lower.

Subgroup analyses

We were able to rate data quality for only half of the intervention evaluations (36,38–40,43,44,46,50–53). The distribution of these rated evaluations across housing types and outcome

variables was such that there was only one outcome variable on which we could compare interventions on the basis of whether the evaluation data were rated as low or high quality. This was the case even when we grouped studies by outcome and ignored housing intervention. Thus we could examine only the association between data quality and effect sizes for supported housing interventions and housing stability outcomes. For two of the supported housing evaluations, the data quality scores were high (51,54), and for another two the scores were low (52,53). Housing stability increased in these studies regardless of data quality. However, the increase was much less in the higher-quality studies (effect size of .22 compared with .79).

When we subdivided interventions into subgroups on the basis of sociodemographic and clinical variables, we also found that the distribution of interventions was such that analyses for combinations of outcomes and interventions were not possible. However, when we examined the association of sociodemographic and clinical variables with outcomes without regard to housing intervention, we found a number of outcomes on which we could compare two or more interventions by sociodemographic and clinical subgroups. Table 4 presents the results of these subgroup analyses.

As shown in Table 4, subgroups with more females uniformly experienced better outcomes than groups with few-

Table 3

Random effects for housing intervention and outcome combinations, by housing category

Outcome ^a	Nonmodel		Residential care and treatment		Residential continuum		Permanent supported housing	
	Number of interventions	Pre-post effect size	Number of interventions	Pre-post effect size	Number of interventions	Pre-post effect size	Number of interventions	Pre-post effect size
Housing stability	2	-.63	2	.48 ^b	3	.80 ^b	6	.63 ^b
Reduction in psychiatric symptoms	2	-.11	3	.65 ^b	1	.68	1	.08
Reduction in hospitalization	2	-.33	3	.34 ^b	0	—	8	.72 ^b
Reduction in alcohol abuse	3	.06	2	.87	2	.07	4	.21
Reduction in drug abuse	2	.2	2	.41	2	.3	3	.51
Increased employment	0	—	2	.27	0	—	2	.27
Increased satisfaction	3	-.38	5	.07	1	.55	2	.73 ^{c,d}

^a Too few evaluations measured incarceration or employment to include these outcomes in the analysis.

^b Significant at $p < .05$ compared with nonmodel housing

^c Significant at $p < .001$ compared with nonmodel housing

^d Significant at $p < .05$ compared with residential care and treatment

er females in terms of housing stability, hospitalization reduction, alcohol abuse reduction, drug abuse reduction, and satisfaction, although none of the individual differences were statistically significant.

Table 4 shows inconsistent findings for comparisons of certain subgroups. In comparisons of the pre-post effect sizes, we noted inconsistencies across all housing models on the five outcome areas for which we had sufficient data. For example, in studies reporting a majority of nonwhites, drug abuse reduction was greater than in studies reporting a majority of whites (.46 and .09, $p < .001$), as was hospitalization reduction (.32 and .22, not significant) and alcohol abuse reduction (.56 and .06, not significant). However, in studies with a majority of whites, housing stability was greater than in studies with a majority of nonwhites (.54 and .37, not significant) and satisfaction was greater (.19 and .09, not significant). We found a similar pattern of inconsistency in comparisons of the pre-post effect sizes of studies with a majority of Hispanics and studies with a majority of non-Hispanics. In the studies with a majority of non-Hispanics, drug abuse reduction (.47 and .09, $p < .001$) and

alcohol abuse reduction (.33 and .06, not significant) was greater, but in studies with a majority of Hispanics, housing stability was greater in this group (.42 and .23, not significant).

Finally, participants without co-occurring substance use disorders experienced greater reductions in alcohol and drug abuse than those with co-occurring disorders (effect sizes of .40 and .06, respectively, for both alcohol and drug abuse, not significant) and greater satisfaction (.22 and $-.34$, $p < .001$). Among the studies included in the analysis, only two were unpublished. Omitting them from the analysis did not alter the results.

Discussion

The comparisons reported here are all “observational”—that is, they are comparisons of housing interventions in which effect estimates from different studies were synthesized for housing models and compared, without controlling for differences among residents and with an imputed control condition that was statistically constructed by synthesizing effect sizes from different studies of nonmodel housing. Thus the associations found should be interpreted as consistent with, but not indicative of, causal rela-

tionships. Nevertheless, the effort expended in collecting data for more than 13,000 individuals makes it especially important to glean as much information as possible to guide practice and research.

We have no explanation for the lower data quality scores for studies of residential continuum model housing and permanent supported housing. The scores should be interpreted with caution because we were able to score only half of the evaluation studies. Our inability to rate the quality of some studies suggests that evaluation researchers should provide more information about study methods in their reports and that reviewers and editors should request this information. The substantial difference in housing stability between supported housing studies that had low- and high-quality data, although not statistically significant, is consistent with the meta-analysis literature, which suggests that lower-quality studies tend to show larger treatment effects (27,55).

The wide variety and sparse distribution of the outcomes studied suggest that the field has not settled on the most important outcomes to measure in housing evaluations. Greater consensus is needed to make housing

Table 4

Random effects for sociodemographic and clinical subgroups across all housing interventions^a

Outcome	Male		Female		White		Nonwhite		Hispanic ^b		Non-Hispanic		Co-occurring substance disorders		No co-occurring substance disorders	
	N ^c	Pre-post effect size	N ^c	Pre-post effect size	N ^c	Pre-post effect size	N ^c	Pre-post effect size	N ^c	Pre-post effect size	N ^c	Pre-post effect size	N ^c	Pre-post effect size	N ^c	Pre-post effect size
Housing stability	4	.31	9	.52	6	.54	7	.37	3	.42	8	.23	6	.63	7	.30
Hospitalization reduction	4	.01	5	.83	5	.22	2	.32	0	—	0	—	5	.56	4	.36
Alcohol abuse reduction	2	.03	7	.26	6	.06	3	.56	4	.06	5	.33	5	.06	4	.40
Drug abuse reduction	2	.05	5	.35	4	.09	3	.46*	4	.09	3	.47*	3	.06	4	.40
Increased satisfaction	6	.01	5	.20	2	.19	9	.09	0	—	0	—	2	-.34	9	.22*

^a For each subgroup, the number of interventions and effect size listed are derived only from the evaluation studies in which that subgroup accounted for more than 50% of the study sample. For example, a total of four intervention studies that reported on housing stability as an outcome had samples in which males constituted more than 50% of the sample.

^b Of any race

^c Number of interventions

* $p < .001$, for comparison of the pre-post effect size estimate

intervention studies more comparable.

Overall our findings support several of our hypotheses and lead to additional important conclusions. Hypothesis 1, in regard to stability as the outcome, is supported. The housing model interventions that we analyzed all addressed housing stability outcomes more than other outcomes, and all of the housing interventions achieved this outcome. Hypotheses 2 and 3, about the achievement of distal outcomes and the differential outcomes of various models, respectively, are also supported by the results. Outcomes other than stability were achieved, and some evidence was found that housing models differ in regard to which of these other outcomes they achieve. The fact that all housing models were effective in achieving stable housing and some achieved other outcomes provides ethical justification for randomly assigning persons to these models to better understand what models work best, in what ways, and for whom.

Our fourth hypothesis was that different housing models may have different outcomes for different types of persons with serious mental illness. The studies reviewed had too few data to allow us to test this hypothesis directly. However, our findings of some associations between outcomes and subgroup variables across housing interventions suggest the importance of collecting data to explore this hypothesis further in future housing evaluations. The consistent, but not statistically significant, effects found for gender are particularly noteworthy. The findings of better outcomes for persons without comorbid substance abuse may speak to the difficulties of serving persons with multiple disabilities. If housing interventions work differently for different groups of persons, these differences should guide program development and clinical decision making. Also, controlling for these variables in evaluations should clarify the effectiveness of different housing interventions.

Our meta-analysis showed that permanent supported housing is receiving increasing attention. It achieves stable housing, and residents are very satisfied with it. The latter finding is not surprising given the low-demand, flex-

ible nature of most permanent supported housing interventions. In concept, permanent supported housing can cost-effectively provide any housing-service bundle required to meet consumers' needs and achieve any outcome as well as or better than any other housing model. Our analysis suggests that this may not yet be the case, either because the concept is incorrect or because implementation has been flawed. Therefore, it may be a mistake to focus too much on just one housing model. Further evaluation research that randomly assigns persons to the three housing models described, investigates multiple outcomes, enrolls consumers in sufficient numbers for planned subgroup analyses, and follows other accepted research conventions should help determine whether the concept is correctly understood the practice implemented.

The limitations of this study stem from the nature of housing interventions and meta-analysis. Housing interventions are "socially complex services" (12) that are difficult to operationally define and categorize. Although our coders achieved reliability, it is possible that our categorizations of housing models were not always accurate. Meta-analysis is handicapped by the quality of the studies analyzed and the details provided in the study reports. Few of the housing studies we reviewed were randomized trials or even quasi-experimental studies. In addition, many reports did not contain detailed information about interventions, participants, or methods. Under these circumstances meta-analysis is a blunt instrument. Nevertheless, this approach has raised issues above and beyond those identified in narrative reviews.

Although we reviewed 30 studies of 44 interventions that included more than 13,000 individuals, the amount of data we obtained was not sufficient to fully answer the questions we set out to investigate. These questions were those designed to enable program planners and caregivers to better meet the needs of persons with serious mental illness. It is important to answer these questions. In 2001 Newman (6) wrote that there is a need for a housing research program that employs common measures and adheres to scientific conventions. This continues to be true.

Conclusions

The findings of this meta-analysis provide quantitative evidence consistent with the interpretation that housing models, compared with nonmodel housing, contribute to the achievement of stable housing and other favorable mental health outcomes. In addition, the results are consistent with the theory that different housing models achieve different outcomes for different subgroups of persons.

Acknowledgments and disclosures

This study was funded by cooperative agreement SM55236-03-2 from the Center for Mental Health Services, Substance Abuse and Mental Health Services Administration. The authors thank Matthew Jameson, B.A., Ben Cichocki, M.S., Laysa Ostrow, B.A., and Elizabeth Fischer, B.A., for their contribution in coding the studies for analysis and Crystal Blyler, Ph.D., for her support.

The authors report no competing interests.

References

1. Braginsky BM, Braginsky DA, Ring K: Methods of madness: the mental hospital as a last resort. *American Sociological Review* 35:951-952, 1970
2. Goldman HH, Morrissey JP: The alchemy of mental health policy: homelessness and the fourth cycle of reform. *American Journal of Public Health* 75:727-731, 1985
3. Banks S, McHugo G, Williams V, et al: A prospective meta-analytic approach to the analysis of multi-site data. *New Directions in Evaluation* 94:45-59, 2002
4. Rog D, Randolph F: A multisite evaluation of supported housing: lessons learned from cross-site collaboration. *New Directions for Evaluation* 94:61-72, 2002
5. Leff HS, McPartland JC, Banks S, et al: Service quality as measured by service fit and mortality among public mental health system service recipients. *Mental Health Services Research* 6:93-107, 2004
6. Newman SJ: Housing attributes and serious mental illness: implications for research and practice. *Psychiatric Services* 52:1309-1317, 2001
7. Rochefort D: *From Poorhouses to Homelessness: Policy Analysis and Mental Health Care*. Westport, Conn, Auburn House, 1993
8. Barrow S, Zimmer R: Transitional housing and services: a synthesis, in *Practical Lessons: The 1998 National Symposium on Homelessness Research*. Edited by Fosburg LD, Dennis DL. Washington, DC, US Department of Housing and Urban Development, US Department of Health and Human Services, 1999
9. Rog DJ: The evidence on supported housing. *Psychiatric Rehabilitation Journal* 27: 334-344, 2004
10. Drake R, Mueser K: A review of treatments for people with severe mental illnesses and co-occurring substance use disorders. *Psychi-*

- atric Rehabilitation Journal 27:360–374, 2004
11. Nelson G, Aubry T, Lafrance A: A review of the literature on the effectiveness of housing and support, assertive community treatment, and intensive case management interventions for persons with mental illness who have been homeless. *American Journal of Orthopsychiatry* 77:350–361, 2007
12. Wolf N: Using randomized controlled trials to evaluate socially complex services: problems, challenges, and recommendations. *Journal of Mental Health Policy and Economics* 3:97–109, 2000
13. Nuttbrock L, Rahav M, Rivera J, et al: Outcomes of homeless mentally ill chemical abusers in community residences and a therapeutic community. *Psychiatric Services* 49:68–76, 1998
14. International Conference on Harmonisation: Guidance on Statistical Principles for Clinical Trials. Federal Register 63:49583–49598, 1998
15. Strayhorn J: Control groups for psychosocial intervention outcome studies. *American Journal of Psychiatry* 144:275–282, 1987
16. Bhui K, Shanahan L, Harding G: Homelessness and mental illness: a literature review and a qualitative study of perceptions of the adequacy of care. *International Journal of Social Psychiatry* 52:152–165, 2006
17. Hogberg T, Magnusson A, Lutzen K: Living by themselves? Psychiatric nurses' views on supported housing for persons with severe and persistent mental illness. *Journal of Psychiatric and Mental Health Nursing* 13:735–741, 2006
18. Nelson G, Wiltshire C, Hall G, et al: Psychiatric consumer/survivors' quality of life: quantitative and qualitative perspectives. *Journal of Community Psychology* 23:216–233, 1995
19. Forchuk C, Nelson G, Hall GB: "It's important to be proud of the place you live in": housing problems and preferences of psychiatric survivors. *Perspectives in Psychiatric Care* 42:42–52, 2006
20. Hennessey C, Grant D: Developing a model of housing support: the evidence from Merseyside. *International Journal of Consumer Studies* 30:337–346, 2006
21. Morrell-Bellai T, Goering P, Boydell K: Becoming and remaining homeless: a qualitative investigation. *Issues in Mental Health Nursing* 21:581–604, 2000
22. Blid M, Gerdner A: Socially excluding housing support to homeless substance misusers: two Swedish case studies of special category housing. *International Journal of Social Welfare* 15:162–171, 2006
23. Lipsey M, Wilson D: *Practical Meta-Analysis*. Thousand Oaks, Calif, Sage, 2001
24. Becker B: Synthesizing standardized mean-change measures. *British Journal of Mathematical and Statistical Psychology* 41:257–278, 1988
25. Robertson C, Ruzni N, Idris N, et al: Beyond classical meta-analysis: can inadequately reported studies be included? *Drug Discovery Today* 9:924–931, 2004
26. Berkey C, Anderson J, Hoaglin D: Multiple-outcome meta-analysis of clinical trials. *Statistics in Medicine* 15:537–557, 1998
27. Wood L, Egger M, Gluud LL, et al: Empirical evidence of bias in treatment effect estimates in controlled trials with different interventions and outcomes: meta-epidemiological study. *British Medical Journal* 336:601–605, 2008
28. Champney TF, Dzurec LC: Involvement in productive activities and satisfaction with living situation among severely mentally disabled adults. *Hospital and Community Psychiatry* 43:899–903, 1992
29. Shern DL, Tsemberis S, Anthony W, et al: Serving street-dwelling individuals with psychiatric disabilities: outcomes of a psychiatric rehabilitation clinical trial. *American Journal of Public Health* 90:1873–1878, 2000
30. McGeary KA, French MT, Sacks S, et al: Service use and cost by mentally ill chemical abusers: differences by retention in a therapeutic community. *Journal of Substance Abuse* 11:265–279, 2000
31. McHugo GJ, Bebout RR, Harris M, et al: A randomized controlled trial of integrated versus parallel housing services for homeless adults with severe mental illness. *Schizophrenia Bulletin* 30:969–982, 2004
32. Leda C, Rosenheck R: Mental health status and community adjustment after treatment in a residential treatment program for homeless veterans. *American Journal of Psychiatry* 149:1219–1224, 1992
33. Rimmerman A, Finn H, Schnee J, et al: Rehabilitation of persons with severe mental illness in adult homes: the NYPCC study. *Psychosocial Rehabilitation Journal* 15:55–65, 1992
34. Hawthorne WB, Fals-Stewart W, Lohr JB: A treatment outcome study of community-based residential care. *Hospital and Community Psychiatry* 45:152–155, 1994
35. Lehman AF, Kernan F, DeForge BR, et al: Effects of homelessness on the quality of life of persons with severe mental illness. *Psychiatric Services* 46:922–926, 1995
36. Lehman AF, Possidente S, Hawker F: The quality of life of the chronic patients in a state hospital and in community residences. *Hospital and Community Psychiatry* 37:901–907, 1986
37. McCoy ML, Devitt T, Clay R, et al: Gaining insight: who benefits from residential, integrated treatment for people with dual diagnoses? *Psychiatric Rehabilitation Journal* 27:140–150, 2003
38. Yanos PT, Barrow SM, Tsemberis S: Community integration in the early phase of housing among homeless persons diagnosed with severe mental illness: successes and challenges. *Community Mental Health Journal* 40:133–150, 2004
39. Okin RL, Dolnick JA, Pearsall DT: Patients' perspectives on community alternatives to hospitalization: a follow-up study. *American Journal of Psychiatry* 140:1460–1464, 1983
40. Arthur Anderson LLP, University of Pennsylvania Health System Department of Psychiatry Center for Mental Health Policy and Services Research, Sherwood K, et al: Connecticut Supportive Housing Demonstration Program. Oakland, Calif, Corporation for Supportive Housing, 2002
41. Harder and Company Community Research: The Benefits of Supportive Housing: Changes in Residents' Use of Public Services. Oakland, Calif, Corporation for Supportive Housing, 2004
42. Mares AS, KasproW WJ, Rosenheck RA: Outcomes of supported housing for homeless veterans with psychiatric and substance abuse problems. *Mental Health Services Research* 6:199–211, 2004
43. Martinez TE, Burt MR: Impact of permanent supportive housing on the use of acute care health services by homeless adults. *Psychiatric Services* 57:992–999, 2006
44. Kessel ER, Bhatia R, Bamberger JD, et al: Public health care utilization in a cohort of homeless adult applicants to a supportive housing program. *Journal of Urban Health* 83:860–873, 2006
45. Pilisuk M: A job and a home: social networks and the integration of the mentally disabled in the community. *American Journal of Orthopsychiatry* 71:49–60, 2001
46. Schutt RK, Goldfinger SM, Penk WE: Satisfaction with residence and with life: when homeless mentally ill persons are housed. *Evaluation and Program Planning* 20:185–194, 1997
47. Sacks S, De Leon G, Sacks JY, et al: TC-oriented supported housing for homeless MICAs. *Journal of Psychoactive Drugs* 35:355–366, 2003
48. Proscio T: *Supportive Housing and Its Impact on the Public Health Crisis of Homelessness*. Oakland, Calif, Corporation for Supportive Housing, 2000
49. Shern DL, Felton CJ, Hough RL, et al: Housing outcomes for homeless adults with mental illness: results from the second-round McKinney program. *Psychiatric Services* 48:239–241, 1997
50. Clark C, Rich AR: Outcomes of homeless adults with mental illness in a housing program and in case management only. *Psychiatric Services* 54:78–83, 2003
51. Lipton FR, Nutt S, Sabatini A: Housing the homeless mentally ill: a longitudinal study of a treatment approach. *Hospital and Community Psychiatry* 39:40–45, 1988
52. Tsemberis S, Gulcur L, Nakae M: Housing first, consumer choice and harm reduction for homeless individuals with a dual diagnosis. *American Journal of Public Health* 94:651–656, 2004
53. Culhane DP, Metraux S, Hadley T: Public service reductions associated with placement of homeless persons with severe mental illness in supportive housing. *Housing Policy Debates* 13:107–163, 2002
54. Drake R, Yovetich N, Bebout R, et al: Integrated treatment for dually diagnosed homeless adults. *Journal of Nervous and Mental Disease* 185:298–305, 1997
55. Wolf F: *Meta-Analysis: Quantitative Methods for Research Synthesis: Series: Quantitative Applications in the Social Sciences*. Newbury Park, Calif, Sage, 1986
56. Folsom D, McCahill M, Bartels S, et al: Medical comorbidity and receipt of medical care by older homeless people with schizophrenia or depression. *Psychiatric Services* 53:1456–1460, 2002
57. Pryce I: The effects of social changes in chronic schizophrenia: a study of forty patients transferred from hospital to residential home. *Psychological Medicine* 7:127–139, 1977