Journal of Urban Health: Bulletin of the New York Academy of Medicine, Vol. 93, No. 4
doi:10.1007/s11524-016-0062-9


# Housing Quality in a Randomized Controlled Trial of Housing First for Homeless Individuals with Mental Illness: Correlates and Associations with Outcomes

Carol E. Adair, Brianna Kopp, Jino Distasio, Stephen W. Hwang, Jennifer Lavoie, Scott Veldhuizen, Jijian Voronka, Andrew F. Kaufman, Julian M. Somers, Stefanie R. LeBlanc, Sonia Cote, Sindi Addorisio, Dominique Matte, and Paula Goering

**ABSTRACT** *Housing quality (HQ) is associated with mental health, and may mediate outcomes in housing interventions. However, studies of housing interventions rarely report HQ. The purpose of this study was to describe HQ in a multi-site randomized controlled trial of Housing First (HF) in five Canadian cities and to examine possible differences by treatment group (HF recipients and treatment-as-usual (TAU) participants who were able to find housing through other programs or on their own). We also examined the association between HQ and the primary trial outcome: housing stability. The performance of a new multi-dimensional standardized observer-rated housing quality scale (the OHQS) in a relatively large cross-site sample was also of interest. HQ was rated by trained research assistants for 204 HF participants and 228 TAU participants using the OHQS. General linear regression models were used to examine unit/building quality scores by group and site adjusting for other group differences, and as a predictor of housing stability outcomes after 24 months of follow-up. The OHQS was found to have good reliability and validity, but because most of the neighborhood subscale items were negatively correlated with the overall scale, only unit and building items were included in the total HQ score (possible scores ranging from 13.5 to 135). Unit/building HQ was significantly better for the HF group overall (91.2 (95 % CI = 89.6–92.9) vs. 88.3 (95 % CI = 86.1–90.5); p = .036), and in one site. HQ in the TAU group was much more variable than the HF group overall (W (mean) = 24.7; p < .001) and in four of five sites. Unit/building HQ scores were positively associated with housing stability: (73.4 (95 % CI 68.3–78.5) for those housed none of the time; 91.1 (95 % CI 89.2–93.0) for those housed some of the time; and 93.1 (95 %*

Adair and Kopp are with the Mental Health Commission of Canada, Ottawa, Ontario, Canada; Adair is with the Department of Psychiatry, the University of Calgary, Calgary, Alberta, Canada; Adair is with the Department of Community Health Sciences, the University of Calgary, Calgary, Alberta, Canada; Distasio and Kaufman are with the Institute of Urban Studies, University of Winnipeg, Winnipeg, Manitoba, Canada; Hwang is with the Centre for Research on Inner City Health, St. Michael's Hospital, Toronto, Ontario, Canada; Lavoie is with the Department of Criminology, Wilfrid Laurier University, Brantford, Ontario, Canada; Veldhuizen and Goering are with the Centre for Addiction and Mental Health, Toronto, Ontario, Canada; Voronka is with the Department of Humanities, Social Science and Social Justice, University of Toronto, Toronto, Ontario, Canada; Somers is with the Faculty of Health Sciences, Simon Fraser University, Burnaby, BC, Canada; LeBlanc is with the Centre de recherche et de développement en éducation, Université de Moncton, Moncton, New Brunswick, Canada; Cote is with the At Home/Chez Soi Montreal Site, Montreal, Quebec, Canada; Addorisio is with the At Home/Chez Soi Vancouver Site, Vancouver, British Columbia, Canada; Matte is with the At Home/Chez Soi Montreal Site, Douglas University Institute for Mental Health, Montreal, Quebec, Canada.

Correspondence: Carol E. Adair, Mental Health Commission of Canada, Ottawa, Ontario, Canada. (E-mail: ceadair@ucalgary.ca)

*CI 91.4–94.9)) for those housed all of the time ($F = 43.9$ $p < .001$). This association held after adjusting for site, housing characteristics, participant ethnocultural status, community functioning, and social support. This study demonstrates that HQ can be as good or better, and less variable, in HF programs in Canada that systematically and predominantly source housing stock from the private sector compared to housing procured outside of an HF program. HQ is also an important predictor of housing stability outcomes.*

**KEYWORDS** *Housing quality, Standardized measures, Housing-related health research, Public/private housing, Homelessness, Mental illness*

## INTRODUCTION

The physical quality of housing has been associated, including more recently in intervention studies, with many physical health conditions, ranging from specific exposures and outcomes including infectious diseases, respiratory diseases, injuries, cardiovascular diseases, and overall mortality.[1] The evidence for connections between housing quality (HQ) and mental health is more recent but is also accumulating.[2,3] Two examples of single exposure/outcome associations are those between insufficient daylight and depressive symptoms and between noise levels and psychological distress.[2] The impact of HQ measured as a composite index of multiple aspects of housing has also been demonstrated for mental health. For example, Hwang et al. (2003) found that mental health status scores were lower for residents of lower quality rooming houses,[3] and Suglia et al. (2011) found a higher risk for depression among women experiencing housing disarray and instability.[4] Harkness and colleagues (2004) measured lower health service use costs for individuals with mental illness living in higher quality buildings,[5] and Wells and Harris (2007) demonstrated reduced psychological distress among low-income women after moves to higher quality housing.[6] Characteristics of neighborhoods have also been associated with health, mental health, and behavioral outcomes.[7] Relationships have been shown between a higher density of alcohol establishments in urban neighborhoods and rates of four types of violent crime,[8] between noise, overcrowding, absence of green space, community facilities, and fear of crime with lower mental health status and vitality scores,[9] and between neighborhoods defined in various ways as being disadvantaged and higher levels of depression and psychiatric distress.[10–12]

While this accumulation of evidence makes it clear that housing and neighborhood quality are important social and physical ecological variables that may mediate or moderate outcomes in housing intervention research, "objective" housing and/or neighborhood quality is rarely measured in studies of housing interventions for homeless individuals including those with mental illness. Studies typically report on the simple provision of housing without elaborating on quality. If included, HQ is usually measured either by self-report or using only a few observed indicators. This circumstance may be attributed to a lack of a standard, reliable, and valid objective HQ instrument.[13] We previously reported on the development and initial validation of an instrument for this purpose (the Observer-rated Housing Quality Scale, or OHQS).[1]

This article describes the systematic measurement of HQ and its relation to outcomes in a multi-site randomized controlled trial (RCT) investigating the effectiveness of Housing First (HF) for adults with serious mental illness (and substance use disorders if concurrent) in five Canadian cities (Vancouver, Winnipeg,

Toronto, Montreal, and Moncton). Housing First involves relatively immediate provision of housing and supports, through rent subsidies and usually in scattered site apartments in the private housing market, that is not conditional on treatment compliance or sobriety.[14] An important feature of HF is the encouragement of the participant to choose among a small number of available units; though in practice, this is sometimes constrained by availability. In the main trial, 2255 participants were randomized to HF and supports at two levels of intensity (either assertive community treatment or intensive case management) and respective treatment as usual (TAU) groups. Each team also had staff that administered rent subsidies (to keep the tenant's housing cost to a maximum of 30 % of their income), procured housing units and furnishings, moved tenants, and worked with landlords and tenants to achieve successful tenancies. Participants were then interviewed every 3 months for up to 24 months, on a battery of measures including housing stability, mental illness symptoms, functioning, and quality of life outcomes. The trial protocol is provided in detail in Goering et al.,[15] and 24-month outcomes are reported elsewhere.[16,17]

The cities differed in rental market conditions, so we knew that finding quality rental units would be more challenging in some sites. We also wanted to ensure that housing units procured in the private market with rent supplements from a publicly funded program were of acceptable quality. Aside from these descriptive objectives, in the present analysis, we also sought to examine the role of HQ as a mediator of the primary outcome of the trial (housing stability).

## METHODS

### Participants, Sampling, and Primary Outcome

For this sub-study, we selected random samples from among participants known to have obtained stable housing over the course of the study. We targeted 500 participants, with numbers at each site proportionate to the full trial sample (80 for Moncton and 120 for the other four sites). This sample size was based on a power analysis for a two-group comparison based on results from our HQ pilot study,[1] with oversampling to allow for the more complex analyses planned. In the main trial, a much smaller proportion of TAU participants obtained stable housing, but for this analysis, we aimed for a 1:1 ratio between groups to maximize analytic efficiency. The primary outcome variable for the trial was housing stability. Housing defined as "stable" included living in one's own room, apartment or house, or with family (with an expected duration of 6 months or more and/or tenancy rights) versus living on the street or other places not suitable for human habitation, temporary accommodations (<6 months and no tenancy rights), emergency shelters, and institutions.

### Measurement of Housing Quality

OHQS items (Table 1) were chosen based on the relevant literature and systematic assessment of stakeholder preferences. The instrument was designed for residences at the low-cost/rental end of the housing market and developed in English then translated into French (with a checked back translation) for units where the preferred language of the research assistant (RA) raters or participant was French. A pilot study was conducted in 55 housing units in Winnipeg and Toronto (see Adair et al. for details[1]). The OHQS has 34 items: 18 that assess features of the unit, 7 that assess building features, and 9 that assess neighborhood features, see Fig. 1 for an

**TABLE 1 OHQS Subscales and Items**

| Subscales and items | Method of rating or coding |
|---|---|
| Unit | |
| Safety/security | Observation plus interview item |
| Natural light | Observation |
| Artificial light | Observation plus interview item |
| Utilities—power | Observation plus interview item |
| Indoor air/ventilation | Observation plus interview item |
| Utilities—heating/cooling | Observation plus interview item |
| Utilities—water | Observation plus interview item |
| Utilities—plumbing | Observation plus interview item |
| Bathroom | Observation plus interview item |
| Structural | Observation |
| Kitchen/food preparation | Observation |
| Kitchen—appliances | Observation plus interview item |
| Bedroom/sleeping space | Observation |
| Noise | Observation plus interview item |
| Pests | Observation plus interview item |
| Storage | Observation |
| Design | Observation |
| Laundry | Observation plus interview item |
| Building | |
| Safety/security | Observation plus interview item |
| Staff | Observation plus interview item |
| Access/"visitability" | Observation |
| Inside condition | Observation |
| Outside condition | Observation |
| Garbage | Observation |
| Access to nature | Observation |
| Neighborhood | |
| Condition of nearby buildings | Observation |
| Condition of nearby streets | Observation plus interview item |
| Availability of core services | Coded using Walk Score® website "amenities"[a] |
| Locations for social involvement | Coded using Walk Score® website "amenities"[a] |
| Transportation access | Transit score from Walk Score® website[a] |
| Human activity | Observation plus interview item coded by consensus |
| Green space | Coded using Walk Score® website "amenities"[a] |
| Human illegal activity/crime | Crime rates coded using site police data[a] |
| Walk score | Walk score from Walk Score® website |

[a]Publicly available data for walk scores on the Walk Score® website http://www.walkscore.com/ was extracted in the Fall 2013 at a link labeled "use old version of Walk Score®". This information was useful because it was collected/calculated in a similar manner across sites. Core services included major grocery stores (with fresh produce and without membership requirements), pharmacies, banks (not "fast cash" businesses), post offices, family physicians/medical clinics, and public libraries. Major grocery stores and pharmacies were coded from standard lists compiled for each site. Locations for social involvement included coffee/tea shops with seating, places of worship, and community centers. Standard terms were used to search for core and social involvement amenities, and unique items found were coded as a count within 1 km radius. Green space was a count of green spaces, parks, river, or natural water front within a two block radius. For these three coded items, three raters independently coded 50 addresses and resolved discrepancies through discussion; then the remainder of addresses was coded by one individual. Transit scores were taken directly from the Walkscore® website for all addresses, except for those in Moncton where they were unavailable. For Moncton, a proxy transit score was calculated by staff at the Institute of Urban Studies at the University of Winnipeg using the distance to the nearest bus stop and frequency of service using methods based as closely as possible on published information.[18,19] Proxy and regular transit scores were then converted to Z scores and quintiles for all sites so that they were standardized for comparison. For the item human illegal activity/crime scores, each address was used to identify the corresponding neighborhood crime frequency as reported by each site's police service website. Neighborhood total, violent, and property crime data were divided by the associated population (Canadian Census) to obtain a crime rate for each participant's neighborhood. A Crime Severity Index Weighting Scheme[20] was applied to standardize resultant crime rates to allow comparisons across sites. For participants who lived outside a site's police jurisdiction, the mean score of the remaining neighborhood items was used in place of the crime rate

| Natural Light | .5 1 | 1.5 2 | 2.5 3 | 3.5 4 | 4.5 5 |
|---|---|---|---|---|---|
| | No source of light – space is dark if not artificially lit | Small or few windows and/or no fixtures for appropriate window coverings; e.g. resident has to use unconventional means to cover windows such as tin foil or plastic. Space is dim. | Smaller or few windows relative to space, with basic fixtures for coverings. Space has average brightness. | At least two moderate to large windows and fixture s that would allow for adequate coverings. Space is bright. | Windows large in relation to space in every or nearly every room (except bathroom) incl. windows on opposite or adjacent walls. Drapery fixtures permit good quality coverings. Space is very bright. |

**FIG. 1** Example item.

example item from the unit scale. Twelve items are rated by direct observation alone, 16 items are rated using observation *and* responses to standard interview questions to the participant occupant, and 6 items are completed using publicly available information.

Each item is rated or coded on a scale ranging from .5 to 5 in half-point increments. Internal consistency and inter-rater reliability were established in a pilot study for the unit and building subscales.[1] The neighborhood items were not populated in the pilot, so less is known about their performance.

## Other Measures

Demographic variables collected in baseline interviews for the main study included gender, education, prior month income, ethnocultural status, homelessness history, hospitalization history, social support, adverse childhood events, and community functioning (see Goering 2011 [15] for details). Ancillary information collected at the time of the HQ visits were photographs of the buildings' facades, building market type (private or public/social); building type (e.g., house, duplex, highrise); unit type (e.g., own unit, room in group home, single occupancy room); unit size (in square feet); resident restrictions (e.g., age, gender, disability status); pets (no, yes, or yes at additional cost); smoking (no, yes—in unit only, yes—anywhere); length of residency (in months); other occupants, if any, sharing the unit (alone, family, friend, or roommate); amount the participant paid toward rent; total rent charged; whether there was a signed lease (if so the term, if not how and how often rent was paid); and whether the participant shared space with the landlord or his/her family.

Quality of housing from the participants' perspective ("subjective" housing quality) was also collected in the main trial using the Perceived Housing Quality (PHQ) Scale, a set of housing-related satisfaction items selected from two sources [21,22] which was administered in the trial in 6-month intervals over the course of the study. The PHQ scale score most proximal to the date of the visit was used.

The primary outcome variable "time in stable housing" was measured using the Residential Time-Line Follow-Back (RTLFB) [23] calendar which documents the number of days in each type of residence, dates of and reasons for moves, and household composition. The RTLFB was administered via interview at 3-month intervals over the two-year follow-up of the main trial. Because of the non-normal distribution of number of days housed (a preponderance of values at both extremes), we operationalized it as the proportion of time spent in stable housing for the 6 months prior to the end of the study and categorized as "all of the time," "some of the time," and "none of the time."

## Setting and Procedure

Housing units were sampled from a list of participants housed for at least 2 months at each interview point from the main trial in each of the five cities. The two-month

period was necessary to ensure that approximately 3 months of residency would have elapsed to allow for the recall period for the interview-based items. HF participant units were randomly sampled from the list, but because a smaller proportion of TAU individuals were able to procure housing over the 24-month follow-up period, *all* TAU participants were included.

Units were excluded for participants who refused participation, could not be reached after three attempts, or for participants who would require an interpreter. Refusal rates were estimated for four sites (average 9.9 %; range 1.6 to 13 %). Participants' housing status was confirmed at the time of visit scheduling, and they were told again about the purpose and process of the visit. They were assured that participation was voluntary; that the researchers were not interested in how the place was kept, but rather the quality of the structure itself (e.g., doors, windows, appliances); that RAs may need to check on specific aspects of the unit (e.g., look under sinks); that they would also be asked some questions about their place during the visit; and that they would receive $20 CAD for their participation. The study was reviewed and approved as a revision to the main trial protocol by academic REBs in each jurisdiction.

Prior to the visits, RAs were formally trained in use of the OHQS including in situ practice. During the visits, RAs made independent ratings; they were permitted to discuss characteristics of the space during the visit but not reveal or compare their actual ratings. Given that no single description can represent all variation in housing characteristics, RAs were instructed to choose the category representing "best fit" rather than "perfect fit." During the visit, ratings were made on paper forms and the data were subsequently entered to the main study database. Interviewer safety protocols were adapted from those for the main trial interview visits including research office check-in calls before and after visits, attending visits in pairs, and training in how to terminate interviews if safety concerns arose. Given the potential for transfer of bedbugs from infestations in some units, a standard prevention protocol was used, and no such incidents occurred. Visits took on average 45 min for assessment of all three components.

### Data Analysis

A maximum of two missing values for both raters were permitted for the unit subscale and one for the building subscale for the case to be included in analysis. This resulted in a loss of eight records (1.9 %). An additional 31 records required minor cleaning such as where an unpermitted value was adjusted to the closest value or where a value from one RA was used where one was missing from the other RA.

Intra-class correlation coefficients were used to calculate inter-rater agreement for the unit and building items (most neighborhood items were not rated). Ratings were then combined across the two raters in each case using the simple arithmetic average. Next item, subscale and total scale score summary statistics and distributions were generated and inspected. Cronbach's alphas were used to calculate internal consistency of subscales. In bivariate analysis, subscale scores were examined graphically by site, demographic, and housing-related variables. Differences between the study sample and the full trial sample and bivariate associations by group and with outcomes were tested using Student's *t* tests or chi-squared tests. Due to the large number of tests, we adopted a more conservative alpha level of 0.01 for bivariate comparisons. Overall differences in HQ were tested between study groups and sites using independent samples *t* tests or analysis of variance and an alpha level of .05. Levene's tests were used to examine differences in variance (based on the mean). Multiple linear regressions were used to examine

predictors of HQ, and ordinal logistic regression was used to explore the relationship between HQ scores and housing stability outcomes, adjusting for variables that appeared to be important in the bivariate analysis. Alpha levels used for the models were .05. Collinearity and fit statistics were also generated.

Regarding the selection of variables for the models, we considered the role of each variable in four steps: (1) the theory/logic model of how HF leads to housing stability outcomes more generally and what predictors are important; (2) differences between the full trial sample and our HQ sample; (3) group differences in the sample (Table 2); and (4) bivariate associations with each dependent variable.

## RESULTS

### OHQS Psychometrics

Inter-rater reliability (applicable to a single rater) was very good to excellent for all 27 rated items. Intra-class correlation coefficients were .8 or above for 20 items and .7 or above for the rest, ranging from .89 for "pests" to .72 for "condition of nearby streets" and "safety/security". Internal consistency of subscales was very good and slightly higher than the pilot for the unit and building subscales (Cronbach's alpha = .94 and .83, respectively, and .81 for the neighborhood subscale). However, six of the seven neighborhood items were negatively correlated with the total set of items, and the items of the other two subscales. Therefore, we did not consider it to be appropriate to combine the three subscales into an overall HQ construct and total score, and for subsequent analyses, we used the unit and building total score (possible scores ranged from 13.5 to 135). Validity of the unit/building scale was supported by a high correlation between it and Perceived Housing Quality (PHQ) total scale scores ($r = .51$; $p < .001$) and a significantly lower unit/building score for participants in both groups who reported a perceived need for major repairs (mean 93.3 SD 12.9 vs. 84.6 SD 16.0; $p < .001$).

### Sample Characteristics Compared to the Full Trial Sample

The final sample was 432 (53 from Moncton, 101 from Montreal, 108 from Toronto, 84 from Winnipeg, and 86 from Vancouver). There were no significant differences between the participants in the HQ study sample and the main trial sample (either group) in gender, age, aboriginal origin, ethnocultural status, education, prior criminal activity, diagnosis (including alcohol and substance dependence and abuse), suicidality, history of learning problems, or past month income at baseline using bivariate analyses.

The proportion of TAU participants with a high need level in the HQ study sample was lower than that in the full sample (34.9 vs. 55.4; $x^2 = 6.66$; $p = .01$) and had higher interviewer-rated mean functioning relative to TAU participants in the full sample (58.9 vs. 61.6; $t = -4.24$; $p < .001$), but the actual differences were not large, and this pattern was not true for HF participants. More HF participants in the HQ sample were born in Canada than HF participants in the full sample (88 vs. 80 %; $x^2 = 8.02$; $p < .005$). These differences helped inform the initial choice of variables for the regression models.

### Demographic Characteristics of the Sample by Group

There were no significant differences between the two groups at the .01 level on a range of baseline demographic and clinical variables. The mean ages at enrolment for HF and TAU were 42.2 and 42.1, and the percentages of males were 63.7 and 63.2. The age at first homelessness was 33.2 and 32.8 years, and the total lifetime month's homeless were 52.8

**TABLE 2 Sample housing characteristics by group**

| Housing characteristics | | | | |
|---|---|---|---|---|
| Variable | | HF ($N = 204$) | TAU ($N = 228$) | $p$ value |
| Site | Freq (%) | | | |
| Moncton | | 30 (14.7) | 23 (10.1) | $x^2 = 6.13$ |
| Montreal | | 53 (26.0) | 48 (21.1) | $p = 0.20$ |
| Toronto | | 42 (20.6) | 66 (29.0) | |
| Winnipeg | | 40 (19.6) | 44 (19.3) | |
| Vancouver | | 39 (19.1) | 47 (20.6) | |
| Building market type | Freq (%) | | | |
| Private market | | 163 (79.9) | 88 (38.6) | $x^2 = 79.3$ |
| Public/social housing | | 37 (18.2) | 126 (55.3) | $p < 0.001^{\#}$ |
| Both/mixed | | 0 (0) | 9 (4.0) | |
| Don't know | | 3 (1.5) | 5 (2.2) | |
| Type of building[a] | Freq (%) | | | |
| High rise | | 53 (27.2) | 63 (28.8) | $x^2 = 9.23$ |
| Low rise | | 121 (62.1) | 117 (53.4) | $p = 0.06$ |
| Fourplex | | 5 (2.6) | 2 (0.9) | |
| Duplex | | 4 (2.0) | 8 (3.7) | |
| House | | 12 (6.2) | 29 (13.2) | |
| Number of units | mean (SD) | 64.4 (104.5) | 57.8 (73.8) | $t = 0.74$ $p = 0.46$ |
| Length of current tenancy (months) | mean (SD) | 16.6 (7.9) | 15.1 (9.3) | $t = 3.40$ $p < 0.001$ |
| Occupant composition | Freq (%) | | | |
| Alone | | 148 (80) | 156 (75) | $x^2 = 1.20$ |
| Family | | 24 (13) | 32 (15) | $p = 0.55$ |
| Friend/roommate | | 14 (7) | 20 (10) | |
| Cost for rent to tenant (per month in CAD) | mean (SD) | 366.2 (224.5) | 374.3 (187.0) | $t = -0.393$ $p = 0.69$ |
| Total cost of rent for unit (per month in CAD) | mean (SD) | 669.5 (233.8) | 462.5 (246.0) | $t = 8.59$ $p < 0.001$ |
| Signed lease | Freq (%) | | | |
| Yes | | 162 (88.0) | 147 (69.7) | $x^2 = 19.5$ |
| No | | 22 (12.0) | 64 (30.3) | $p < 0.001$ |
| Unit type | Freq (%) | | | |
| Room in group home | | 4 (2.1) | 11 (5.3) | $x^2 = 72.5$ |
| Room in rooming house | | 8 (3.9) | 36 (17.3) | $p < 0.001^{\#}$ |
| SRO room, alone | | 5 (2.7) | 39 (18.8) | |
| Other | | 0 (0) | 6 (2.9) | |
| Own unit, alone | | 149 (79.3) | 84 (40.4) | |
| Own unit, shared | | 22 (11.7) | 32 (15.4) | |
| Unit size (in square feet) | mean (SD) | 464.2 (242.6) | 307.7 (222.9) | $x^2 = 6.74$ $p < 0.001$ |

[a]Percentages are based on non-missing data. Missing data was minimal except in the cases of type of building (9 missing values in HF, 9 missing values in TAU); signed lease (20 missing values in HF, 17 missing values in TAU); unit type (16 missing values in HF, 20 missing values in TAU)

[#]$p$ values are based on Fisher' exact test because of zero or very small cell frequencies

and 49.1. 89.2 and 91.2 % were unemployed, respectively. Country of birth, ethnocultural group (including Aboriginal), adverse childhood experiences, receipt of help in school, educational attainment, psychiatric diagnosis, mental illness symptoms, substance use problems, suicide risk, community functioning, comorbid physical health conditions, history of traumatic brain injury, prior psychiatric hospitalizations, and prior justice system involvement were also not significantly different between groups. As expected, there were some important between-group differences in the characteristics of the housing they obtained during the study (Table 2).

Distributions of unit/building HQ scores are shown overall by group and for sites by group in Figs. 2 and 3, respectively. Overall mean HQ scores were higher for the HF group (91.2 (95 % CI = 89.6–92.9) vs. 88.3 (95 % CI = 86.1–90.5)), and the difference was significant ($t = 2.11$, $p = .036$). However, at the site level, HQ was statistically significantly better in only one site (Vancouver) where the mean unit/building scores were 95.9 ((95 % CI 91.4–100.4) (HF) and 81.9 (95 % CI 75.4–88.4) (TAU); $t = 3.58$, $p < .001$, respectively). Very notably, *variability* in housing quality was much greater in the TAU group than the HF group (Levene's test for homogeneity of variance ($F = 24.7$, $p < 0.001$)). Variability was also significantly greater for the TAU group in every site except Winnipeg (Moncton $F = 10.01$, $p < .003$; Montreal $F = 7.58$, $p < .008$; Toronto $F = 5.71$, $p < .02$; Winnipeg $F = 3.66$ $p = .059$; Vancouver $F = 15.31$, $p < .001$). Note that the pattern was present in Winnipeg as well, but the difference was not quite statistically significant.

Table 3 provides model parameter values for HQ by group for all sites after adjusting for variables considered to be important in the bivariate analysis. After adjustment, HQ was no longer significantly different by study group overall. In the model, site (Winnipeg), market type, unit type, and unit size were significant co-predictors of HQ, but building type, cost of rent to the tenant, co-occupant composition, community functioning, and social support were not.

In bivariate analysis, unit/building HQ was strongly associated with the primary outcome of the main trial: housing stability in the final 6 months of the study. Mean unit/building scores were 73.4 (95 % CI 68.3–78.5) for those housed none of the time; 91.1 (95 % CI 89.2–93.0) for those housed some of the time; and 93.1 (95 % CI 91.4–94.9)) for those housed all of the time ($F = 43.9$ $p < .001$). The association held after adjusting for other potentially important confounders (Table 4). Treatment group, Winnipeg site, unit rent cost, presence of a signed lease, and Aboriginal status were also significant co-predictors of housing stability but market type, unit size, participant community functioning, and social support were not. An interaction between group and HQ was also included in this model to test whether the association between housing quality and stability differed by group, but it was not significant so was not retained.

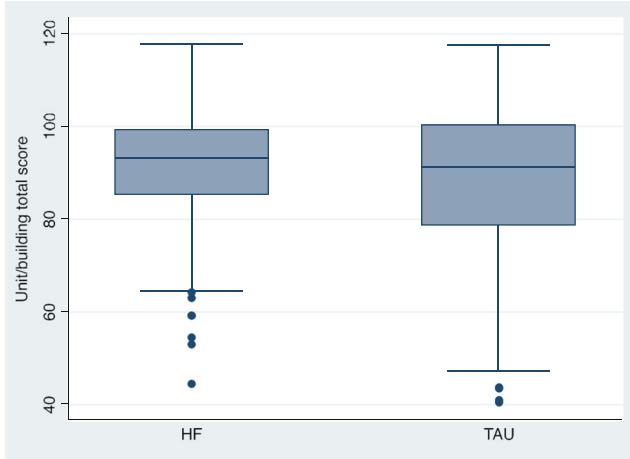


**FIG. 2** Housing quality (unit and building combined) by treatment group.

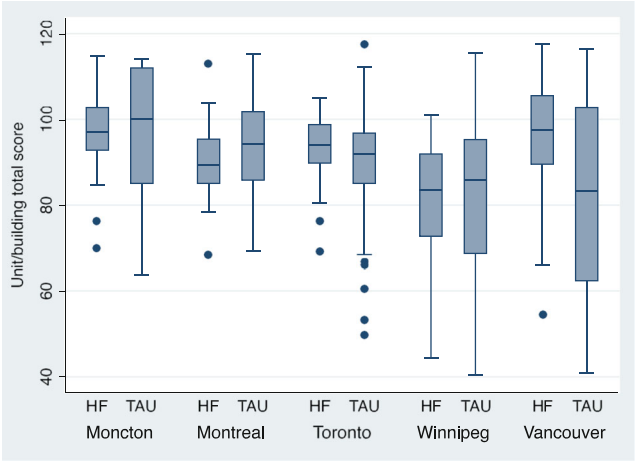


**FIG. 3** Housing quality (unit and building combined) by treatment group and site.

## DISCUSSION

In bivariate analysis, HQ was found to be better, on average, for the cross-site sample of individuals who received a HF intervention as part of the At Home/Chez Soi trial compared to those in the trial's TAU group who became housed through other means. The superiority of HF unit/building HQ held at the site level for only one site; at the other sites, average HQ was not different by group. There was no significant difference in HQ between groups after adjustment for other variables. However, and more importantly, the quality of the housing obtained by TAU participants was significantly more variable, with some units at higher and (of greater concern) at the lower ends of the HQ distributions. Our findings suggest that the presence of specialized staff dedicated to sourcing housing and rent subsidies can make a small difference in HQ and a substantial difference in housing consistency.

In our multivariate model of HQ other, even stronger, correlates were market and unit type and unit size and site. In terms of market type, public/social housing was found to be strongly associated with HQ on average after consideration of site, and other housing variables. However, given the low supply of this type of housing, at least for the foreseeable future, it is reassuring that the levels of HQ for the HF group, who were predominantly placed in private sector units, were of similar quality overall and were less variable overall. Additionally, some individuals prefer to be in regular market housing. The HF participants were offered some choice over housing, and it was observed by housing teams that some participants accepted a lower quality unit in a preferred neighborhood. Given the negative association between neighborhood items and unit/building items, this tendency could have lowered average unit/building scores for the HF group, although it could be argued that the same preferences may have been in play for TAU group participants who were able to exercise choice. Our development work showed that on average individuals considered aspects of the unit to be most important to them, then the neighborhood, then the building, but the dynamics of individual choice as well as family, community, and system factors that result in particular living circumstances are in need of greater elucidation. Rooming house rooms, single occupancy rooms, and smaller units were also of lower mean HQ, likely because of fewer amenities, shared amenities, and less privacy. The specific OHQS items that seem to discriminate features of different types of accommodation warrant further examination.

**TABLE 3 Housing quality (unit and building) by study group**

| Variable: | Coefficient (95 % CI) | *p* value |
|---|---|---|
| Group | | |
| TAU (referent) HF | 1.74 (−1.11–4.60) | 0.23 |
| Site | | |
| Moncton (referent) | – | – |
| Montreal | −1.04 (−5.46–3.38) | 0.64 |
| Toronto | −4.00 (−9.12–1.15) | 0.13 |
| Winnipeg | −12.68 (−17.5–7.9) | <.001 |
| Vancouver | −2.19 (−7.22–2.84) | 0.39 |
| Building type | | |
| Highrise (referent) | – | – |
| Lowrise | −.973 (−2.20–4.15) | 0.55 |
| Fourplex | 3.27 (−5.00–11.54) | 0.44 |
| Duplex | −2.43 (−9.03–4.17) | 0.47 |
| House | −1.85 (−6.73–3.03) | 0.46 |
| Market type | | |
| Private (referent) | – | – |
| Public/social | 9.00 (5.95–12.06) | <.001 |
| Both/mixed | 17.61 (8.43–26.8) | <.001 |
| Unit type | | |
| Own unit alone (referent) | – | – |
| Room in group home | −7.91 (−16.7–.893) | 0.08 |
| Room in rooming house | −13.53 (−18.03–9.04) | <.001 |
| Single occupancy room | −26.41 (−30.9–21.95) | <.001 |
| Own unit (shared) | −3.38 (−8.97–2.21) | 0.24 |
| Other | −11.09 (−20.70–1.48) | 0.02 |
| Cost for rent to tenant (per month in CAD) | .004 (−.001–.01) | 0.14 |
| Unit size (in square feet) | .008 (.003–.013) | <.001 |
| Occupant composition | | |
| Alone (referent) | – | – |
| With family | 1.36 (−3.68–6.40) | 0.60 |
| With friend or roommate | −4.21 (−9.94–1.54) | 0.15 |
| Community functioning (MCAS mean total score) | −.021 (−.189–.148) | 0.81 |
| Social support (presence of confidante) | 1.91 (−.462–4.27) | 0.11 |

$N = 340$; $F\ (22, 317) = 15.19$, $p < .001$; $R^2 = .51$; adjusted $R^2 = .48$. Fit–log-likelihood full model—1275.35 LR test 244.741 $p < .001$

Mean VIF value 1.84; range 1.06–3.57

The market and social contexts in each site and the mix of housing procured by HF teams and otherwise available varied widely across sites in bivariate analysis. Significant differences in HQ between groups in Vancouver are likely attributable to the concentration of very low quality accommodations in the city's urban core, where homelessness is most visible.[24,25] After adjustment for other important variables, site was a significant predictor of HQ only for Winnipeg. There are important contextual differences with this city that can explain these findings. It is relatively younger city with slower growth historically and an inner city that has not yet reached the stage of revitalization that is characteristic of larger cities. The supply of housing overall as well as affordable housing is constrained, with almost no secondary ring of affordable suburban housing stock. In contrast, the rent subsidies enabled access to a larger stock of suburban apartments in Vancouver despite high average rental costs, and some participants reportedly preferred these places to the inner city options because of light, noise, and safety issues. There is also greater residential mobility in Winnipeg, a lower than average household income, and a low cost of living. Winnipeg has the highest percentage of citizens of Aboriginal heritage which in the main trial were intentionally oversampled to test a culturally relevant adaptation of HF. Their history of

**TABLE 4 Unit/building quality as a predictor of housing stability**

| Variable: | OR (95 % CI) | *p* value |
|---|---|---|
| Unit/building score | 1.04 (1.02–1.06) | <.001 |
| Group | | |
| TAU (referent) | – | – |
| HF | 2.31 (4.12–1.30) | .004 |
| Site | | |
| Moncton (referent) | – | – |
| Montreal | 0.89 (.34–2.33) | 0.81 |
| Toronto | 1.54 (0.51–4.72) | 0.45 |
| Winnipeg | 3.51 (1.11–11.06) | 0.03 |
| Vancouver | 0.68 (0.23–2.05) | 0.50 |
| Market type | | |
| Private (referent) | – | – |
| Public/social | 1.40 (0.73–2.70) | 0.31 |
| Both/mixed | 1.58 (0.27–9.25) | 0.61 |
| Cost for rent for the unit (per month in CAD) | 1.0017 (1.003–1.0031) | 0.02 |
| Cost for rent to the tenant (per month in CAD) | 1.0004 (0.999–1.0019) | 0.57 |
| Signed lease | | |
| No (referent) | | – |
| Yes | 2.38 (1.25–4.51) | 0.01 |
| Unit size (in square feet) | 1.0001 (0.999–1.0013) | 0.79 |
| Ethnocultural status | | |
| Non-aboriginal/ non-ethnocultural (referent) | – | – |
| Aboriginal | 0.46 (0.2–1.08) | 0.07 |
| Ethnocultural | .97 (0.56–1.68) | 0.91 |
| Community functioning (MCAS total score) | 1.02 (0.98–1.05) | 0.30 |
| Social support (presence of confidante) | | |
| No (referent) | | – |
| Yes | 1.04 (0.65–1.65) | 0.88 |

$N = 317$; log-likelihood full model—259.4; LR test 98.65 $p < .001$ mean VIF value 1.85; range 1.07–3.43

colonialism and associated systemic and individual trauma and discrimination was observed to play a role in their access to housing by the housing teams at that site. Finally, at the time of the study, Winnipeg was entering a period of economic and population growth, and the vacancy rate was very low.

The neighborhood items of the OHQS were negatively correlated with the unit and building items. A "first blush" explanation is that properties with better access to transit and amenities tend to be located in the inner city where buildings may be older, in poorer condition, and which are more affordable to individuals with fewer resources. This is likely an oversimplification of the complexity of the built urban environment, and greater examination of the neighborhood variables in relation to the other findings, particularly the site findings, is planned.

In our second model, HQ was also found to be strongly and positively associated with housing stability, even after adjusting for a range of housing characteristics, participant functioning, and social support. In this model, important correlates of housing stability were also Winnipeg site, unit cost, and presence of a signed lease. These associations were all in the expected direction. In bivariate analyses, one variable (unit type) was so strongly associated with *both* housing quality and housing stability that it was considered to be on the causal pathway of this association and as such was not included in the model to avoid obscuring the association of primary interest.

The person-level variables including baseline community functioning and social support did not predict HQ or housing stability. For the HF group, this is more easily explained as the housing staff was in control of housing unit procurement. It is not clear why no personal characteristics predicted HQ for the TAU group when they might be expected to play some role. Our results support the idea that, at least in this population, access to and procurement of housing likely have more to do with opportunity and external circumstances such as market familiarity, vacancy rates, financial resources, and property management than personal aptitude. Only some of these higher level variables were measured directly in our study. The good news is that the important variables seem to be modifiable and HF programs can take direct action on them.

It is important to underscore that our group comparison was between HF participants who received HF program housing (and had some choice among units), and TAU participants *who were able to become housed* on their own or using other available housing programs, rather than *all* TAU participants. Many TAU participants (and a few HF participants) remained homeless in the larger trial. Had we compared the living conditions of HF participants with *all* TAU participants; there is no question that the HF group would have had substantially better conditions simply due to a much smaller proportion living in shelters or on the street. However, our interest was in the HQ of units located by housing teams in the intervention relative to units located by homeless individuals with mental illness on their own or through pre-existing programs in the respective communities.

### Limitations and Strengths

While embedded in a RCT, the research design reported here is quasi-experimental. Group differences and therefore potential selection biases came in at both the stage of selecting the participants from the larger sample and in the portion of the subgroups able to be housed for the HF group and those able to procure housing other ways (for the TAU group). However, differences on baseline variables were few. Our study included a comprehensive array of variables but did not capture several key aspects of the process of housing attainment, and some were not optimally measured. For example, while our measure of community functioning was observer-rated and standardized, our measure of social support was very superficial.

A large number of statistical tests were used in the study, especially for bivariate comparisons. While we set the alpha level to .01 for all bivariate analyses to keep the focus on the most meaningful differences, there is an increased probability that some of the differences reported are attributable to chance alone.

The OHQS is very new and although reliability and validity seem to be very promising, the instrument is as yet without a long track record and has not been independently validated. Further examination of its psychometric properties via classical test theory and item response theory is warranted and we have not yet sufficiently analyzed ancillary (including qualitative) data that were collecte. We also have yet to develop evidence-based cut points that would aid interpretation of scale scores for practitioners. In particular, better understanding is needed of the neighborhood items and how they relate to HQ for this population. The implications of this inverse relationship for practice are not clear at this point; practitioners and clients are best placed to consider trade-offs between quality of units and locations with important individual amenities or supportive social connections. Even so, the items on the OHQS could help guide such discussions. With further validation, the OHQS may also play a role in more systematic and comparable measurement in housing and health-related research. At the level of policy, HQ measurement can ensure accountability for a standard of HQ in programs funded with public dollars. The OHQS is

in the public domain. Minor adaptation of the tool for practice use, including a pre-occupancy version, and further psychometric analysis is underway.

## CONCLUSION

This is the first time, to our knowledge, that systematic and objective measurement of HQ has been undertaken in the context of a HF intervention trial. Our analysis confirmed that HQ was at least as good, and less variable, for units and buildings provided in a HF intervention than those obtained by TAU participants. HQ was also found to be significantly associated with housing stability. With further validation, the OHQS may be a useful tool for HF and related programs for procuring quality housing for individuals with mental illness and reducing housing instability and related homelessness. These findings provide initial confirmation that it is possible, even in tight rental markets, for HF program staff to locate quality housing for clients, and that the quality of that housing is an important contributor to housing stability.

## ACKNOWLEDGMENTS

The authors wish to dedicate this paper to the memory of our senior co-author Dr. Paula Goering, who passed away on May 24, 2016. Dr. Goering was the lead investigator on the At Home/Chez Soi trial – the largest study of its kind in the world, from 2007 to 2016. She was a consummate collaborator, leader and visionary, lighting the way for our research and practice teams to succeed, and also had an unwavering commitment to the involvement of people with lived experience in research. Through the study she not only oversaw the provision of housing and supports to more than 1000 Canadians, but her legacy lives on in national policy, including funding support for Housing First. In the study, and more broadly, she was a champion for improving the lives of marginalized and vulnerable people and their families.

We are grateful for the privilege of working with her and learning from her. We also thank Jayne Barker (2008–11), Cameron Keller (2011–12), and Catharine Hume (2012–2015), Mental Health Commission of Canada At Home/Chez Soi National Project Leads, as well as the National Research Team, the five site research teams, the site coordinators, the numerous service and housing providers, and people with lived experience who contributed to the original At Home/Chez Soi research demonstration project. We would, most especially, like to acknowledge the contributions of At Home/Chez Soi participants, whose willingness to share their lives, and open their homes for quality assessments were central and essential to the project.

The original research was made possible through a financial contribution from Health Canada. The views expressed herein are solely those of the authors.

## REFERENCES

1. Adair CE, Kopp B, Lavoie J, Distasio J, Hwang SW, Watson A, Veldhuizen S, Chislett K, Voronka J, Ahmad M, Ahmed N, Goering P. Development and initial validation of the Observer-Rated Housing Quality Scale (OHQS) in a multisite trial of Housing First. *J Urban Health*. 2014; 91(2): 242–55.

2. Evans GW. The built environment and mental health. *J Urban Health*. 2003; 80(4): 536–55.
3. Hwang SW, Martin RE, Tolomiczenko GS, et al. The relationship between housing conditions and health status of rooming house residents in Toronto. *Can J Public Health*. 2003; 94(6): 436–40.
4. Suglia SF, Duarte CS, Sandel MT. Housing quality, housing instability and maternal mental health. *J Urban Health*. 2011; 88(6): 1105–16.
5. Harkness J, Newman SJ, Salkever D. The cost-effectiveness of independent housing for the chronically mentally ill: do housing and neighborhood features matter? *Health Serv Res*. 2004; 39(5): 1341–60.
6. Wells NM, Harris JD. Housing quality, psychological distress, and the mediating role of social withdrawal: a longitudinal study of low-income women. *J Environ Psychol*. 2007; 27: 69–78.
7. Lindberg RA, Shenassa ED, Acevedo-Garcia D, et al. Housing interventions at the neighborhood level and health: a review of the evidence. *J Public Health Manag Pract*. 2010;16(5) E-Supp: S44–52.
8. Toomey TL, Erickson DJ, Carlin BP, et al. The association between density of alcohol establishments and violent crime within urban neighborhoods. *Alcohol Clin Exp Res*. 2012. doi:10.1111/j.1530-0277.2012.01753.x.
9. Guite HF, Clark C, Ackrill G. The impact of the physical and urban environment on mental well-being. *Public Health*. 2006; 120(12): 1117–26.
10. Matheson FI, Moineddin R, Dunn JR, et al. Urban neighborhoods, chronic stress, gender and depression. *Soc Sci Med*. 2006; 63: 2604–16.
11. Ludwig J, Duncan GJ, Gennetian LA, et al. Neighborhood effects on the long-term well-being of low-income adults. *Science*. 2012; 337: 1505.
12. Wright PA, Kloos B. Housing environment and mental health outcomes: a levels of analysis perspective. *J Environ Psychol*. 2007; 27(1): 79–89.
13. Keall M, Baker MG, Howden-Chapman P, et al. Assessing housing quality and its impact on health, safety and sustainability. *J Epidemiol Community Health*. 2010; 64: 765–71.
14. Tsemberis S, Gulcur L, Nakae M. Housing first, consumer choice, and harm reduction for homeless individuals with a dual diagnosis. *AJPH*. 2004; 94(4): 651–6. doi:10.2105/AJPH.94.4.651.
15. Goering PN, Streiner DL, Adair CE, et al. The At Home/Chez Soi trial protocol: a pragmatic, multi-site, randomized controlled trial of a Housing First intervention for homeless individuals with mental illness in five Canadian cities. *BMJ Open*. 2011; 1: e000323. doi:10.1136/bmjopen-2011-00323.
16. Stergiopoulos V, Hwang SW, Gozdzik A, Nisenbaum R, Latimer E, Rabouin D, et al. Effect of scattered-site housing using rent supplements and intensive case management on housing stability among homeless adults with mental illness: a randomized controlled trial. *JAMA*. 2015; 313(9): 905–15. doi:10.1001/jama.2015.1163.
17. Aubry T, Goering P, Veldhuizen S, et al. A randomized controlled trial in five Canadian cities of the effectiveness of Housing First with assertive community treatment for persons with serious mental illness and a history of homelessness. *Psychiatr Serv*. 2016; 67: 275–281.
18. Carr LJ, Dunsiger SI, Marcus BH. Validation of Walk Score for estimating access to walkable amenities. *Br J Sports Med*. 2011; 45(14): 1144–8.
19. Duncan DT, Alstadt J, Whalen J, Melly SJ, Gortmaker SL. Validation of Walk Score for estimating neighborhood walkability: an analysis of four US metropolitan areas. *Int J Environ Res Public Health*. 2011; 8(11): 4160-4179.
20. Babyak C, Alavi A, Collins K, Halladay A, Tapper D. The Methodology of the Police-Reported Crime Severity Index. Working Paper #HSMD_2009-006E/F; Ottawa: Statistics Canada; September 2009.

21. Tsemberis S, Rogers ES, Rodis E, et al. Housing satisfaction for persons with psychiatric disabilities. *J Community Psychol*. 2003; 31(6): 581–90.
22. Toro PA, Bellavia CW, Daeschler CV, et al. Evaluating an intervention for homeless persons: results of a field experiment. *J Consul Clin Psychol*. 1997; 65(3): 476–84.
23. Tsemberis S, McHugo G, Williams V, et al. Measuring homelessness and residential stability: the residential time-line follow-back inventory. *J Commun Psychol*. 2007; 35: 29–42.
24. Somers JM, Patterson ML, Moniruzzaman A, Currie L, Rezansoff SN, Palepu A, et al. Vancouver at home: pragmatic randomized trials investigating Housing First for people who are homeless and mentally Ill. Trials. *Trials*. 2013; 14: 365. doi:10.1186/1745-6215-14-365. http://www.trialsjournal.com/content/14/1/365.
25. Somers JM, Moniruzzaman A, Rezansoff SN. Urban migration in a cohort of mentally ill homeless adults: a fourteen-year retrospective study. *BMJ Open*. 2016; 6: e009043. doi:10.1136/bmjopen-2015–009043. http://bmjopen.bmj.com/content/6/1/e009043.full.