Response shift in quality of life ratings in homeless individuals with mental illness: a residuals analysis of the At Home / Chez Soi study

Guido Powell

Department of Epidemiology, Biostatistics, and Occupational Health

McGill University

June 2015

A thesis submitted to McGill University in partial fulfillment of the requirements of the degree of Masters of Science, Epidemiology

Table of Contents

List of Figures and Tables

Figures

Tables

Abstract

Background: The At Home/Chez Soi project was a pan-Canadian randomized controlled trial of a Housing First intervention among 2,148 homeless individuals with mental illness. The trial provided subsidized private-market apartments with recovery-oriented support services to treatment (HF) participants, while Treatment As Usual (TAU) controls were left to obtain services available in the community. Although the HF group reported significant improvements in quality of life (QOL) compared to controls, greater improvements were expected given observed group contrasts in both housing stability and a subsample's qualitative interviews. Recognized in the literature as a bias affecting self-reported QOL measurements, response shift occurs when a life event (e.g., an illness or a treatment) changes the meaning of an individual's QOL rating. Their internal standards may change or different values defining QOL may be reprioritized or redefined, affecting the accuracy of longitudinal quality of life comparisons, such as those of the At Home/Chez Soi project.

Following reviews of the literature on QOL among individuals with mental illness and among the homeless, as well as a review of theories of response shift and methods for its identification, the current thesis seeks to adjust for response shift in comparisons of QOL in the At Home/Chez Soi participants.

Methods: A secondary analysis method developed in previous literature allows for the identification of response shift using the residuals from an explanatory model of the QOL outcome. Based on a random intercept model explaining variability in the 7-point global QOL item (of the 20-item QOL index [QOLI-20]), participants' degree of response shift was defined by the standard deviation (*sd*) of their respective residual values. These *sd* values were accounted for in group comparisons of both the QOLI-20 total score and global item outcomes, allowing for estimates of the intervention's effect on QOL to vary according to degree of participants' response shift. A sensitivity analysis reassessed findings after excluding observations for which interviewers felt responses were invalid or insincere, under two levels of exclusion criteria (A and B).

Results: A model explaining 62% of QOL variability was estimated using 23 covariates. Participants' mean *sd* of residuals was 0.95 (min: 0.00, max: 3.93). HF treatments effects on QOLI-20 total scores diminished when estimated at higher levels of response shift (1 point increments in *sd* of residuals), for most follow-up periods (interaction at 6 mo. β = -3.97, p = .052; at 12 mo. β = -3.20, p = .109; 18 mo. β = -4.15, p = .043) but not at 24 months (β = 0.14, p = .946). A non-significant decrease was seen in the HF treatment odds ratio (OR) from an ordinal logistic regression of the QOLI-20 global item, collapsed across follow-up periods (interaction β = 0.83, p = .175). Following sensitivity analyses, the HF treatment x *sd* of residuals interactions no longer met significance levels as exclusion criteria were applied to the QOLI-20 total score analyses. However, the interaction term of the QOLI-20 global item analysis became significant under the more stringent exclusion criterion B (β = 0.69, p = 0.027), while criterion A had little impact on the original estimate.

Conclusion: These findings suggest that unexpectedly modest effects of the HF intervention on QOL may be explained by response shift, a novel observation in the homelessness literature. A more conservative conclusion from the results is that QOL improvements from a HF intervention are more evident in individuals for whom reported and expected QOL differ by an amount that is relatively consistent over time. HF treatment effects are smaller in participants whose residual patterns fluctuate over time. This fluctuation is interpreted as response shift, in accordance with previous literature, though alternative interpretations may also explain some of these patterns.

Résumé

Contexte : Le projet Chez Soi est un essai randomisé contrôlé pancanadien d'une intervention Logement d'abord, auprès de 2148 personnes itinérantes ayant un trouble de santé mentale, offrant des appartements subventionnés sur le marché privé et des services de soutien à un groupe d'intervention (HF), tandis que le groupe témoin (TAU) ont obtenu les services habituels. Bien que le groupe HF a rapporté des améliorations significatives en termes de qualité de vie (QDV) par rapport aux témoins, un plus grand effet était attendu étant donné les différences de stabilité résidentielle ainsi que des résultats d'entrevues qualitatives. Reconnu dans la littérature comme étant un biais qui affecte les mesures de QDV, le changement de position se produit quand un événement change le sens de l'évaluation de QDV d'un individu. Leurs normes internes peuvent changer ou les valeurs qui définissent la QDV peuvent prendre de nouvelles priorités ou être redéfinies. Cela affecte la précision des comparaisons de mesures de QDV à travers le temps, tel que dans le projet Chez Soi.

À la suite de revues de littérature sur la QDV parmi les personnes ayant un trouble de santé mentale et sur le changement de position, cette thèse cherche à ajuster les comparaisons de QDV en tenant compte des effets du changement de position dans les participants du projet Chez Soi.

Méthodes: Une méthode d'analyse secondaire permet d'identifier le changement de position par les valeurs résiduelles d'un modèle de régression des mesures de QDV. À partir d'un modèle à l'ordonnée à l'origine aléatoire expliquant la variabilité d'un item de QDV globale de 7 points (de l'index de qualité de vie à 20 items [QOLI-20]), le degré de modification de réponse pour chaque participant est identifié par l'écart-type (*sd*) de leurs valeurs résiduelles. Ces valeurs *sd* sont intégrées dans les comparaisons entre groupes des scores totaux du QOLI-20 et des résultats de l'item global, permettant les estimations de l'effet de l'intervention sur la QDV de varier selon le degré de modification de réponse. Les résultats ont été révisés sous une analyse de sensibilité excluant les observations estimés comme invalides ou malhonnête, selon deux niveaux de critères d'exclusion (A et B).

Résultats: Un modèle expliquant 62% de la variabilité de QDV a été estimé avec 23 variables. La moyenne de l'écart-type de valeurs résiduelles des participants était de 0.95 (minimum : 0.00, maximum : 3.93). L'effet de l'intervention sur les scores totaux du QOLI-20 diminue avec la modification de réponse (augmentations de 1 écart-type des valeurs résiduelles), pour la plupart des suivi (l'interaction à 6 mois β = -3.97, p = .052; à 12 mois β = -3.20, p = .109 ; et à 18 mois β = -4.15, p = .043) sauf à 24 mois (β = 0.14, p = .946). Une diminution non-significative a été estimée pour le rapport de cotes de l'effet de l'intervention dans une régression logistique ordinale de l'item global du QOLI-20, couvrant tous les suivis (interaction β = 0.83, p = .175). À la suite des analyses de sensibilité appliquant les critères d'exclusion, l'interaction des effets de l'intervention et de l'écart-type des valeurs résiduelles n'était plus significative dans l'analyse des scores totaux du QOLI-20. Par contre, l'interaction pour l'item global du QOLI-20 est devenu significative sous le critère d'exclusion plus rigoureuse (B), tandis que les estimés n'ont pas changé sous le critère A.

Conclusions : Ces résultats suggèrent que les effets modestes de l'intervention sur la QDV s'expliquent possiblement par le changement de position, une observation originale dans la littérature sur l'itinérance. Selon une interprétation plus prudente, les résultats suggèrent que de plus grandes améliorations de QDV sont perçues chez les participants pour qui la QDV résiduelle est constante à travers le temps, tandis qu'une fluctuation des valeurs résiduelles pourrait être interprétée autrement que le changement de position.

Acknowledgements

I would first and foremost like to thank my supervisors, Dr. Eric Latimer and Dr. Nancy Mayo, for their guidance and support throughout the entirety of this project. I have spent the past few years taking steps through the doors Dr. Latimer has opened for me, happily receiving his pat on the back along the way. I have also been lucky to benefit from Dr. Mayo's enthusiasm and endless knowledge. Both Dr. Latimer and Dr. Mayo developed the blueprint and foundation of this research as a result of ir collaborative expertise and creativity. The undertaking of the thesis also could not have of been possible without their diligence and wisdom.

I must also acknowledge the great assistance I have received from members of both my supervisors' research labs. I would like to thank Susan Scott for help with the statistical analyses and her admirable receptiveness. I am also grateful for the insight, methodological help, and general support from Erika Braithwaite, Zhirong Cao, Ashley Lemieux, Christian Méthot, and Daniel Rabouin.

Finally, I could not have completed this thesis without the constant support of my greatest ally, Ruth Jacobs. Thank you for being wonderful and patient, but not too patient.

Author's Contribution

The following thesis was written by myself, Guido Powell, with feedback from Dr. Latimer and Dr. Mayo. The literature reviews and design of the study's analyses were also developed with guidance from my supervisors.

Introduction

A common measure of the success of an intervention for helping homeless people suffering from mental illness is their quality of life (QOL) ratings [1–4]. However, true change in QOL can be difficult to track accurately. Important life events, such as an illness or even an intervention, can affect the way respondents understand the very construct of QOL. Within one individual, the internal standards, values, and meaning of QOL can change over the course a study, leading to a phenomenon known as "response shift"[5–7]. Response shift frequently explains paradoxical findings in medicine, such as unexpectedly high QOL ratings that conflict with deteriorating health[8]. Fortunately, advanced statistical tools have been developed to identify response shift through secondary analyses of existing data[9–11]. Using data from the At Home/Chez Soi Project in Canada, the current study is the first to examine the impact of response shift on QOL ratings in a housing intervention for homeless individuals with mental illness. The analysis sorts participants according to the extent of their identified response shift, allowing for original QOL outcome analyses to be adjusted for this measure.

The first chapter of the introduction provides a brief overview of homelessness in Canada, the Housing First model of housing interventions and its implementation under the At Home/Chez Soi project. The second chapter is an in depth review of QOL in terms of its measurement and associations among individuals with mental illness, its responsiveness to treatment, with comparisons of these findings to those among a homeless population. The last section of the introduction reviews the concept of response shift, both in terms of the theory of the phenomenon, and how it is identified in the literature. The subsequent methods and results sections each reflect a three-step process consisting of; 1) a model explaining QOL ratings, 2) a calculation of a response shift metric for each participant, and 3) two different QOL outcome analyses adjusted for response shift (on a global QOL item rating and the

total score of the same instrument). Validation of the QOL measure as well as a sensitivity analysis are presented in the results.

Finally, the discussion examines the relevance, strengths, and limitation of both the response-shift-adjusted QOL outcomes and the findings of the explanatory model.

CHAPTER 1: Homelessness

Given its complexity and frequent misrepresentations, an accurate portrayal of homelessness is warranted. Some aspects of homelessness, however, can only be described through approximations. For instance, estimates of the size of homeless populations vary by city and by estimation methodology. A frequently cited figure for the nation-wide homeless population in Canada sets the yearly number of homeless individuals as high as 235,000, with a daily prevalence of around 35,000 Canadians [12]. Possibly more accurate figures can be obtained from city-specific homeless counts, enumerating people living in the streets and shelters. For example, a 2013 report by the City of Toronto identified 5,253 individuals who were on the streets and shelters in one night[13], while a Vancouver street count of the same year counted 1,600 sheltered and unsheltered homeless individuals[14]. In 2005, the city of Winnipeg counted 350 individuals living on the streets and 125 using emergency shelters, of which a severely disproportionate amount identified as Aboriginal (62%)[15]. In Moncton, over the course the year 2013, 781 unique individuals were identified as users of emergency shelters, without any figures for number living on the streets[16]. (The results of a 2015 street count in Montreal are yet to be announced).

However, such prevalence rates ignore an important distinction between the different types of homelessness. Homelessness, in reality, fits along a continuum where at one extreme are the *absolutely homeless* individuals who live in the street, places that are not developed for housing (e.g., abandoned or unguarded infrastructure) or who make use of emergency shelters[17]. Less visible are the *hidden homeless*, a category in the middle of the continuum consisting of individuals without a home who stay temporarily with friends and family (i.e. “couch-surfing”) or in their own cars. Similarly, another category of homeless individuals are described as “provisionally accommodated”, or those who are in settings such as institutional care or transitional housing with guarantee of any permanent tenure. At the other end of this spectrum are individuals experiencing *relative homelessness*. These individuals may be housed but facing eviction, living in substandard, unsafe housing, or paying a dangerous portion of their

income on rent. In addition to the degree of homelessness, distinctions can be made by the duration or frequency of the experience. Homelessness can be *chronic* (long-term and frequent), *cyclical* (occasional [e.g., following a hospitalization]), or *temporary* (brief and non-recurring [e.g. following a catastrophe])[12]. This range in definitions has broad effects on policy and research. For example, it is estimated that for every homeless individual living in the street, there are up to 4 individuals experiencing *hidden* homelessness[18]. Thus, any concern about a visibly growing problem of homelessness is amplified by rates that are frequently undetected.

Considering common myths about the homeless, it is important to emphasize that homelessness is not a choice. When asked whether they wished to be permanently housed, the vast majority (93%) in Toronto's street count stated this was what they desired[13]. Given this overwhelming intention to be housed, the persistence of homelessness can only be understood by examining its causes.

The paths to homelessness are described as an interplay of individual/relational factors, system failures, and structural factors. Among many individual factors (childhood abuse, recent traumatic events), most commonly acknowledged are challenges with mental health or substance use[17]. Short of fully capturing the complexity of mental illness in the homeless, some estimates are nonetheless informative of this critical reality. Of 300 individuals experiencing homelessness for the first time in Toronto [19], 67% had a diagnosis of mental illness in their lifetime. While other Toronto homeless studies found surprisingly low lifetime prevalence of schizophrenia (6%), a higher prevalence has been reported for affective disorders (between 20%-40%)[20]. Lifetime alcohol and substance abuse disorders were equally frequent, at a rate of 68% based on the study of the first-time homeless in Toronto [19]. It is important to note that mental illness and substance problems can both be causes as a well as consequences of homelessness. Frequently, the experience of homelessness can trigger or aggravate

mental health and substance problems (as well as causing problems with nutrition, sleep, and a variety of health problems that can drastically increase risk of death)[20].

The second category of causes, system failures, is equally important in explaining homelessness[17]. One frequent path to homelessness involves care services' poor planning for individuals being discharged from hospitals and psychiatric or addiction recovery institutions. This failure follows a history of deinstitutionalization since the 1950s in North America, where many incentives led to the downsizing or closure of psychiatric institutions[21]. While successful in cutting down long-term psychiatric hospitalizations, the process of deinstitutionalization often failed to complement discharges with community treatment or housing. This trend contributed to a rise over several decades in individuals with mental illness experiencing different degrees of homelessness, from relative to absolute[21].

Closely linked to system failures are structural factors, the third category of causes of homelessness. Examples of these factors are discriminatory social and economic environments or policies that affect poverty levels and affordable housing[17]. Federal policies have contributed to housing access problems, with per capita investments in social and affordable housing decreasing significantly since the 1980s. At the same time economic factors have encouraged the supply of privately owned condominiums and homes at the expense of more accessible rental units[20].

Many of these factors, along with the significant national rates of homelessness populations, created a context in which government agencies were motivated to take action to end homelessness. With funding from the federal government, the Mental Health Commission of Canada (MHCC) undertook the At Home/Chez Soi project. The project was a multi-site randomized controlled trial of Housing First that took place across Vancouver, Winnipeg, Toronto, Montreal, and Moncton between 2009 and 2013.

1.1 Housing First and the At Home/Chez Soi Project

Housing First (HF) is an approach to help homeless individuals with mental illness by offering immediate access to permanent housing of their choice, through rental supplements, along with clinical support services as needed. The approach was popularized by Sam Tsemberis whose version of the philosophy was implemented in New York City by *Pathways to Housing* in the 1990s[22].

There are five core principles of HF that underlie the approach [22]. These principles distinguish HF from mainstream "continuum of care" models, in which access housing with onsite support is offered only to individuals who meet certain criteria. First, clients are provided with immediate permanent housing without any condition of readiness tied to mental health, functioning, or sobriety. Second, HF is guided by a principle of self-determination such that clients choose the location and type of housing they are offered (scattered site or congregate) and also choose the clinical services they receive. Third, all clinical services provided are "recovery-oriented" in that they seek to help clients beyond managing symptoms and meeting basic needs. Recovery from mental illness can be defined in different ways, but individual well-being is always a focus. Recovery from substance problems under HF allows for a "harm reduction" approach where abstinence is not enforced. Fourth, HF services are customized to meet the needs of the clients, both in terms of level and frequency of clinical supports as well as the amount provided as rental supplements (clients pay no more than 30% of their income on housing). Finally, the fifth core principle of HF is its focus on social and community integration. Clients are encouraged to engage in meaningful activity in their communities, be they social, vocational, occupational, or recreational. Similarly, scattered site housing with *offsite* supports (as is the case for most clients) is said to facilitate better integration in the community and prevent the stigmatization created by other housing models[22].

These five principles directed the implementation of the At Home/Chez Soi intervention[23]. In comparison, participants randomized to the "treatment as usual" (TAU) control group were left to

obtain services offered in the community. HF participants assessed as having "high" needs were provided with Assertive Community Treatment while those with "moderate" needs were offered Intensive Case Management.

Evaluations of the intervention from the study's cross-site final report[24] demonstrated that Housing First was a feasible and effective method to reduce homelessness, requiring relatively minor investments from society (given economic offsets in other services used). Two different papers report the results of the study for the High Needs group [25] and the Moderate Needs group [26]. In both studies, the primary outcome of the trial, housing stability was much higher for HF than TAU participants. Over the course of 24 month following baseline, HF groups in the high needs arm were housed 71% of the time compared to 29% for TAU participants. For the Moderate Needs arm, HF participants were housed between 63% and 77% of the time (depending on site) while TAU participants were housed between 24% and 39% of the time. For Moderate Needs groups, the HF group also had a lower percent of participants who remained absolutely homeless (5%) compared to the TAU group (31.5%).

Positive impacts of immediate housing were apparent in glimpses provided by qualitative studies. Early on in the trial, a small subsample of 27 participants in the HF group participated in narrative interviews, one month after being housed but prior to their first visit with the clinical team[27]. Analyses revealed that for most of these participants, immediate housing instilled hope about the future. Reclaiming a positive identity was also one of the major themes that emerged from these data. Both these themes are considered important elements for recovery from mental illness. However, for a small but important minority, interviews revealed themes of demoralization and concerns about social isolation (19% and 15% respectively)[27].

Another subsample of 197 participants in both groups provided insights into the effects of the intervention after 18 months[24]. Experiences from narrative interviews were classified as either positive, negative, or mixed/neutral life courses. Participants in the HF group were over twice as likely to have

experienced a positive life course, whereas TAU participants were over four times as likely to have a negative life course (similar rates of mixed/neutral life courses were reported in each group). Housing, social contacts, reduced substance use, and hope were all important factors related to positive life courses. A theme of "wellbeing" emerged from interviews, where participants with positive outcomes reported themselves as being "more secure", "peaceful", and "in a better place" [24].

These positive findings from the subsample's qualitative measures raised the question as to why the whole sample's *quantitative* outcomes showed more modest treatment effects than expected. Surprisingly, despite being more successfully housed and treated, participants from the HF group did not differ significantly from the TAU group over time in terms of their substance use (Global Assessment of Individual Need scale) or mental health (Colorado Symptoms Index). These measures similarly improved for both groups from the baseline period onward.

Two other measures highlighted in the final report did distinguish the two groups significantly, but not with the large effects that were expected. Community functioning, as measured by the Multnomah Community Assessment Scale (MCAS), was one outcome for which HF participants showed modest but significantly more improvement than TAU participants. This difference was most evident in terms of items measuring behaviour such as impulsivity and medication compliance. For other domains of the MCAS, improvements were seen for both groups, thus showing little treatment effect (e.g. health impairment, social network, etc.).

The other important outcome on which the treatment effect was surprisingly modest, though statistically significant, was quality of life (QOL)[24].

CHAPTER 2: Quality of Life

The At Home/Chez Soi study based measurements of QOL on a 20-item version Lehman's Quality of Life Index [28,29] (see Appendix A for instrument details). The QOLI-20 measures self-reported QOL on several domains, such as everyday activities, leisure, social relationships, finances, and safety. An additional item assesses the participant's *global* QOL by asking, "How do you feel about your life (as a whole)?" All items are scored as either 1 ("Terrible"), 2 ("Unhappy"), 3 ("Mostly dissatisfied"), 4 ("Mixed"), 5 ("Mostly Satisfied"), 6 ("Pleased"), or 7 ("Delighted"), with each score indexed by face illustrations ranging between a frown (1), a neutral face (4) and a smile (7).

Over the course of the study, the two groups did not differ as greatly as expected in terms of QOL. The cross-site final report of the At Home / Chez Soi study reported an increase over time in the total score of the QOLI-20 for both groups (see Appendix B for QOL outcome analyses). A significantly higher average score for the HF group was seen early on at 6 and 12 months after baseline. This effect seemingly tapered off as both groups showed relatively similar QOLI-20 total scores at 18 and 24 months from baseline[24].

In comparison to total QOL scores, greater treatment effects were seen in domains of QOL. Treatment was linked to greater improvements in QOL in the domains of living situation, and to lesser extents domains of perceived safety and finances. Satisfaction with social and family relationships improved equally in both HF and TAU groups (no treatment effect). Furthermore, comparing the effects of treatment on QOL for different need level groups revealed that Moderate Need groups showed larger treatment effects than did the High Need groups[25]. These effects were seen especially for the domains of leisure activity, living situation, and perceived safety [24], while improvements in family, finances, social domains and overall QOL were not significantly greater for the HF group [26].

2.2 Measuring QOL in individuals with mental illness

Regardless of some domain-specific differences, the modest treatment effect on total QOL scores are worth examining further, especially given the emphasis of QOL measures in mental health research. Clinicians and researchers in the field of mental health as well as other chronic illnesses encourage the use of self-reported QOL to assess the effectiveness of services provided[30]. One reason is that they capture a more complete picture of the subject's internal state than other indirect measures of QOL. In many instances, reported QOL scores may be very different from one individual to the next despite their similarities in more "objective" measures, such as hospitalizations or income. Similarly, the evaluation of a treatment solely based on symptoms of a chronic illness can be less reliable given the cyclical or persistent nature of many disorders, especially in mental health and substance abuse. Finally, measuring self-reported QOL is in accordance with the ultimate goals of important mental health organizations, like the Mental Health Commission of Canada [31], which have stated the improvement of QOL as an essential element of their mental health strategies[31].

Reviewing the literature on QOL in mental illness, one recurring issue worth clarifying is the diversity of concepts that are used interchangeably with QOL. A review by Evans[30] found that the QOL construct is frequently referred to as either wellbeing, happiness, life satisfaction, health-related QOL, health status, or functioning. While life satisfaction, happiness, and wellbeing are more closely related to the concept of QOL adopted in this study, terms such as health status, functioning, and health-related QOL diverge from this construct. Health status captures more of patients' perspectives on their physical or mental health, functioning captures an individual's independence, social relationships, activities, while health-related QOL measures specifically the impact of illness (mental or physical) on different aspects of QOL[30].

The broader concept of QOL is an expansion of the narrower, but often used health-related QOL and is argued to be a better measure for evaluating the impact of treatment [30]. The broadening of QOL

measurement to one that is multidimensional, beyond just the impact of an illness or an illness-specific intervention, captures the personal values and preferences that determine an individual's QOL ratings. Non-health domains of QOL may demonstrate otherwise unmeasured impacts of a treatment that address the many needs of patients outside mental or physical health, such as social and family relationships, living situation, or finances [28]. The many QOL measures developed for both general and mental illness populations often incorporate these various QOL domains as well as a global QOL measurement of overall life satisfaction[32]. Comparisons have shown agreement between researchers and participants with regard to the set of domains that are deemed relevant to QOL and worthy of measurement[32].

Evidently, there is a trade-off involved in the propagation of global, non-health related QOL measures applied across both healthy and mentally ill populations. One disadvantage is that fewer disease-specific outcomes are being measured in QOL assessments, while a benefit is the potential for QOL comparisons across these two groups[30].

Despite the promotion and growth of QOL measurement in individuals with mental illness, there remain concerns about the validity of QOL scales in this population. A review by Lehman [33] examines the psychometric properties of several such instruments. Among the 10 instruments reviewed, only 4 were found to have evidence of acceptable reliability and validity. Most relevant among these 4 is Lehman's original QOLI, which has 143 items and covers objective indicators of QOL on top of the life satisfaction scales included in the QOLI-20. Of interest for the current study, psychometric properties of the life satisfaction scales of the QOLI include a median internal consistency reliability of 0.85 (with a range across domains of 0.79 to 0.88)[33]. A subsample of 45 participants in one study provided a median one-week test-retest reliability of 0.72[34]. While the measure's reliability on the safety domain, at 0.41, was considerably low in this sample, the other 9 domains had a median reliability of 0.72 (with the general life satisfaction item's reliability at 0.71)[33]. The QOLI, both in terms of its life satisfaction scales and

objective indicators, has been evaluated as having good construct and predictive validity based on confirmatory factor analyses and multivariate predictive models. Different studies support the QOLI's predictive validity, given its differentiation of patients based on residential circumstance, clinical symptoms and mental health status, gender, and age [33]. In another review of 11 QOL instruments for mental health research, Van Nieuwenhuizen and Wolf[35] conclude that two instruments are best suited for measuring QOL in a mental illness population, Lehman's QOLI and the very closely related Lancashire Quality of Life Profile (LQOLP).

2.3 Determinants of QOL

Given the validation of the QOLI and similar instruments, multiple studies examine QOL among individuals with mental illness using these measures. Several predominant associations between QOL and other factors stand out in this literature (similar studies with a focus on homeless individuals are discussed further below).

Personal Characteristics: Demographics

Lehman originally developed a theoretical model of global QOL with personal characteristics as one of three proposed determinants, along with objective and subjective QOL indicators in different domains[28] (see Figure 1). His analyses however established that personal characteristics played only a minor role in determining QOL. Of the demographic characteristics, only being married and having *less* education showed a small, but positive association with global QOL in an early test of the model. Gender, age, and parental social class were not significantly related to global QOL (personal characteristics of psychiatric diagnosis are discussed further below) [28]. Partly corroborating these findings, Trauer and colleagues [36] found that of several demographic characteristics, only being married or in a de facto relationship (the status of 9 out of 55 participants) predicted higher QOL. In another

study of demographic correlates in the QOL of 1,805 individuals with mental illness, Lehman found some different modest associations [37]. Men reported slightly higher satisfaction with life in general and most life domains (daily activities, family, and safety) compared to women. He also found that non-white participants rated higher average life satisfaction with life in general and with family and social relationships, in comparison to white counterparts [37].

However, some inconsistencies exist in the literature regarding demographic characteristics as determinants of QOL of individuals with mental illness. In a brief literature review of QOL determinants in people with mental illness, Hansson [38] identifies only marital status and employment as important demographics contributors to higher QOL. One study included in the review found no link in a sample of 49 patients between QOL and either age or education [39]. However, older age was found to be associated with slightly higher global QOL in a Greek study of 54 psychiatric residents [40], as well as in a Swedish study of 288 substance abusers with mental illness [41]. This latter study also reported only a minor association of marital status [41]. Age was however related to *lower* global QOL in a review by Holloway and Carson [42]. Furthermore, a study of 162 Los Angeles board and care home residents by Mares and colleagues [43] reported no link between QOL and marital status and instead found that gender impacted QOL, with male participants reporting higher QOL. In a review of QOL studies in schizophrenia patients, Pinikahana and colleagues [44] found that age and gender were generally not important QOL determinants of QOL in this literature, with higher QOL being associated with younger age in only one study[45] and being female in one other[46]. Equivocal findings were reported on the importance of marital status in the studies reviewed [44]. These conflicting findings may be explained by slightly different mental health populations in every study.

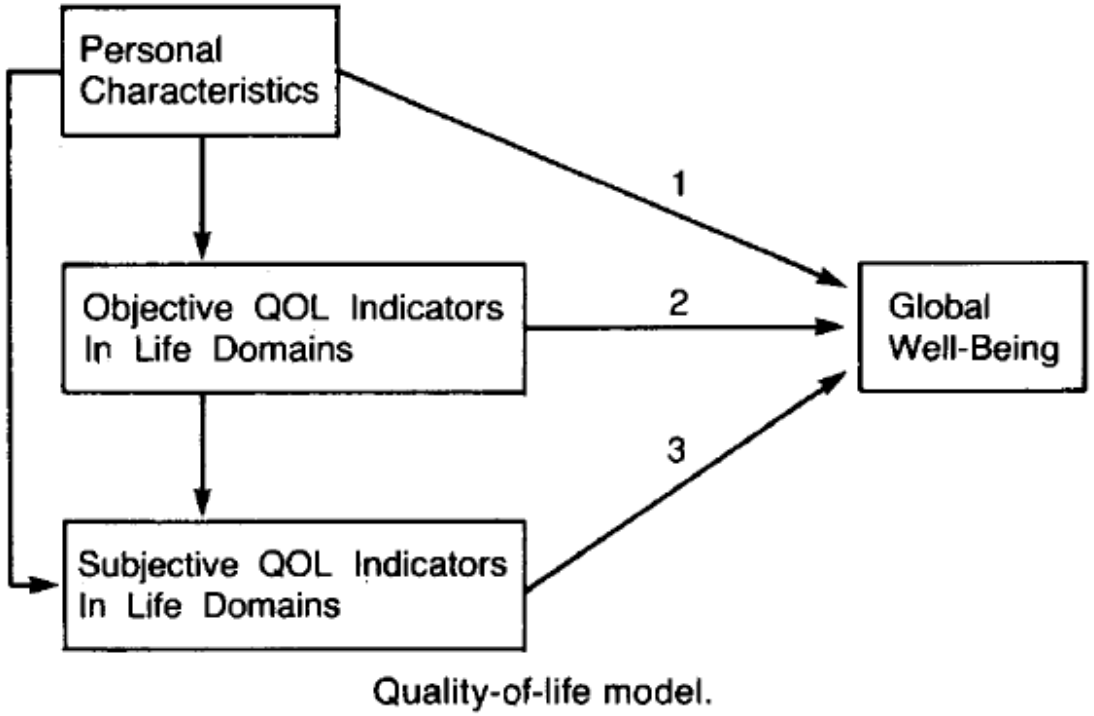


Figure 1 Lehman's 3-part theoretical model of global quality of life[28]

Personal Characteristics: Psychiatric symptoms

Although in Lehman's study of his theoretical model psychiatric diagnoses correlated as weakly as demographic characteristics with global QOL (approximately .14)[28], stronger links between psychiatric symptoms and QOL are often reported in the literature. In a study of 70 severely mentally ill patients, Holloway and Carson[47] identified both symptoms of depression and negative psychosis symptoms as the strongest predictors of global QOL. Kaiser and colleagues [48] also reported that among 440 patients in Germany and Wales, psychopathology (self-reported symptoms and assessed behaviour) consistently correlated with lower QOL in all domains. Similarly, a meta-analysis of the impact of psychiatric symptoms on QOL in schizophrenia patients [49] found that symptoms of depression and anxiety were stronger predictors of lower QOL than were positive or negative psychosis symptoms (though negative and positive symptoms did predict lower QOL specifically in outpatient settings). One reviewed study of

62 Canadian schizophrenia patients identified the severity of symptoms and side effects of antipsychotic drugs (akathisia and neuroleptic dysphoria) accounted for nearly half of global QOL variance [50].

These associations are also reflected in baseline QOL differences between diagnostic groups seen in the Swedish study of psychiatric and social services by Schaar and Ojehagen [41]. Among four diagnostic groups, the psychosis group reported higher average QOL, followed by a group representing an assortment of other diagnoses, while individuals with depression and borderline personality disorder reported the lowest QOL. In this study [41] and many similar studies [39,50–52], a greater number of symptoms and their greater severity are consistently reported as important determinants of lower QOL ratings. Similar patterns have also been identified in a synthesis of qualitative research on the topic [53].

Given this association of QOL and psychiatric symptoms, some authors express concerns that QOL measurements in individuals with mental illness achieve little more than assessing their mental health (mood state, anxiety, etc.). One study concludes that QOL ratings in several psychiatric groups has questionable validity, based on findings of incompatible QOL ratings and objective indicators of QOL in 69 patients with schizophrenia in Calgary [54]. It is argued that poor insight and affective bias may excessively influence QOL ratings, irrespective of objective life circumstance (e.g. adverse life events). Other authors argue that this conclusion ignores a similar disparity between objective life circumstances and subjective QOL in the general population[38]. Additionally, in an early validation study of the original QOLI by Lehman [55], analyses demonstrated that the associations of personal characteristics, objective QOL indicators, and domain-specific subjective QOL with general QOL ratings were unaffected by psychopathology. Coefficients of the three predictors were not significantly different across regression analyses that either included or excluded a combined measure of depression, anxiety and self-control (of thoughts, behaviour, and emotions). Similarly, in examining the reliability and validity of a QOL instrument derived from Lehman's QOLI, the Lancashire Quality of Life Profile (LQOLP), Oliver and colleagues [56] concluded that QOL measures among individuals with mental illness can indeed be

distinguished from findings on other measures. The authors also state that although mental health does not alter QOL findings, one should still control for the effect of psychiatric symptoms in QOL measurement.

Objective QOL Indicators in Life Domains

The second component of the Lehman's theoretical QOL model, objective indicators of QOL, proved only somewhat better than demographic variables at explaining global QOL [28]. The indicators included items from the QOLI measure that assessed participants' experiences in the domains of living situation, family, social relations, leisure, work, law-safety, finances, and health. Indicators in the domain of work had the highest bivariate correlations with global QOL (up to .47), while other domains showed lower correlations (generally between .1 and .2). In a multivariate model with personal characteristics, objective indicators that did predict *higher* QOL were less use of medical services, not being the victim of a crime, more frequent and more intimate social relationships, number of hours worked, seeking work, and privacy. Along with personal characteristics, these variables explained a modest portion of global QOL variance (R^2=.23)[28].

The study by Trauer and colleagues [36] examining Lehman's QOLI similarly found that objective QOL indicators again showed small correlations (maximum of -.37) with global QOL. Less service use, fewer problems with access to services, not being the victim of a crime, functional community living skills, and more social contacts correlated with greater global QOL. However, despite mostly adverse objective indicators of QOL (social isolation, poverty, and unemployment), individuals with mental illness rated their QOL relatively high between 4.2 and 5.0 (out of 7) in all domains[36].

Similarly, in QOL research from Greece by Zissi and colleagues [40] using Lehman's original QOLI measure, many of the same objective life indicators (family, social contacts, leisure, physical health, safety) were once again weakly associated with global QOL. A multivariate model containing only

objective indicators explained only 9% of global QOL variance. The study did report, however, that some of these objective indicators were indirectly related to global QOL through their substantial impact on self-concept and perceived change[40].

An important objective indicator of QOL found in several studies was that of social relationships. There are examples in the literature of the strong influence of social network size on global QOL, both in terms of bivariate correlations (.48) [39] and QOL variance explained from multivariate models with network size as strongest predictors (R^2 = .41) [57]. A complementary perspective in a study of 73 psychiatric patients by Aubry and colleagues [58] demonstrates that global QOL was positively associated with features of an individual's housing environment that indicate greater "social role valorization" (e.g., encouraging integration, participants' choices, respectful interactions, etc.). This association was however mediated by psychological integration (sense of belonging, emotional investment) in participants' neighbourhoods[58].

Subjective QOL Indicators in Life Domains

The final component in Lehman's theoretical model, subjective indicators of satisfaction in different domains, proved to be the most reliably strong correlate of global QOL[28]. Global QOL correlated with satisfaction with life domains of living situation (.45), family (.35), social relations (.58), leisure (.59), finances (.40), law-safety (.42), unemployment (.33), and health (.66). The few employed participants (42 of 278) showed a low correlation between life satisfaction with work and global QOL (.17). The addition of subjective QOL indicators doubled the explained global QOL variance of a model including only personal characteristics and objective indicators (from R^2=.23 to R^2=.58)[28]. As discussed above, this improvement of model fit from subjective indicators was not affected by adjustments for psychiatric symptoms in a subsequent study by Lehman [55]. The study by Trauer and colleagues [36]

replicated most of the bivariate correlations between global QOL and life satisfaction in QOL domains (within a range of .01 to .09), except for a much higher correlation in domain of leisure (.76).

Intuitively, the primacy of subjective indicators of QOL in explaining global QOL can be understood by the fact that both are assessed from equivalent life satisfaction scales. However, these associations persist even when subjective indicators are assessed through different types of questions. The Los Angeles board and care resident study found strong associations between perceptions of their social living environment (cohesion, conflict, etc.) and their global or living domain QOL[43]. Another example is given by results of the UK700 group's finding that the strongest predictor of QOL in 708 individuals with mental illness is a sense of having serious problems with meeting basic, social and functioning needs, explaining 20% of QOL variance [59]. This finding is corroborated by those from a study 576 patients by Skinner and colleagues [60] where important predictors of lower QOL were unmet needs with management of money, health, and family concerns.

2.4 Responsiveness of QOL to change

The associations described so far mostly reflect predictors of QOL from cross-sectional analyses. Of greater interest to the current study is understanding how QOL can change over time, especially following treatment.

Concerns have however been expressed in the literature about the responsiveness to change of self-reported global QOL ratings. For example, no QOL change (global or domain-specific) was reported over 6 months in a study of 260 patients in 4 UK cities examining community mental health services for various mental illnesses [61]. In another British study of mental health services, both groups (*n* = 138) of schizophrenia patients receiving intensive or standard case management showed no QOL improvements over the course of 2 years, either in terms of global QOL or average domain-specific QOL[62]. Both these studies with null findings assessed QOL using the LQOLP, a similar measure as the QOLI.

Other studies illustrate that only modest QOL change can sometimes result from treatments. For example, Barry and Crosby[63] found that over the course of a year, 29 former psychiatric inpatients in Wales resettled in community residences with mental health services showed improvements only in the living situation domain of QOL. In an Italian study of community mental health services among 261 patients, any significant global and cross-domain QOL improvements over a period of 6 years (also assessed through the LQOLP) were surprisingly small [64].

Nonetheless, several studies do establish that measures of QOL are responsive to change from treatment. Huxley and colleagues [65] reported 2-year QOL outcomes from the UK700 multi-site randomized trial of intensive case management (with controls received standard case management) using the LQOLP. They observed significant improvement in global QOL in 3 out 4 sites, while improvements in domain-specific measures of QOL were seen in one site for all domains, in 5 out of 8 domains for another site, and in only one domain for the two other sites. A study by Urbanoski and colleagues [52] of 133 individuals with mental illness in Ontario (55% of which had substance abuse problems) found that intensive case management and assertive community treatment improved global QOL as well as QOL in the domains of living conditions, safety, and daily activities. Finally, in the Swedish study examining the effects of cooperation between psychiatric and social services[41], participants with mental illness and substance use problems showed significant improvement in QOL over 18 months.

Mediational Model of QOL Change

In her review, Evans [30] discusses the different range of QOL outcomes following treatments and proposes that some studies with null findings may have unwarranted expectations for positive results. One assumption made in many studies is that QOL ratings will improve following treatment irrespective of improvements in other measures. This assumption ignores frequent findings that QOL change is generally observed in unison with change in other factors like symptoms, self-related constructs (e.g.,

self-esteem), and to some extent objective indicators of QOL. Examples discussed above include the mediation between change in objective QOL indicators and change in global QOL through self-concept in Zissi and colleagues [40] or through psychological integration in Aubry and colleagues [58]. Likewise, perception and beliefs about objective circumstances often plays a key role in mediating global QOL differences. Perceived family criticism [51], beliefs about social rejection [66], and the value of activities participated in [67] were identified in different studies as predictors of QOL, regardless of actual family contacts, rejection, or hours of activity, respectively. Additionally, Rosenfield [68] identifies a mediating role of a perceived sense of mastery (or control) in enabling services that improve economic resources and decision-making power to actually improve QOL in study participants. These findings are consistent with those of Boyd and Bentley [69], who examines the role of empowerment in explaining differences in QOL between groups of mental health consumers. Similarly, Markowitz [70] describes a reciprocal model representing the relationship between changes in psychiatric symptoms, self-esteem and QOL. These patterns all point to the limitation of a three-part theoretical model of QOL as defined by Lehman [28], in which each component directly predicts global QOL, without necessary mediation. It is clear that some components, such as certain objective indicators, may only be linked to QOL through an individuals' perceptions or appraisal.

The evidence of this type of mediation is precisely what motivated authors Zissi and colleagues [40] to expand on Lehman's theoretical model of QOL in order to capture the process of appraisal (comparison standards, aspirations, expectations, values and beliefs) in perception of QOL. As seen in Figure 2**,** the role of objective life circumstances is theorized to be dependent on mediating variables such as personal characteristics, self-related constructs(self-esteem, self-efficacy, perceived control), clinical characteristics, and other variables which influence the appraisal process. This theoretical model was empirically tested and the authors confirmed that indeed mediating variables such as autonomy, self-related constructs, and the perception of change predicted global QOL and mediated other

objective variables. The relevance of these mediating factors is supported by a review of qualitative research on QOL, which highlights the importance of domains such as autonomy, self-perception, choice, control, and hope in determining global QOL, according to people with mental illness[40].

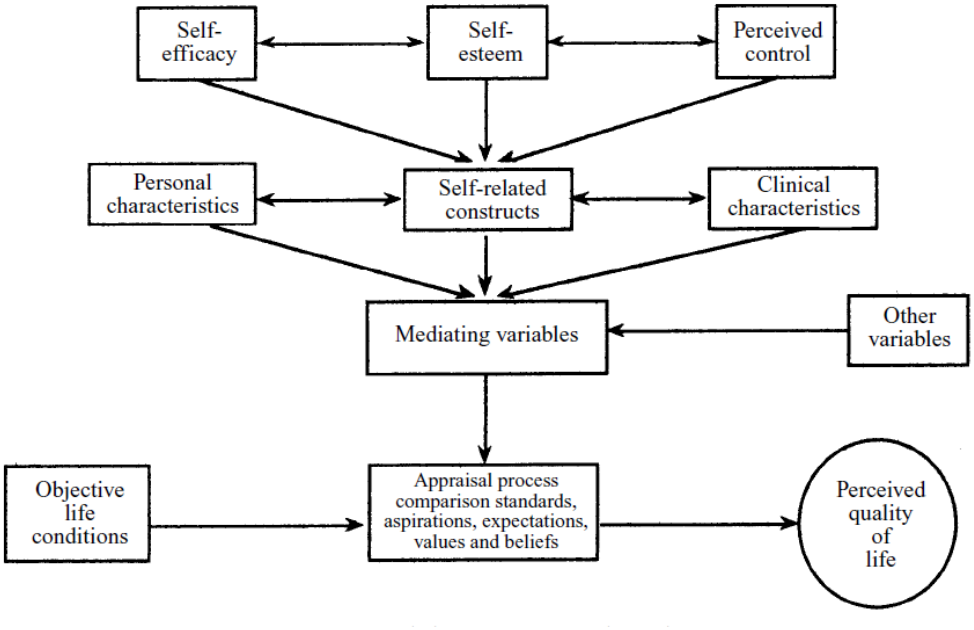


Figure 2. A mediational model of QOL devleoped by Zissi et al.[40]

Summary

In understanding QOL in a population with mental illness, factors that play the most important roles in determining global QOL are life satisfaction across different domains, self-related constructs (e.g., self-esteem, autonomy, mastery), psychiatric symptoms, appraisal of objective life indicators, and to a lesser extent objective indicators themselves, perhaps playing a more indirect role on QOL. Demographics variables (e.g., age, sex, maritral status) may have some influence depending on the specific population, but in general are not as important in determining QOL.

2.5 Quality of Life in Homeless Populations

Though these patterns contribute to an understanding of QOL in a population with mental illness, none of the studies cited above specifically sample individiuals experiencing homeless. Several of those studies do however examine effects of housing on QOL among individuals with mental illness who have not specifically experienced homelessness. Some observe these effects from comparisons following discharge from hospital [40,71], or across different types of residential treatment settings [43,58]. Other studies that do not include homeless participants have reported that QOL is neither significantly predicted by the type of housing provided (public, private, group home, etc.) nor by the perception of choice in housing[72].

Studies of individuals with mental illness who are homeless show that QOL ratings and change in QOL ratings are associated with similar correlates or predictors as some of the broader mental research discussed above. For example, Calsyn and colleagues [73] found that stronger predictors of QOL in 178 St. Louis homeless individuals with mental illness were psychopathology, self-related constructs (e.g., self-esteem), and social variables (interpersonal adjustment and alienation). Conversely, demographic and objective indicators of QOL only weakly correlated with either total QOL scores or domain-specific ratings. A related study however found that men in this population had lower QOL ratings than women across 7 domains of a QOL measure similar to the QOLI, the Satisfaction with Life Domain Scale (SLDS) [74].

Although homelessness may not significantly alter what types of measures correlate with global QOL, it is consistently proven to itself be associated with lower QOL. For instance, a study by Lehman and colleagues [37] found that homeless participants in four US cities reported lower QOL ratings compared to their housed counterparts in domains of finances, daily activities, family relationships, and, evidently, living situation (lower global QOL ratings were not statistically different). These findings were supported by a study comparing different groups of health clinic service patients by Stein and Gelberg [75], where 747 Los Angeles homeless patients reported lower life satisfaction (on a single general item) than housed patients. However, one diverging finding was reported by Van der Plas and colleagues [76] who

found no global or domain-specific QOL differences between homeless and housed schizophrenia patients in the Netherlands, except for the domain of health, in which the homeless rated themselves *higher.* Aside from this aberrant finding, homelessness does seem to negatively impact the QOL of individuals with mental illness when comparing to housed counterparts.

In contrast to analyses across housing status groups, somewhat more definitive conclusions can be made about effects of housing from longitudinal QOL comparisons, within groups of individuals with mental illness progressing out of homelessness. A study by Wolf and colleagues of 485 Los Angeles homeless individuals [77] found that QOL improved across different domains as individuals exited out of homelessness, either into dependent or independent housing. Those exiting into independent housing showed highest QOL improvement specifically in domains of housing, leisure, and money. Similarly, a cohort study by Kertesz and colleagues [78] looked at mental health-related QOL (SF-36) in 274 Boston individuals with substance addiction problems showing different patterns of homelessness over time. Participants who remained homeless (chronically homeless) over 2 years of follow-up showed smaller improvement in mental health-related QOL, in comparison to those who transitioned out of homeless, themselves improving less than a group who were always housed [78]. Findings on QOL in homelessness from both these longitudinal studies agree with findings from studies comparing across groups cross-sectionally, namely that homelessness negatively impacts QOL.

Given this evidence, one can expect that housing would improve QOL in this population. Therefore, it may not be surprising that treatments offered to homeless individuals with mental illness *without* a housing component would not consistently show QOL improvements. A UK-based randomized controlled trial of case management for 80 homeless individuals with mental illness failed to show QOL changes for treatment group, who along with controls did not improve after 14 months[79]. Similarly, a study by Lehman and colleagues [80] found that Assertive Community Treatment (ACT) (the same treatment provided to high need participants in the current study) did not improve QOL in 77

participants over 12-months any more than improvements seen in 75 TAU controls. Significant group differences in QOL were only seen at the 6-month follow-up, in terms of global QOL and domains of neighbourhood and health. On the other hand, a large study by Lam and Rosenheck [81] did find significant 12-month improvements in global QOL among over 4000 homeless individuals with mental illness participating in case management programs across the US. However, no comparison group provided evidence of this QOL improvement as being an effect of treatment. Despite this, it is worth noting that one significant predictor of improvement in global QOL was an increase in stable housing.

Findings of improved QOL are somewhat more consistent when interventions provide both support services and housing to homeless individuals with mental illness. One study by Bebout and colleagues [82] examined the effect of clinical services and housing support on the QOL of 158 homeless individuals with both mental illness and a substance use disorder in Washington, DC. However, given that housing was provided along a continuum model (placement depended on staff approval), only 41% of participants were stably housed by the end of the study. Differences in QOL were seen between these groups, with those housed showing higher global and domain-specific QOL (housing, family, leisure, and finances)[82]. One study of 78 New Jersey homeless women with children by Camasso[83] found that participation in a transitional housing program improved QOL, with improvements predicted by hope at baseline and participation in program activities. Though the study did not recruit homeless women specifically with mental illness, the sample showed high average scores of depressive symptoms. Finally, research by Schutt and colleagues[84] revealed only partial QOL improvements following random assignment of 89 Boston homeless individuals to either independent housing or group homes, both with intensive case management. Satisfaction with housing features and residence overall improved over 18 months, while global QOL did not change significantly change. Any improvements were however equivalent across assigned housing type.

The literature therefore establishes that QOL in homeless populations with mental illness is responsive to interventions and that there is possibly better evidence for QOL improvements when housing is provided in addition to clinical services.

2.6 Quality of Life and Housing First

Most relevant to the current study is the effects on QOL of interventions under the Housing First model. In the interest of understanding modest QOL results of the At Home / Chez Soi study, a focus on similar studies reveals some evidence that Housing First improves QOL. This evidence is however not without limitations.

One study of a Housing First project in San Diego [1] demonstrated a link between a HF intervention and higher overall and domain-specific QOL. 161 clients receiving HF were compared to 86 homeless individuals initiating outpatient services around the same period. These controls were matched by propensity scores on age, sex, race/ethnicity, living situation, and clinical diagnosis (for a quasi-experimental study). QOL measurements were examined cross-sectionally, from a biannual survey of clients. All QOL domains as well as global QOL ratings were higher for the HF group in comparison with controls, with largest group differences seen for the domain of living situation. It is important to note, however, that no baseline measure of QOL was provided for either group, limiting conclusions about the effect over time of HF in improving QOL any more than existing community services[1].

A Canadian Housing First study shows similarly positive effects on QOL, with some limitations. The City of Toronto assessed the outcomes of the Streets to Homes project[85], providing HF services to 88 formerly homeless individuals. In almost all important measures grouped as domains of QOL, a majority of clients reported improvement (rather than “stayed the same” or “worsened”). Health improvements

were reported in 70% of clients and 72% reported improvements in personal security. 60% said their level of stress had improved, while 57% reported improved mental health. Social interaction showed the least improvement, improving only for 40% of clients, while 26% reported worsened social interaction and 34% "stayed the same". Of course, one limitation of these findings is the lack of a control group to infer these changes as effects of HF. Additionally, an important distinction of the Streets to Homes project is that, in comparison to other HF interventions, participants were not recruited on a basis of mental illness[85].

Another study based out of Denver, Colorado reported HF's positive effects on QOL for homeless individuals with disabilities, again without comparison to a control group[3]. Based on a subset of 19 participants out of 137 enrolled in HF services, the authors reported a 2-year improvement in overall QOL for 64% of study participants. This improvement was the most successful of the study's secondary outcomes (beyond housing stability). Comparatively, substance use was reduced in only 15% of participants, while mental health improved for only 43% and health status improved for 50%. The report of the project was unfortunately more detailed in terms of housing and cost-effectiveness, while little was provided to describe QOL measures. Additionally, the lack of a control group again prevents comparisons between HF's effect and those that would have otherwise been observed with usual services[3].

Notably, in the only study of Housing First comparing QOL outcomes over time to those of control group, HF did not show a significantly greater effect on QOL. A study by Mares and colleagues [2] did report 12-month improvements in mean overall QOL in 296 chronically homeless participants in HF programs across multiple sites in the US. However, these QOL improvements were no greater than those of a comparison group of 118 homeless individuals receiving usual care over the same period. Additionally, in both groups, QOL improvements over the 12-month period were modest. One limitation

of the study was the lack of randomization to either group. However, comparisons were adjusted for group differences in baseline measures[2].

Similar to the broader literature revealing that both overall QOL and specific domains of QOL among the homeless are generally responsive to treatments that provide housing, the literature on Housing First also points to evidence of QOL improvements. However, studies have either failed to compare QOL improvements to those of a comparison group, or have only compared between HF and controls at one point in time. Furthermore, the one study that compared two such groups longitudinally did not find group differences in OQL improvements.

One can then assume that the modest QOL results of the At Home/Chez Soi study are not so surprising, given their partial agreement with previous literature. However, results of the At Home/Chez Soi study and other HF literature are nonetheless puzzling. All these studies suggested that treated participants were more successfully housed and often benefitted from treatment in terms of measures other than QOL. Additionally, the greater positive outcomes demonstrated in narrative interviews of the At Home/Chez Soi subsamples point to larger QOL treatment effects than those observed[24].

Different explanations are offered as to why treatments did not have as large an effect on QOL. Authors of the National Report of the At Home/ Chez Soi project suggest that "regression to the mean" may underlie the parallel increase in QOL of the control group[24]. They describe the recruitment criteria of recent or current homelessness as a "crisis" circumstance from which most participants are expected to generally improve, regardless of treatment. This may explain why, even in the absence of permanent stable housing and clinical services, QOL seems to increase on average.

In Lehman and colleagues' study of ACT among the homeless [80], it is suggested that the control group showed similar improvements in QOL because they may have benefitted from additional investments in housing and shelters that were tied to the grant of the ACT study itself. The city also later

began its own ACT program in the community, thus possibly leading to “contamination” among the control group. Data to support these assumptions were not recorded, however.

An alternative phenomenon may perhaps better explain why the two groups’ improvements in QOL did not differ more meaningfully or consistently over 2 years, considering that the HF group was much more stably housed and received individualized clinical services. The current study examines the possibility that the treatment effect on QOL was moderated by response shift. As explained in detail in the next chapter, response shift is a consequence of adaptation to life events leading to reappraisal of QOL’s determinants, and thereby a reinterpretation of QOL itself [5]. Over the course of the study period, both HF and TAU participants may have experienced events (whether the intervention itself or other experiences) that initiated response shift. In such cases, response shift would mean that measurements of QOL at different time points (before and after being housed) and across treatment groups (HF and TAU) may not always be comparable. A more accurate effect of the treatment would be evident in participants whose ratings of QOL are based on similar interpretation of the measure at every time point. This study sets out to identify treatment and control individuals for whom these QOL outcomes are indeed comparable, that is, for those showing no response shift, or “true change”. Our expectation is that the extent of response shift in other individuals will modify the QOL group differences seen over time (i.e. the treatment effect).

Such a hypothesis requires a deeper understanding of how response shift influences QOL outcomes, how it can be identified, and how doing so can possibly elucidate treatment effects.

CHAPTER 3: Response Shift

Response shift is defined by a change over time in the meaning of self-reported QOL, specifically when this construct is targeted for evaluation[5]. Though first described in the fields of educational training [86] and organizational development [87] , response shift has more recently been studied in clinical research, where patient-reported outcomes are increasingly emphasized. According to its most common definition by Schwartz and Spranger[5], response shift may follow a change in an individual's (1) internal standards of measurement (recalibration), (2) values (reprioritization), or (3) definition of the construct (reconceptualization).

An overview of response shift by Barclay-Goddard and colleagues [8] illustrates each of the three response shift mechanisms as they would occur following changes in health. Recalibration is described by a situation where a pain scale of 0 (no pain) to 10 (worst pain imaginable) is reinterpreted following a patient's kidney stone experience, shifting her earlier rating of a bruised knee from a score of 8/10 to one of only 4/10. Reprioritization is described by an example where a patient's health scare reorders the value of family interaction ahead of physical function, whereas he previously would have ranked the values in the opposite order. Reconceptualization is described by the case of a woman whose health-related QOL is originally determined mainly by her energy level, physical function, and mental health. A cancer diagnosis and treatment causes these 3 contributors to health-related QOL to be replaced (and not just reprioritized) by family, bodily pain, and fatigue[8].

Each one of the response shift mechanism is by nature a deviation from *true* change. In earliest descriptions of response shift, true change (then labelled *alpha* change) was defined as "a variation in the level of some existential state, given a constantly calibrated measuring instrument related to a constant conceptual domain" [87]. This "phenomenon" is precisely the type of change assumed in most longitudinal comparisons of self-reported measurements, regardless of the size or significance of the change, whether it is random or not, or whether comparisons of change are made across groups. While

true change was initially defined in the context of change following an intervention, different life events are now recognized as leading to true change, be they positive or negative, incidental or assigned. A simple example of true change given by Golembiewski and colleagues[87], where the feet a child have grown according to conventional concepts of "length", estimated over time by a constant scale of measurement. More conventionally, this would correspond to examples in research where change occurs without response shift.

In comparison to true change, the various types of changes from response shift may act alone or simultaneously in causing paradoxical findings in clinical research. Among the many examples of response shift in the literature, most cited are cases where the *catalyst* for response shift is an illness or a change in health state. These include examples where patients suffering from severe disability or chronic or even terminal illness show similar QOL ratings over the progress of the illness or in comparison to healthier counterparts[5,88]. Such effects of response shift have been observed in a range of health conditions such as cancer, stroke, multiple sclerosis, diabetes, and dental disorders, among many others[9].

On the other hand, the catalyst for response shift may be a treatment seeking to improve a health condition. In fact, in many clinical settings, such as multiple sclerosis or palliative care, response shift may actually be the desired effect of an intervention. Such an intervention would target a patient's psychological adaptation to an illness where little hope exists for true QOL improvement, such as a degenerative condition [11]. However, in instances where a true change in QOL is expected following a treatment, the different types of response shift can limit the accuracy with which on can capture such change in self-reported QOL. Specifically, following response shift, comparisons are no longer made on a concept of QOL that is the same from one time to another[8].

For example, such undesired effects of response shift explained counterintuitive findings in a psychosocial intervention that sought to promote self-esteem among 22 cancer survivors [89]. Surprisingly

detrimental effects on participants' QOL, at 3-month follow-up, were explained by a recalibration of their standards of QOL, resulting from the intervention heightening of their expectations for daily life. Similarly, participants were found to have "normalized" their concept of QOL to be closer to that of 54 healthy age-matched controls. In doing so they reprioritized the values contributing to their QOL, approaching that of controls[89].

3.1 Theoretical Models of Response Shift

To better understand the process by which response shift can impact evaluations of QOL, several theoretical models propose different components of the response shift process and the links between them. A frequently cited theoretical model of health-related QOL that integrates response shift was developed by Sprangers and Schwartz [5]. As seen in Figure 3**, t**he key components of the model include

1) a catalyst: a change in health status, either an illness or a treatment,

2) antecedents: personality, expectations, or stable sociodemographic variables,

3) mechanisms: processes that occur in response to the catalyst such as social comparison, coping, goal reordering, or reframing expectations,

4) response shift, and

5) perceived QOL: defined as being multidimensional (physical, psychological, and social functioning).

The process of the model is initiated by a catalyst's triggering of mechanisms to accommodate the change in QOL (this process was described for health status by the authors). These behavioural, cognitive, and affective mechanisms lead to response shift by changing standards (scale recalibration), values (reprioritization), or conceptualization of QOL (reconceptualization). Antecedents will influence

whether individuals adopt one mechanism or one type of response shift over another. The effect of perceiving low QOL can reinitiate accommodating mechanisms, causing a cycling of the process[5].

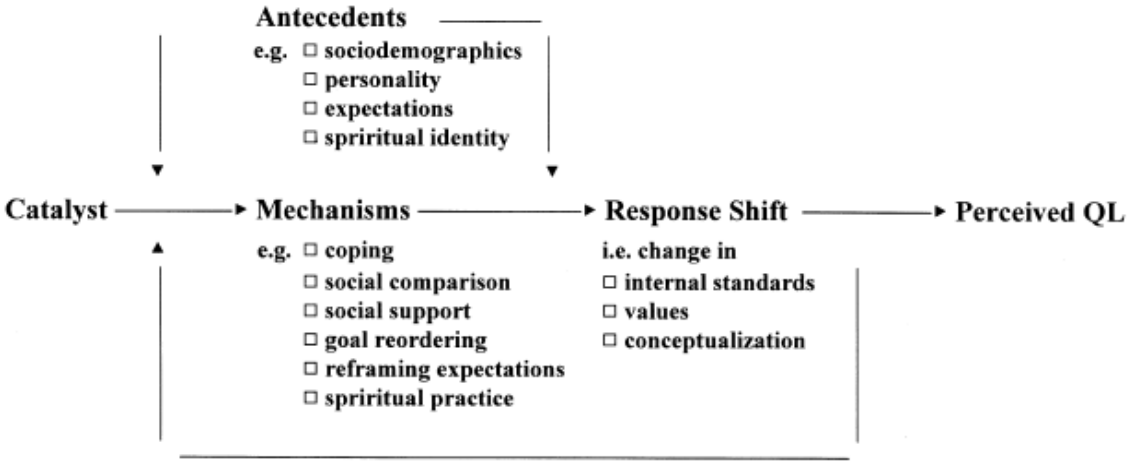


Figure 3. A theoretical model integrating response shift in self-reported QOL by Sprangers and Schwartz[5]

Though the focus of this model is on the process of response shift, there are similarities to existing QOL models in mental illness, such as the mediational model defined by Zissi and colleagues [90]. The component of antecedents includes the role of personality, under which Sprangers and Schwartz include self-esteem, as well as sociodemographic factors that influence the mediation by response shift. A similar mediation is the key component of the Zissi model. Additionally, elements of the mechanism component in this model actually overlaps fittingly with Zissi's appraisal process component, itself defined by values, beliefs or comparison standard reconceptualization. If anything these similarities are evidence of the importance of response shift, given that the Zissi model approximated its definition without seeking to do so. It also points to the likelihood that response shift is relevant in a population

with mental illness, even though no mental health-specific models have sought to account for this phenomena.

An expansion on the Sprangers and Schwartz model has been developed by Rapkin and Schwartz [6]. As seen in Figure 4**,** the more recent model sought to better take into account the appraisal process in ratings of QOL and to address some issues with circularity and ambiguity in elements of the original model. The theoretical model proposed by the authors contains 3 families of hypotheses as to how QOL ratings can change.

1) The first family (S_1-S_3), refers to the standard assumptions in QOL research, that is, the assumption of true change. Both catalysts (life events, illness, etc.) and antecedents (demographics, personality, etc.) influence QOL directly, though catalysts may also mediate the influence of antecedents (S_2).
2) The second family of hypotheses (C_1-C_3) addresses the mechanisms adopted to accommodate catalysts, which can also be influenced by antecedents. The authors expand on the original Sprangers and Schwartz model by hypothesizing that these mechanisms may also moderate the direct impact of catalysts on QOL (C_3).
3) The third family of hypotheses (A_1-A_3) addresses the process of appraisal of QOL. This process can be influenced by mechanisms, catalysts, or antecedents. The authors make the distinction between direct response shift in which changes in appraisal directly influences self-reported QOL and moderated response shift in which the appraisal interacts with the direct link between catalysts and QOL.

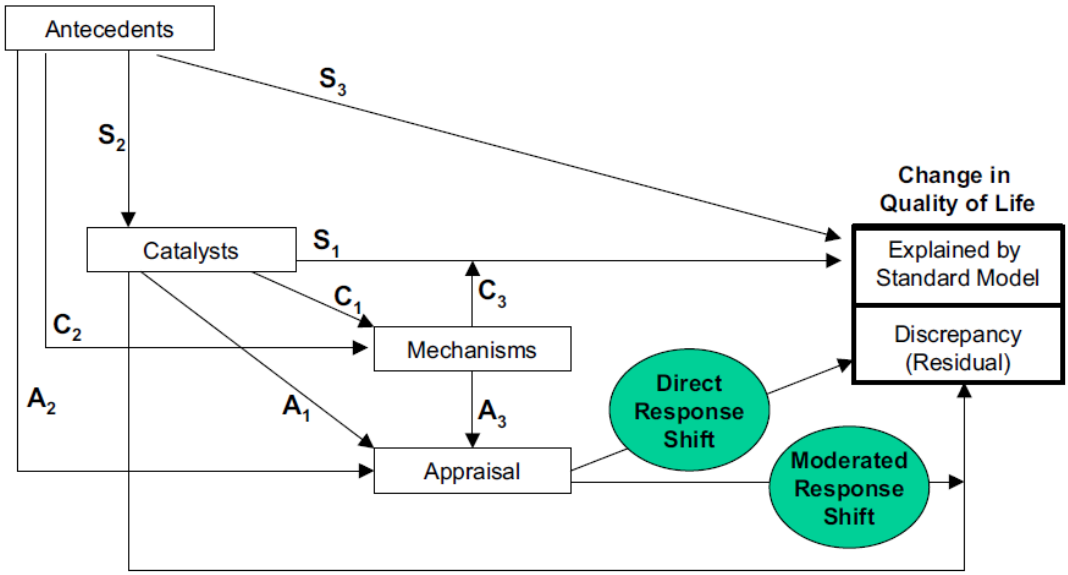


Figure 4. A theoretical model of response shift in QOL by Rapkin and Schwartz[6] :"Partitioning response shift effects in the Sprangers and Schwartz (1999) model using a linear regression paradigm: Accounting for changes in Standard influences (S), Coping processes (C), and Appraisal (A) variables"

With considerable detail, the authors describe this appraisal process as an elaborate algorithm. Specifically, the process is a weighted combination of evaluations of experiences or events sampled from a frame of reference. This frame of reference refers to a given aspect of individual's life that is personally deemed important for their QOL evaluation[6].

Each response shift mechanism is defined by a change in a specific element of this algorithm. Reconceptualization is defined as a changes in the frame of reference. Reprioritization can be defined as a change in strategies for sampling experiences within the frame of reference or as changes in the factors that give salience to these experiences. Finally, recalibration is defined as a change in the standards of comparison under which experience are compared.

3.2 Response Shift as Discrepancies/Residuals

The most important element of the Rapkin and Schwartz model is its definition of response shift as a discrepancy from expected QOL. The authors distinguish between true QOL change (explained by standard influences) and any QOL change that is discrepant from the change expected from standard influences[6]. This latter change represents response shift that is explained by changes in appraisal. Exploiting this interpretation enables response shift to be identified in regression-based analyses of QOL change. Specifically, variance in observed QOL change that is discrepant from expected change, i.e. the residuals of a model, is said to represent response shift. They stipulate that this inference can be made when the variance of residuals is explained by measures of appraisal. As will be discussed in detail further below, this interpretation of residual variance largely forms the basis on which the current study identifies and adjusts for response shift[6].

Another study also interprets response shift from indicators of discrepancies in QOL. Building on earlier work on well-being [91], Evans and Huxley [7] describes response shift as any change in QOL that deviates from observed change in life situation or circumstance. That is, when circumstances improve, deteriorate, or remain stable, reported QOL is expected to change accordingly. Discrepancies from this pattern is assumed to reflect an internal accommodation or adaptation to these circumstances. In a similar, though perhaps less nuanced interpretation than seen in other models, response shift is said to occur when this adaptation impacts self-reported QOL[7].

In their study, Evans and Huxley[7] identify response shift through inferences about adaptation for each of 1,086 participants across the UK. In each QOL domain, participants responded about their desire for improvement within the domain. Adaptation was inferred when a desire for improvement was discrepant with actual change in circumstance. Adaptations were either positive (aspiration) or negative (resignation). The authors found that the number of aspirations (“raising one’s sights”) actually

decreased QOL ratings in multiple domains, while resignations ("lowering one's sights") mainly increased global QOL.

An additional notable feature of the Evans and Huxley study was its sample, consisting of a healthy general population sample, a common mental disorder sample (based on the health questionnaire score of a survey), as well as individuals with severe mental illness (from both groups of the UK700 case management trial)[7]. Surprisingly, no earlier study was identified in searches of the response shift literature with a mental illness sample. Beyond this paper, only one other recent study discussed in the next section[92] examines response shift in a sample of schizophrenia patients.

Though rarely studied within populations with mental illness, response shift may still have occurred within the At Home/Chez Soi study. In fact, the many examples of response shift mechanisms and processes described in the literature suggest different possible deviations from true QOL change, both in HF and TAU groups. After obtaining independent housing, some HF participants may have reprioritized or completely reconceptualized the factors that are important to QOL, perhaps demoting the salience of basic needs in favour of expectations that are more difficult to meet. The HF clinical services may have also contributed to this response shift, given the model's philosophy focusing on hope, recovery, and autonomy instead of symptoms and survival. The philosophy of community integration may also have changed participants' standards of comparison, leading to recalibration.

Examples of response shift under the adaptation model by Evans and Huxley are also feasible for HF participants, but it is to TAU participants that their definition of response shift is most easily applied. Specifically, participants who either remained homeless or obtained unsatisfactory housing may have become resigned to their housing status. In doing so they may have heightened their QOL ratings beyond their unchanging circumstance. On the other hand, the various response shift examples suspected in HF participants may have occurred in TAU participants following other important life

events in the 2-year follow-up period. Similarly, other examples of response shift in HF participants also may have occurred in opposite directions from the ones described due to the heterogeneity of the

3.3 Identifying response shift

If such examples of response shift occurred in the At Home/Chez Soi study, their effect on QOL ratings need to be disentangled from that of true change in order to evaluate the outcome of the study. However, accurately capturing a variety of response shift as extensive as those suggested above would require assessment of appraisal with every QOL rating of the At Home/Chez Soi study. Such assessment is what various response shift identification methods aim for in directly inquiring about appraisal. Of course, such methods must be integrated in the design of a study. For example, in the study of cancer survivors [89], the frequently used *then-test* asked participants to retrospectively reassess their baseline QOL, once they experienced the intervention. Though a popular tool, the *then-test* requires that participants be able to perform reliable long-term reappraisal, without any additional recall bias. Several other methods exist that frame QOL ratings with additional questions on appraisals (e.g., Patient Generated Index, the card sort approach, etc.), all of which require planned implementation beforehand[93].

The other family of approaches infers response shift from statistical trends in longitudinal data[93]. An example of this approach was also seen in Schwarz's study of cancer survivors [89]. Reprioritization was detected using covariance analysis, identifying how the relative weights of components of QOL changed over time. Their small sample size precluded formal tests, however. Other advanced statistical methods allow for different kinds of secondary analysis of response shift. Approaches with similar aims as covariance analysis include *structural equation modelling* (SEM), latent class *growth modelling*, or *Relative Importance Analysis* (based on logistic regression and discriminant analysis models)[9].

The advantage of these statistical methods is that underlying response shift can be identified post hoc. As long a study collects information beyond QOL ratings, such as "objective" measures of QOL (e.g., functioning) or measures of domains of QOL (e.g. satisfaction with family or finances), response shift can be inferred from changes in the association of these measures with QOL ratings.

One very relevant example of such secondary analysis, given its study population, is the analysis of reprioritization response shift in a cohort of 233 schizophrenia patients across Europe by Boucekine and colleauges. The authors used a *random forest* method, a bootstrapping of several recursive partitioning trees, to predict QOL over 24 months (based on the general health item of the SF36). Reprioritization of QOL predictors was identified as a change in the importance over time of a predictor variable, such as functioning or mental health, based on the average variable importance (AVI) statistic. Different response shift patterns emerged depending on whether patients' psychotic symptoms improved, worsened, or remained stable. Among patients who improved, mental health increased in importance as a predictor of QOL, while vitality decreased in importance. Among patients who worsened, vitality and bodily pain increased in importance, while mental health decreased in importance. For the stable patients, functioning increased in importance as a predictor of QOL. The authors suggest several clinical implications of their findings, such as the tailoring of therapeutic approaches based on how the meaning of QOL changes for patients with different profiles[92].

While it may be among the rare examples of secondary analysis of response shift in a mentally ill population, the Boucekine study illustrates the main disadvantage of most secondary analyses of response shift: that they usually only describe patterns of response shift at the group level. In the study reprioritization of QOL's predictors cannot be identified within an individual, but only for the whole patient group (according to symptom stability)[92]. The aforementioned study by Schwartz and colleagues[89] also uses a group-level identification of response shift, requiring the direct, individual-level *then-test* to permit response shift adjustments. One of the most commonly used secondary analysis

methods, SEM, involves iterative testing of models under different constraints[9]. The goal is to locate response shift inferred from changes in model parameters (e.g. factor structure). Not only does this method similarly restrict conclusions about response shift to the group level, but it also requires that the majority of participants in the study experience response shift for it to be detected at all[9]. Just as in other secondary analyses of response shift, identifying responses shift at the group level serves mostly a descriptive purpose, rather than facilitating individual-level adjustments for response shift.

3.4 Identifying Response Shift in Residual Values

A novel approach has been developed by Mayo and colleagues [10] to identify response shift at the individual level. Referred to *as Latent Trajectory Analysis* (LTA), or *Latent Trajectory of Residuals*, this method detects response shift based on the pattern of discrepancies , i.e. residual values, between observed and expected patient-reported outcomes within an individual. Without identifying the exact type of response shift in the data (reprioritization, reconceptualization, or recalibration), response shift is inferred from any change in person-specific QOL ratings that are incongruent with their change in expected QOL estimated from an explanatory model [10].

Such incongruence results in a fluctuation in an individual's residual values. Both LTA and another descriptive approach have been used to quantify or describe such fluctuation. However, the focus on residual values in both approaches fits with the theoretical interpretations of response shift from the model by Rapkin and Schwartz [6] as well as the discrepancy interpretation in Evans and Huxley[7]. Specifically, response shift is inferred when a change in QOL is incongruent with changes expected from standard influences [6] or is discrepant from changes in circumstance [7].

The original paper by Mayo and colleagues [10] that developed the LTA method of identifying response shift was conducted with 678 stroke patients across Canada, assessed four times over the course of a year (baseline, 1, 3, 6, and 12 months). The outcome measure was self-reported health, from

the EQ-5D 100-point Visual Analog Scale (VAS). Expected EQ-VAS scores were obtained from a random intercept model, containing explanatory covariates fitting under a specific set of criteria. These criteria and their adaptations for the current study are explained in greater detail in the methods section of this document. However for an exact breakdown of the methodological approach applied in the Mayo et al.[10] paper, see Appendix C.

Briefly, following the estimation of expected EQ-VAS scores, differences between reported and expected scores (i.e., residual values) are calculated for each individual at each time point. These residual values are centered on their person-specific average, enabling a latent categorization of residual patterns over time using group-based trajectory analysis (GBTA) [94].

The results of this study demonstrated various forms of response shift in 28.5% of participants (2 negative and 3 positive response shift trajectories), while the majority (67.4%) showed stable residuals over time (no response shift or true change)[10]. The remaining 4.1% of participants with excessive variation in residuals did not fit any of the response shift latent groups identified and were classified as an “unstable groups”.

Steps taken to validate the method enhanced the value of the study. The authors compared the LTA results to both a direct assessment of response shift using the *then-test* and results of the LTA under simulated data. They concluded first that the LTA of residuals does indeed identify similar response shift as that identified directly. Specifically, retrospective reassessments indicative of response shift in the *then-test* generally matched the direction of the response shift group (positive, negative, and no response shift). Second, using 20 randomly generated data sets, LTA methods did not identify the same patterns that occurred with actual response shift data. Mainly, far fewer trajectory groups were identified, in only 1 out of 20 datasets was a stable residual pattern identified for the majority of participants, and model fit of the LTA was considerably worse from simulated data [10].

Given the benefits of identifying response shift at the individual level through secondary analysis, two other papers have applied the LTA model: one in a multiple sclerosis (MS) sample [95] and another in inflammatory bowel disease (IBD) [96]. A surprising finding of the study of MS patients was the very small portion of patients identified as showing any response shift[95]: 99.7% of the sample of 1,566 were classified in a no response shift group (stable residuals). This near absence of response shift may reflect the voluntary nature of SF-12 general health assessments in the sample (web and mail surveys), possibly oversampling responses when patients were not encountering any medical symptoms [96]. The IBD study also identified a predominant (but smaller) no response shift group (82.1% of 388 participants), while in other participants negative (8.6%), positive (6.4%), and "rebounded" (3.1%) response shift was detected[96].

An additional paper by Mayo and colleagues [97] illustrates the flexibility of an approach inferring response shift from residual patterns. The study again included stroke patients, this time 190 participants representing both groups of a randomized controlled trial of case management. The authors detected response shift through a descriptive measurement of residual values over time. Short of the minimum number of follow-up assessments needed to apply LTA, patterns of residuals were instead described based on differences between each follow-up residual and the baseline residual (e.g. lost 10, gained 10, stayed within 10). Participants were also classified as either realistic, optimistic, or pessimistic based on the direction and magnitude of baseline residual values (within, greater, or less than 10 points of expected EQ-VAS, respectively). Once again, the largest portion of participants (50-58%) showed no response shift, without much difference between treatment and control groups [97].

3.5 Response Shift Adjustment

The identification of true change and response shift in these studies enable adjustment for response shift in analyses of the reported outcomes. This adjustment may elucidate comparisons of

patient-reported outcomes over time or between groups if these comparisons differ as a function of response shift. It is worth keeping in mind that the identified response shift may not necessarily be undesired, especially in cases of positive response shift. However, the aim is to better understand changes in these outcomes by highlighting this reported change when the measure is not reconceptualized, reprioritized, or recalibrated.

In the first paper developing the method by Mayo and colleagues [10], the authors compared EQ-VAS changes in individuals with no response shift to those with any response shift (averaging across the 6 response shift groups). This comparison demonstrated that response shift does not significantly affect estimates of EQ-VAS change when response shift is ignored, given that the outcome improved similarly regardless of identified response shift.

In the IBD study, the classification of response shift trajectories also allowed for a response shift adjustment of the SF-36 general health perception outcome [96]. It was found that the no response shift and negative response shift groups did not differ in terms of reported change, with both groups' participants showing stable outcomes over the study period. However, patients with positive and rebounding response shift were the only groups for whom improvements in the reported outcome were identified, suggesting that response shift explains any positive change in health ratings.

Finally, the other studies using similar residual-based approaches did not adjust reported outcome comparisons for response shift. In the multiple sclerosis study [95], the 0.3% of participants showing response shift was too small of a group to make meaningful comparisons of general health ratings to the no response shift group. In the study examining response shift in a post-stroke trial [97], the authors also did not analyze changes in reported EQ-VAS ratings as a function of residual patterns. This omission leaves unanswered the question of whether such response shift adjustments can explain modest treatment effects observed over the study period.

Hypothesis

The hypothesis of the current study is that indeed a response shift adjustment may help identify larger treatment effects on QOL in the the At Home/Chez Soi study. Specifically, it is expected that the extent to which a person shows response shift through fluctuating residual values, as opposed to true change (stable residuals) will modify estimated group differences in QOL, with more true change demonstrating larger differences. This hypothesis will be tested for the QOLI-20 total score as well as the QOLI-20 global item rating.

CHAPTER 4: Methods

The current study is a secondary analysis of results from the At Home/Chez Soi project. See Goering et al. [23] for a more detailed description of the study (enrolment criteria, specifics about the intervention, and all questionnaires administered).

4.1 Participants

A total of 2148 individuals experiencing homelessness and mental illness were enrolled for the At Home/Chez Soi project, across Vancouver, Winnipeg, Toronto, Montreal, and Moncton. Participants were recruited from referrals by agencies offering services to the homeless in the community and in institutions (shelters, drop-in centres, outreach and mental health teams, hospitals, and criminal justice programs).

To be eligible for the study, participants had to be at least 18 years old (19 for Vancouver), present a current mental disorder, be experiencing either precarious housing or absolute homelessness, and not be receiving Assertive Community Treatment (ACT) or Intensive Case Management program services. A current mental disorder (major depressive, manic, or hypomanic episode, or a post-traumatic stress, panic, or psychotic disorder) was identified either by the Mini International Neuropsychiatric Interview (MINI 6.0) or from a documented diagnosis from the past year (or up to 5 years in the case of a psychotic illness).

4.2 Intervention

Housing First participants obtained access to housing of their choice through rental supplements ensuring they spent no more than 30% of their income on rent. In addition, HF participants received one of two forms of individualized, recovery-oriented clinical support services based on assigned level of need. Assertive Community Treatment (ACT) was provided for the High Needs group and Intensive Case

Management (ICM) for the Moderate Needs group (only ACT was offered in Moncton given the site's smaller sample). The only condition for receiving both rental supplements and clinical services was agreement to meet a support team member at least once a week. Participants randomized to the TAU group were left to seek out their community's existing services for housing and/or support.

Assessment periods (21 vs 24 months)

While most participants were administered complete interviews every 6 months, the project's endpoint was moved from 24 to 21 months, as a cost-saving measure, for approximately half of the participants (recruited later) in Moncton, Montreal, and Toronto for budgetary purposes. For these participants 21 month assessments are included in both the explanatory model and the response-shift-adjusted QOL analysis stages and treated as their 24 month assessments.

4.3 Measures

Quality of Life

Quality of life outcomes were obtained from the 20-item Quality of Life Index (QOLI-20)[29]. The QOLI-20 measures life satisfaction in the domains of everyday living, leisure, family, social relationships, finances, and safety. Every item is scored from 1 ("Terrible") to 7 ("Delighted"). A global QOL item asks "How do you feel about your life (as a whole)?". The original At Home/Chez Soi analysis reported QOL outcomes using the QOLI-20 total score. The current analysis reports on changes in both the total score and the global item.

The original, longer version of the QOLI-20, the QOLI, was developed and has been validated for use in populations with mental illness [28,33,37,55]. It has also been used in previous literature on the QOL of homeless individuals[80–82]. A review by Lehman[33] examines the validity of the QoLI through its psychometric properties. Psychometric properties of QOLI items adopted in the QOLI-20 include internal

consistency reliabilities between 0.79 and 0.88 (median = 0.85) and one-week test-retest reliabilities between 0.41 and 0.95 (median = 0.72). The QoLI has been evaluated as having good construct and predictive validity based on confirmatory factor analyses and multivariate explanatory models. Different studies support the QOLI's predictive validity, given its differentiation of patients based on residential circumstance, clinical symptoms and mental health status, and gender, and age [33].

In order to validate the QOLI-20 global item within the study sample of the At Home/Chez Soi study, Spearman correlations were performed for this item with other general, self-reported items at baseline; general health ("poor" to "excellent") from the Short Form-12 (SF-12) Health Survey[98], the EuroQol (EQ-5D)[99] overall, physical, and mental health items of the (on a 100-point "visual analog scale" scored from "worst-" to "best imaginable health state"), and agreement (1 "strongly disagree" to 5 "strongly agree") with the statement "I like myself" on the Recovery Assessment Scale (RAS) [100]. Spearman correlations were also examined between the global item and the total score of the QOLI-20.

As many as 36 other measures were used in different stages to identify response shift in QOLI-20 global item ratings. Demographic and other variables measured only at baseline included: gender, age, education, ethno-racial status, immigrant status, and total time in life spent homeless from a Demographics, Housing, Vocational and Service Use History (DHHS) questionnaire, total physical comorbidities from a Comorbid Condition List (CMC), the total score from the Adverse Childhood Experiences (ACE) questionnaire[101], as well as the baseline total score of the RAS. Time-varying covariates, recorded concurrently (every 6 months) with the QOLI-20, included: the percent of the last 90 days stably housed from the Residential Timeline Follow-Back (RTLFB)[102], interviewer-assessments of health impairment, psychosis/thought disorder, mood abnormality, response to stress/anxiety, acceptance of illness, meaningful activity, social network and social interest from the Multnomah Community Ability Scale (MCAS)[103], self-reported anxiety, cognitive trouble, feelings of self-harm/suicide,

and difficulty fitting in from the Colorado Symptom Index (CSI)[104], unmet health need from a Health Service Access Items (ACC) scale, pain/discomfort and problems with performance of usual activities from the EQ-5D, emergency/ambulance use and victimization from the Health, Social and Justice Service Use Inventory (HSJSU), different sub-domains and the personal confidante item from the QOLI-20, current employment and past-month income from the Vocational Timeline Follow-Back (VTLFB)[105], general health from the SF-12, as well as past-month alcohol and drug use problems from the Global Assessment of Individual Need – Substance Problem Scale (GAIN-SPS)[106].

4.4 Statistical Analyses

The analysis involves three main steps to adjust for response shift bias in QOL comparisons of HF and TAU controls: 1) the development of an explanatory model of QOL, derived from the protocol developed in Mayo and colleagues [10], 2) the scoring of every individual's extent of response shift based on the standard deviation (*sd*) of their residual values, and 3) the analysis of the HF treatment effect on QOL adjusting for person-specific *sd* of residuals.

Additional sensitivity analyses were performed to assess the influence of response validity and sincerity based on the Interviewer Impressions Index (III).

4.4.1 Explanatory model

The first step of the response shift identification method was the development of an explanatory model that best predicted ratings on QOLI-20 global item (response shift is identifiable on individuals' ratings on a single item, rather than the total score summing domains of an instrument).

The explanatory model consisted of a multivariate random-intercept model, with subjects treated as clusters, and was estimated using restricted maximum likelihood (REML). Certain adaptation were made to the model building guidelines in Mayo et al.[10]. For example, the current analysis model fit

was estimated using a pseudo-R^2 statistic recently developed by Nakagawa and Schielzeth[107], complemented with values of the Alkaike Information Criterion (AIC). Additionally, the current analysis did not perform any substitution for missing data but rather restricted the model to complete cases. Further deviations from the original methodology will be indicated below.

Several criteria directed the selection of covariates. First, multiple theoretical models of QOL in individuals with mental illness guided for determining the most relevant covariates among the many available candidate measures [28,40]. Furthermore, the original residual analysis methodology stipulated that QOL not be modeled by using covariates that are themselves vulnerable to response shift. For example, reprioritization may also affect an individual's rating of self-reported health over time. Instead of excluding all such covariates, the current analysis included them at their baseline measurements only. For other covariates that varied over time, covariates from a different cognitive framework than that of the QOL measure were selected (e.g., interviewer's observations reflected in the MCAS). The original methodology also prescribed that covariates be excluded if their effect on QOL changed over time. The current analysis includes the main effect of these covariates regardless of time interactions, in order to avoid ignoring substantial explainable variance in QOL.

The most important constraint on the model is the exclusion of a variable representing time, to avoid capturing any response shift in expected QOL estimates. Instead, it is presumed that any changes in QOL over time discrepant from those expected from change in covariates will reflect response shift.

As proposed in Mayo et al.[10], the selection of explanatory covariates was based on their statistical significance ($p < .1$) in a multivariate random-intercept model with participants' repeated QOLI-20 global item ratings as the outcome. Several covariates were previously excluded in earlier modeling stages if they showed no meaningful bivariate association with QOL, nor any association within

a multivariate model explaining only baseline QOL. In some instances, explanatory covariates excluded after bivariate tests were added to the multivariate models if they improved the pseudo-R^2 or the AIC.

4.4.2 Standard deviation (*sd*) of residual values

Residual values were obtained for participants included in the analysis of step 1, calculated as the difference from linear predictions based on fixed effects only (Observed QOL_{ij} – Expected QOL_{ij}). Participants with at least two residual values, including a baseline value, were given a response shift score corresponding to the fluctuation or *sd* of their residual values. This response shift metric ignores the magnitude of residuals in individuals with consistently high, low, or small residuals. The *sd* of residuals represents instead the average deviation of an individual's residual values from the mean of their residuals and not from the mean expected QOL. That is, the *sd* metric is not to be confused with an individual's average residual value.

Preliminary analyses revealed that the LTA method was not suitable for these data. The large sample size, tailored clinical services in the HF group, variable community services in the TAU group, and potentially cyclical psychiatric symptoms meant that a high number of highly fluctuating trajectories were identified longitudinally. The uncertainty in modeling these fluctuating trajectories may impede the accuracy with which one can identify stable residuals, or true change. The identification of individuals showing no response shift may in fact be more valuable to a response shift adjustment than then identification of different response shift trajectories. Therefore, a small *sd* of residual was found to be a better measure to estimate this true change. The extent to which people deviated from this true change, interpreted under the framework of this analysis as response shift, is then indicated by an increasing *sd* of residuals.

4.4.3 Response-shift-adjusted QOL analysis

The final step of the analysis involved a response shift adjustment of QOL comparisons, both for the QOLI-20 total score and the QOLI-20 global item.

The total QOLI-20 total score was analyzed using a random effects mixed model (random intercept for participants and random slope for time, with unstructured covariance), estimated based on REML using the *mixed* command from Stata 13 [108]. The model examined the main effects of the HF treatment, time (as a factor variable), and *sd* of residuals, as well as all the 2-way and 3-way interactions of these main effects (HF treatment x time, HF treatment x *sd* of residuals, time x *sd* of residuals, and HF treatment x time x *sd* of residuals). Of interest are the effects of *sd* of residuals on the simple effects of HF treatment at different follow-up periods (i.e., the HF treatment x *sd* of residuals two-way interactions at each period). Significant levels of this effect would suggest that the effect HF treatment on QOL is dependent on the extent of true change or response shift observed in an individual.

The response shift adjustment for the treatment effect on the QOLI-20 global item was estimated by a generalized ordered logistic regression using the *gologit2* command of Stata 13[108]. The item originally scored 1-7 was categorized into 3 broader QOL categories (poor: 1-3, fair: 4-5, and good: 6-7) to adjust for sparse data. The analysis collapsed across follow-up periods (post-baseline) to improve power. The model estimated the main effect of HF treatment and *sd* of residuals, the interaction of HF treatment x *sd* of residuals, while adjusting for baseline QOL category. The coefficients estimated represent each term's effect on the odds of a higher QOL category (cumulative odds of poor vs fair or good QOL and of poor or fair vs good QOL). A significant interaction term would mean that the treatment effect on the odds of a higher QOL category depends on the participants' *sd* of residuals, i.e., response shift. The *gologit2* command performs Wald tests on the model fit with proportional odds constraints compared to that of an unconstrained model, as well as tests comparing the proportionality of each coefficient at different levels of the regression equation (allowing for a partial proportional odds

model). The command does not fit mixed effects but corrects standards error and variance-covariance matrix estimates to adjust for within individual correlations.

Sensitivity Analysis

Each of the three steps of the analysis was carried out again under while excluding responses based on the interviewer's impression of their validity and sincerity. Interviewers finished every follow-up assessment by scoring the question "How confident are you in the overall validity of the information collected in this interview?" as "completely confident", "some doubts", or "no confidence". They also scored the participant's "truthfulness/sincerity" from 1 "very poor" to 5 "very good". Initially, a less stringent criterion (A) was applied on all three steps of the analysis which included responses with a validity rated as either "completely confident" or "some doubts" and sincerity scored 4 or 5. A more stringent criterion (B) included only responses with validity rated as "completely confident" and sincerity scored at 5.

CHAPTER 5: Results

5.1 QOL outcomes

Baseline Spearman correlations of the QOL global item were generally weak to moderate: *rho* = .338 ($p < .0001$) with EQ5D physical health, *rho* = -.371 ($p < .0001$) with SF-12 general health, *rho* = .419 ($p < .0001$) with EQ5D overall health, *rho* = .439 ($p < .001$) with EQ5D mental health, and *rho* = .473 ($p < .001$) with the RAS item. Strong convergent validity was demonstrated from a Spearman correlation with the QOLI-20 total score, *rho* = .722 ($p < .0001$).

As shown in Table 1, ratings on the QOLI-20 global item increased for both HF participants (means (sd) at baseline = 3.5 (1.8), 6 months = 4.3 (1.7), 12 months = 4.5 (1.7), 18 months = 4.5 (1.7), and at 24 months = 4.7 (1.7)) and TAU participants (means (sd) at baseline = 3.5 (1.8), at 6 months = 4.1 (1.7) , at 12 months 4.3 (1.7), at 18 months 4.3 (1.7), and at 24 months = 4.5 (1.7)). The table also describes improvements seen for both groups in the QOLI-20 total score, as well as changes over time in the categorized (poor-fair-good) QOLI-20 global item used in the response-shift-adjusted analysis of the item.

Table 1. Means of QOLI-20 Total Score & Global Item; Frequencies of Global Item Categories

		Baseline	6 months	12 months	18 months	24 months
QOLI-20 total score						
Mean (sd)						
	HF	71.5 (22.2)	86.0 (21.0)	88.6 (21.5)	87.5 (22.2)	89.2 (22.1)
	TAU	72.0 (22.8)	80.5 (22.1)	84.9 (21.5)	85.2 (21.1)	86.8 (22.0)
	Total	71.7 (22.4)	83.9 (21.6)	87.1 (21.5)	86.6 (21.8)	88.2 (22.1)
QOLI-20 global item						
Mean (sd)						
	HF	3.5 (1.8)	4.3 (1.7)	4.5 (1.7)	4.5 (1.7)	4.7 (1.7)
	TAU	3.5 (1.8)	4.1 (1.7)	4.3 (1.7)	4.3 (1.7)	4.5 (1.6)
	Total	3.5 (1.8)	4.2 (1.7)	4.4 (1.7)	4.4 (1.7)	4.6 (1.6)

		Baseline *	6 months	12 months	18 months	24 months
Global item category						
Frequency (%)						
Poor (1-3)	HF	520 (48.9)	276 (28.7)	242 (24.8)	230 (23.8)	201 (20.1)
	TAU	356 (48.5)	205 (33.9)	189 (29.4)	191 (29.3)	140 (20.9)
	Total	876 (48.7)	481 (30.7)	431 (26.7)	421 (26.0)	341 (20.5)
Fair (4-5)	HF	361 (33.9)	427 (44.3)	432 (44.3)	432 (44.7)	449 (45.0)
	TAU	259 (35.3)	260 (43.1)	298 (46.4)	299 (45.8)	336 (50.2)
	Total	620 (34.5)	687 (43.8)	730 (45.2)	731 (45.2)	785 (47.1)
Good (6-7)	HF	183 (17.2)	260 (27.0)	301 (30.9)	304 (31.5)	348 (34.9)
	TAU	119 (16.2)	139 (23.0)	155 (24.1)	163 (25.0)	193 (28.9)
	Total	302 (16.8)	399 (25.5)	456 (28.2)	467 (28.8)	541 (32.5)

5.2 Explanatory Model

Baseline descriptive statistics of demographic variables and all variables included in the explanatory model are presented in Table 2. The results of the explanatory model are presented in Table

3. A total of 23 covariates were selected for the explanatory model based on statistical significance (p<.1), the pseudo-$R^{2,}$ and AIC statistics. They included 9 covariates measured at baseline or fixed at their baseline value to prevent response shift in covariates; age, total physical comorbidities, adverse childhood experiences, baseline RAS score, baseline SF-12 general health, and baseline measures of QOLI-20 sub-domains of social relationships, family, leisure and finances. Fourteen different time-varying covariates were included: stable housing, assessed psychosis/though disorder, assessed mood abnormality, self-reported anxiety, cognitive trouble, feelings of self-harm/suicide, difficulty fitting in, unmet health need, current employment, past-month income, personal safety, victimization, substance problems, and the absence of a personal confidante.

The final model explained 61.8% of variance in the QOLI-20 global item (AIC = 23746.4). A total of 272 individuals had one or more observations dropped (1,128 observations in total) from the explanatory model due to missing data on at least one covariate.

5.3 Standard Deviation (*sd*) of Residuals

The standard deviation of residual values was calculated for 1812 participants (64 participants had fewer than 2 observations sampled in the explanatory model). The mean *sd* of residuals was 0.95 (0.53), from a range between 0.00 and 3.93. The HF participants' mean *sd* of residuals (mean = 0.94, *sd* = .52, *n* =1073) was not significantly different from that of TAU participants (mean = 0.96, *sd* = 0.55, *n* = 739).

Table 2. Baseline demographic and descriptive statistics of participants included in explanatory model

Variable		HF		TAU		Total	
Need level assigned, freq.(%)	High need	429	(38.9)	322	(41.6)	751	(40.0)
	Moderate need	673	(61.1)	452	(58.4)	1,125	(60.0)
Age, mean(*sd*)	[18.1 to 71.6]	40.9	(11.0)	41.0	(11.2)	40.9	(11.0)
Gender, freq.(%)	Male	731	(66.3)	526	(68.0)	1,257	(67.0)
	Female	364	(33.0)	239	(30.9)	603	(32.1)
	Other	7	(0.6)	9	(1.2)	16	(0.9)
Ethnoracial, freq.(%)	No	843	(76.5)	585	(75.6)	1,428	(76.1)
	Yes	259	(23.5)	189	(24.4)	448	(23.9)
Past month income, mean(*sd*)	[0 to 9987]	701	(653)	680	(675)	692	(662)
Currently working, freq.(%)	No	1062	(96.4)	745	(96.3)	1807	(96.3)
	Yes	40	(3.6)	29	(3.7)	69	(3.7)
Education, freq.(%)	Less than high school	629	(57.3)	417	(54.1)	1,046	(56.0)
	High school, no university	413	(37.6)	310	(40.2)	723	(38.7)
	Bachelor or more	55	(5.0)	44	(5.7)	99	(5.3)
Stable housing (% past 3mo), freq.(%)	Never	869	(83.2)	624	(85.5)	1493	(84.2)
	< 50% of days	79	(7.6)	49	(6.7)	128	(7.2)
	>50% of days	55	(5.3)	28	(3.8)	83	(4.7)
	Always	41	(3.9)	29	(4)	70	(3.9)
Chronic medical conditions, mean(*sd*)	[0 to 19]	4.8	(3.5)	4.7	(3.4)	4.7	(3.4)
Unmet health need, freq.(%)	No	584	(53.5)	395	(51.8)	979	(52.8)
	Yes	508	(46.5)	368	(48.2)	876	(47.2)
Self-reported general health, mean(*sd*)	1 (poor) to 5 (excellent)	3.5	(1.1)	3.4	(1.1)	3.4	(1.1)
Adverse childhood experiences, mean(*sd*)	[1 to 10]	4.5	(3.0)	4.3	(3.0)	4.4	(3.0)
Recovery Assessment Score, mean(sd)	[1 to 5]	3.6	(0.6)	3.6	(0.6)	3.6	(0.6)
Anxiety (past month), mean(*sd*)	[1 (never) to 5 (≥daily)]	3.8	(1.2)	3.8	(1.2)	3.8	(1.2)
Cognitive trouble (past month), mean(*sd*)	[1 (never) to 5 (≥daily)]	3.3	(1.4)	3.4	(1.4)	3.3	(1.4)
Difficulty fitting in (past month), mean(*sd*)	[1 (never) to 5 (≥daily)]	3.2	(1.5)	3.2	(1.5)	3.2	(1.5)
Feelings of self-harm/suicide (past month), mean(*sd*)	Never	692	(63.7)	467	(61)	1159	(62.5)
	≥daily	395	(36.3)	299	(39)	694	(37.5)
Assessed psychosis/thought disorder, freq.(%)	Normal	390	(35.4)	233	(30.1)	623	(33.2)
	Slight	346	(31.4)	269	(34.8)	615	(32.8)
	Moderate	263	(23.9)	196	(25.3)	459	(24.5)
	Marked/extreme	103	(9.3)	76	(9.8)	179	(9.5)
Assessed mood abnormality, freq.(%)	Normal	57	(5.2)	36	(4.7)	93	(5)
	Slight	346	(31.4)	273	(35.3)	619	(33)
	Moderate	546	(49.5)	344	(44.4)	890	(47.4)
	Marked/extreme	153	(13.9)	121	(15.6)	274	(14.6)
# of substance problems, mean(*sd*)	[0 to 5]	1.8	(1.9)	1.9	(1.9)	1.8	(1.9)
Victim of a crime, freq.(%)	No	473	(45.4)	287	(39.8)	760	(43.1)
	Yes	569	(54.6)	435	(60.2)	1004	(56.9)
Social QOL domain, mean(*sd*)	[3 to 21] (3 items)	12.4	(4.2)	12.3	(4.2)	12.3	(4.2)
Family QOL domain, mean(*sd*)	[4 to 28] (4 items)	13.6	(6.5)	13.9	(6.5)	13.8	(6.5)
Leisure QOL domain, mean(*sd*)	[5 to 35] (5 items)	18.6	(6.9)	18.7	(7.2)	18.7	(7.0)
Finances QOL domain, mean(*sd*)	[2 to 14] (2 items)	5.3	(3.2)	5.1	(3.1)	5.2	(3.1)
Personal safety (QOL item), mean(*sd*)	[1 to 7]	4.1	(1.8)	4.1	(1.8)	4.1	(1.8)
Having a close confidante, freq.(%)	Yes	545	(49.9)	397	(51.6)	942	(50.6)
	No	548	(50.1)	373	(48.4)	921	(49.4)
Total		**1,102**	**(100.0)**	**774**	**(100.0)**	**1,876**	**(100.0)**

Table 3. Coefficients Estimates of Explanatory Model of QOLI-20 Global Item

Random intercept model		Coefficient	*SE*	*p*-val	95% Conf. Int.
Covariates measured or fixed at baseline					
Age		-0.005**	0.002	.001	-.009, -.002
Adverse childhood experiences		0.012*	0.006	.050	.000, .024
Chronic medical conditions		0.026**	0.006	.000	.015, .037
Social QOL domain		0.064**	0.004	.000	.056, .073
Family QOL domain		0.018**	0.002	.000	.013, .023
Leisure QOL domain		0.073**	0.002	.001	-.009, -.002
Finances QOL domain		0.057**	0.005	.000	.047, .067
Self-reported general health		0.103**	0.015	.000	.075, .132
Recovery Assessment Score		0.233**	0.029	.000	.177, .290
Time-varying covariates					
Stable housing (% past 3mo)	*Never*	-	-	-	-
	< 50% of days	0.104[a]	0.055	.056	-.003, .211
	>50% of days	0.096[a]	0.051	.060	-.004, .195
	Always	0.148**	0.033	.000	.083, .213
Unmet health need		-0.080*	0.032	.011	-.142, -.019
Psych. symptoms freq. score (CSI)					
Anxiety		-0.060**	0.013	.000	-.086, -.034
Cognitive trouble		-0.024[a]	0.012	.055	-.048, .000
Difficulty fitting in		-0.053**	0.011	.000	-.076, -.031
Feelings of self-harm/suicide		-0.211**	0.036	.000	-.282, -.140
Assessed Psychosis/ thought disorder	*None*	-	-	-	-
	Slight	0.062[a]	0.033	.056	-.002, .126
	Moderate	0.183**	0.041	.000	.103, .263
	Marked/Extreme	0.274**	0.074	.000	.128, .420
Assessed mood abnormality	*None*	-	-	-	-
	Slight	-0.112**	0.042	.008	-.194, -.029
	Moderate	-0.275**	0.047	.000	-.368, -.183
	Marked/Extreme	-0.410**	0.066	.000	-.540, -.280
Personal safety		0.094**	0.010	.000	.074, .113
Victim of a crime (past 6 mo; any vs never)		0.063*	0.031	.041	.003, .123
Past month income ($1000 units)	Spline ≤$1800	0.091*	0.0386	.019	.000, .000
	Spline >$1800	0.016[a]	0.0329	.063	.000, .000
Currently working (vs not working)		0.117*	0.046	.011	.027, .207
# of substance problems		-0.026**	0.009	.003	-.043, -.009
No close confidante (vs close confidante)		-0.118**	0.030	.000	-.177, -.059
	Intercept	0.567	0.179	0.002	.216, .918
	var(intercept)	0.219	0.018		.186, .258
	var(residuals)	1.178	0.022		1.135, 1.222
Observations: 7,483, # of Individuals: 1,876					
AIC: 23746.4 Pseudo R^2: .618					

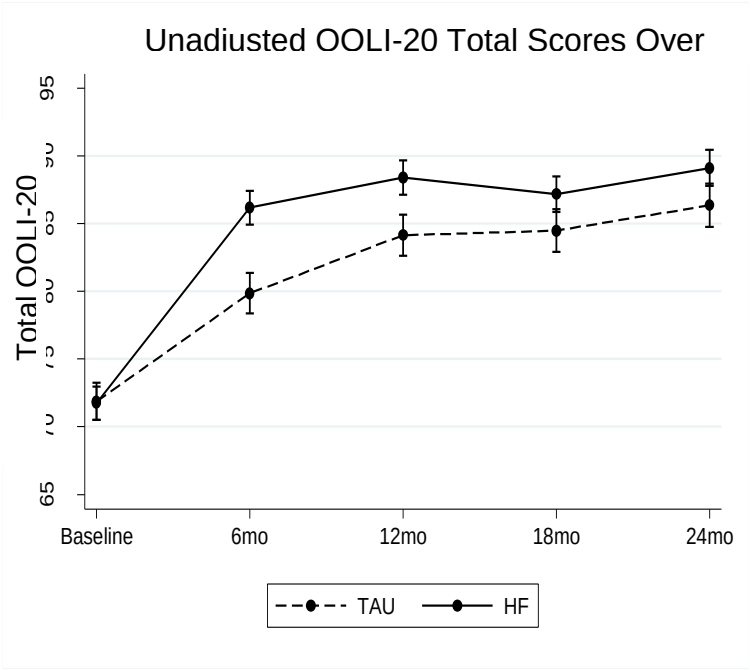
[a] $p < .1$, *$p < .05$, ** $p < .01$

5.4 Response-Shift-Adjusted QOL Analysis

The response-shift-adjusted QOLI-20 total score analysis suggested that the HF treatment effect diminished somewhat consistently as a function of *sd* of residuals. Comparing total scores at higher values of *sd* of residuals (1-point increments) decreased the HF treatment effect at 6 months by 3.97 points (*SE* = 2.04, *p* = 0.052, 95% CI = -7.97, 0.03), at 12 months by 3.20 points (*SE* = 1.99, *p* = .109, 95% CI = -7.1, 0.71), and at 18 months by 4.15 points (*SE* = 2.05, *p* = 0.043, 95% CI = -8.17, -1.13). At 24 months, however, the HF treatment showed little effect, increasing non-significantly by 0.14 points (*SE* = 2.1, *p* = 0.946, 95% CI = -3.97, 4.26) at higher values of *sd* of residuals. Also worth noting was a significant decrease (β = -3.89, *SE* = 1.96, *p* = .047, 95% CI = -7.74, -0.05) in baseline differences between HF and TAU groups at higher values of *sd* of residuals.

However, this effect corresponded to estimated baseline group differences that were only significant for the highest decile of the *sd* of residuals range. At an *sd* of residuals of 1.65 (the 90th percentile) the HF group had QOLI-20 total score 3.36 points lower than the TAU group (*SE* = 1.72, *p* = .051, 95% CI = -6.74, .015). Figure 5 (B) presents the HF treatment effect over time and at different levels of *sd* of residuals (0, 0.5, 1, and 1.5), in comparison to the unadjusted analysis of the full At Home/Chez Soi sample in (A). The analysis was performed on 1,812 participants with a total of 8,302 QOLI-20 total score observations. For complete estimates of main effects and all other interactions see Appendix D.

(A)



(B)

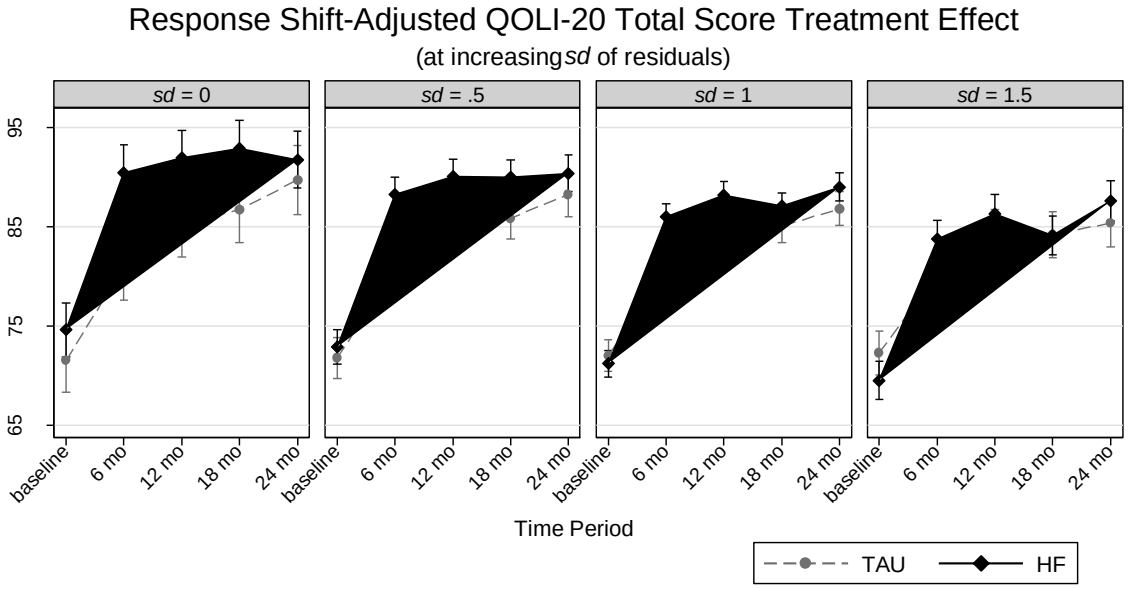


Figure 5. (A) Unadjusted compared to (B) response-shift-adjusted QOLI-20 total score Housing First treatment effect. The Housing First treatment effect over time is weakened by increasing person-specific *sd* of residuals, interpreted as the response shift effect on the outcome. For all time periods except 24 months, the group differences decrease as individuals deviate from the true change, or totally stable residuals, represented by a *sd* of 0.

In terms of the response-shift-adjusted QOLI-20 global item analysis the HF treatment did decrease as a function of *sd* of residuals, but this interaction was not statistically significant. As *sd* of residuals increased, the HF treatment odds ratio (OR) decreased by a factor of 0.83 (*SE* = 1.15, *p* = 0.175, 95% CI = 0.63, 1.09). Proportional odds across both levels of the ordinal logit equation were estimated for the main effects of baseline QOL category (OR= 2.66, *SE* = 1.05, *p* < .001, 95% CI = 2.43, 2.92), HF treatment (OR = 1.58, *SE* = 1.16, *p* = .002, 95% CI = 1.17, 2.11), and the HF treatment x *sd* of residuals interaction, but not for the main effect of *sd* of residuals (at poor vs fair or good: OR = 0.69, *SE* = 1.12, *p* = .001, 95% CI = 0.55, 0.86); at poor or fair vs good: OR = 1.40, *SE* = 1.12, *p* = .003, 95% CI = 1.12, 1.73). Figure 6 presents the estimated treatment effect (on the log-odds scale) at different levels of *sd* of residuals (0, 0.5, 1, and 1.5). The analysis was performed on 1,798 participants (14 participants had were missing QOLI-20 global item observations) with a total 6,470 post-baseline observations. For complete estimates of main effects and all other interactions see Appendix E.

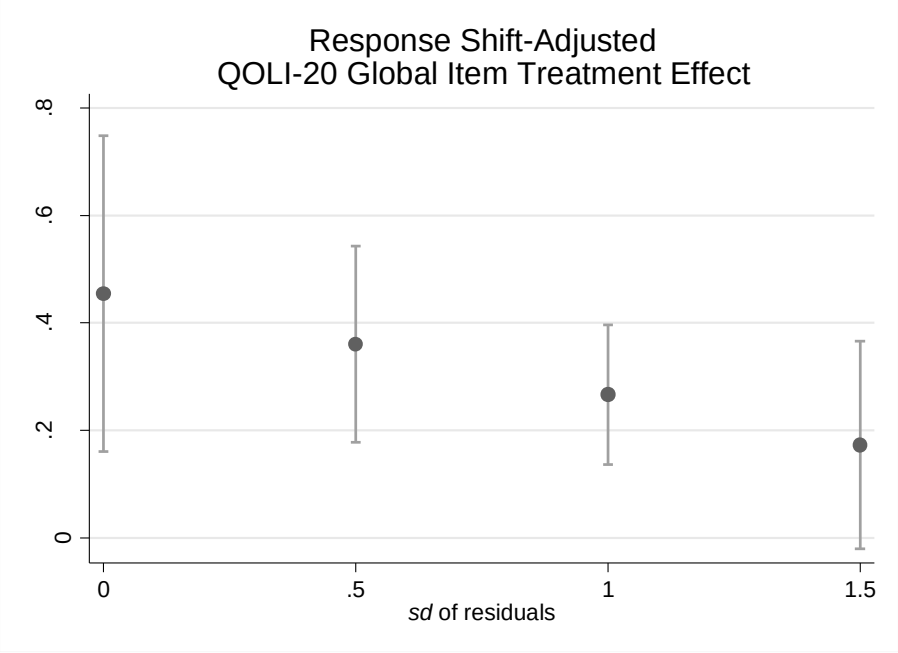


Figure 6. Response-shift-adjusted QOLI-20 global item Housing First treatment effect (log-odds ratio for higher QOLI-20 global item ratings). The treatment log-odds ratio decreases modestly with increasing values of person-specific *sd* of residuals (at 0, 0.5, 1, and 1.5). Estimates are graphed on the log-odds scale for linearity of effects

5.5 Sensitivity Analyses

The exclusion of observations based on the validity and sincerity of responses resulted in similar parameters estimates (criteria A: R^2 = .626; criterion B: R^2 = .636) as the inclusive explanatory model. The resulting *sd* of residuals under criterion A (mean = 0.93, *sd* = 0.55) were similar to those of the inclusive analysis, but were somewhat smaller under criterion B (mean = 0.88, *sd* = 0.56). Criterion A restricted the analysis to 1,692 individuals for 6,951 observations, while under criterion B only 1,164 participants for 3,878 observations were included.

The exclusion criteria had different impacts on each of the QOLI-20 total score and global item analyses.

For the QOLI-20 total score analysis, the HF treatment x *sd* of residuals interactions no longer met significance levels as exclusion criteria were applied, but maintained similar estimates as those seen in the inclusive analysis. Figure 7 presents the impact of the exclusion criteria on the QOLI-20 total score outcomes. Under criterion A, the HF treatment effect mostly decreased with increasing *sd* of residuals, though never significantly. This decrease was seen at baseline (β = -2.40, *SE* = 2.04, *p* = 0.24, 95% CI = -6.40, 1.60), at 6 months (β = -2.71, *SE* =2.15, *p* = 0.206, 95% CI = -6.92, 1.49), at 12 months (β = -2.38, *SE* = 2.02, *p* = 0.24, 95% CI = -6.35, 1.59), and at 18 months (β = -2.03, *SE* = 2.09, *p* = 0.332, 95% CI = -6.13, 2.07), but not at at 24 months where again little interaction occurred (β = 0.24, *SE* = 2.20, *p* = 0.912, 95% CI = -4.07, 4.56). Under the more stringent criterion B, the interaction revealed decreasing HF treatment effects with increasing *sd* of residuals at all time periods, but the interaction estimates were again non-significant throughout: at baseline (β = -4.62, *SE* =2.74, *p* =.092, 95% CI =-9.99, 0.75), 6 months (β = -1.81, *SE* =2.52, *p* =.472, 95% CI =-6.74, 3.12), 12 months (β = -2.80, *SE* =2.44, *p* =.251, 95% CI =-7.59, 1.98), 18 months (β = -3.14, *SE* =2.46, *p* =.203, 95% CI =-7.96, 1.69), and at 24 months (β = -1.07, *SE* =2.69, *p* =.689,

95% CI =-6.34, 4.19). That these estimates of interaction coefficients mirror the inclusive analysis in direction but not in significance may be the result of the reduction in sample sizes after exclusion.

Like the results of the QOLI-20 total score sensitivity analysis, the exclusion criterion A analysis of the QOLI-20 global item yielded similar estimates as the inclusive analysis. However criterion B showed a stronger interaction than both the inclusive analysis and criterion A analysis, despite lower power. Figure 8 presents the impact of exclusion criteria on QOLI-20 global item outcomes. Under criterion A, an increase in the *sd* of residuals decreased the HF treatment odds ratio (OR) by the same factor as the inclusive analysis (OR = 0.83, *SE* = 1.16, *p* = 0.21, 95% CI = 0.63, 1.10). In contrast, under the more stringent criterion B, the *sd* of residuals decreased the HF treatment effect to a greater extent, proving statistically significant (OR = 0.69, *SE* = 1.18, *p* = 0.027, 95% CI = 0.49, 0.96). Similar to the inclusive analysis, the coefficients under both criteria were proportional at both contrast levels of the order logit equation, except for a similar reversal in direction of the main effect of *sd* of residuals. Again, model sample sizes diminished under exclusion criteria, with criterion A restricting the global item analysis to 1,680 individuals and 5,490 observation and criterion B restricting the analysis to 1,158 participants and 3,204 observations. See Appendix F for complete sets of model estimates.

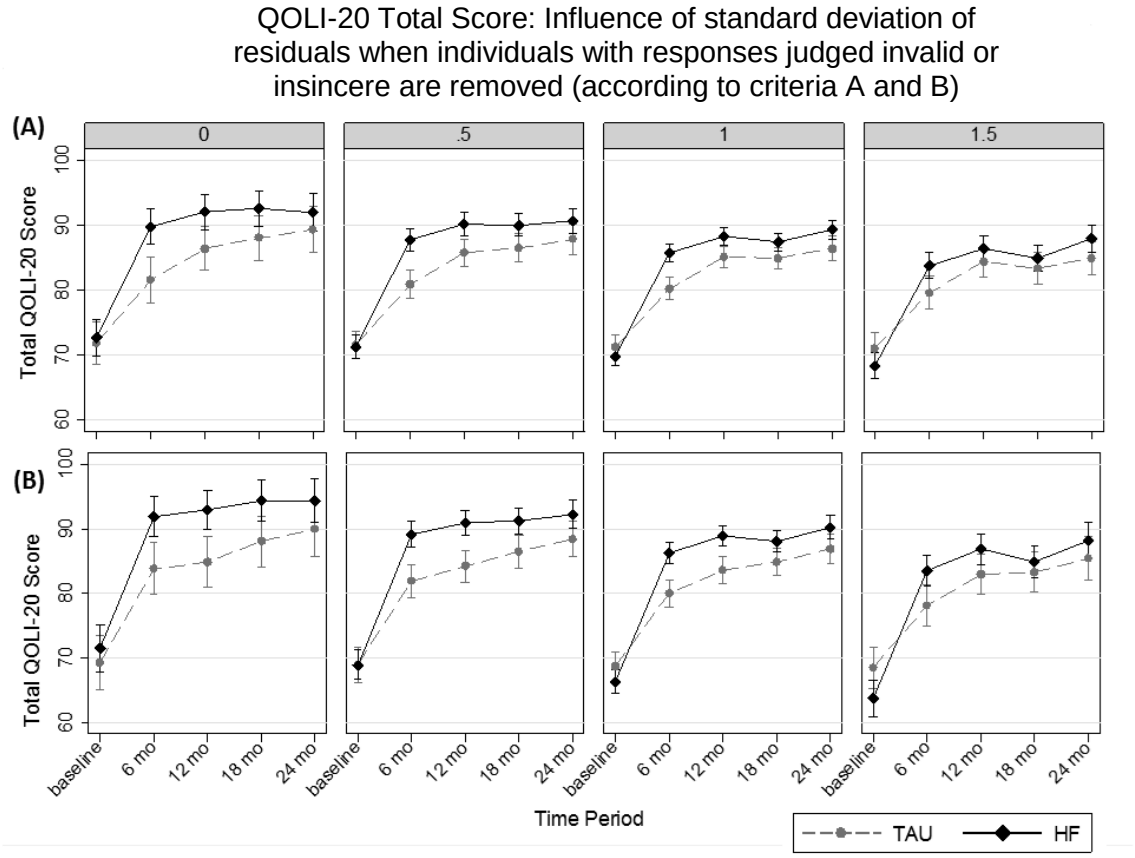


Figure 7. The impact of the sensitivity analysis on the QOLI-20 total score Housing First treatment effect. By excluding responses under both validity and sincerity criteria (A) "complete confidence" or "some doubts" in validity of responses and score 4 or 5 for sincerity/truthfulness and criteria (B) "complete confidence" and sincerity/truthfulness score of 5 only, treatment effects resemble those of the inclusive analysis, decreasing further as *sd* of residuals increase. However, the decrease in treatment effects over *sd* of residuals is no longer significant.

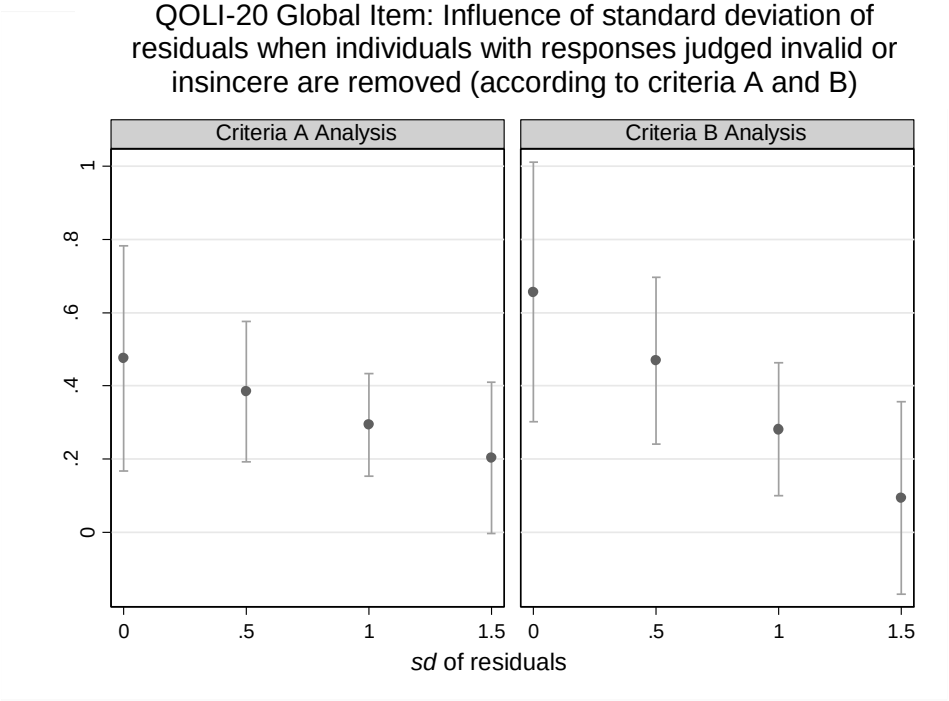


*Figure 8***.** The impact of the sensitivity analysis on the QOLI-20 global item Housing First treatment effect (log-odds ratio for higher QOLI-20 global item rating). By excluding responses under validity and sincerity criterion A ("complete confidence" or "some doubts" in validity of responses and score 4 or 5 for sincerity/truthfulness), treatment effects resemble the inclusive analysis, while under criterion B ("complete confidence" and sincerity/truthfulness score of 5 only) treatment effects decrease further as *sd* of residuals increase (a significant response shift adjustment).

CHAPTER 6: Discussion

This study examined the influence of response shift on QOL treatment effects in the At Home/Chez Soi project, suggesting that modest QOL improvements in the intervention group may be partly attributable to response shift. The analysis adjusted originally reported QOL comparisons to reveal mostly larger HF treatment effects in participants showing more true change than response shift.

6.1 Summary of results

Response shift was inferred from the degree of fluctuation in a participant's residual values, based on the *sd* of their residuals obtained from a model explaining substantial variance in QOLI-20 global item ratings. The effects of response shift adjustment were seen both in the total score of the QOLI-20 and the instrument's global item. The HF treatment effect on the QOLI-20 total score diminished with increasing *sd* of residuals across most time periods (a non-significant decrease was seen at 12 months). At 24 months, however, the HF treatment effect did not change as a function of *sd* of residuals. Analyses of the QOLI-20 global item ratings (estimating the effect of treatment on odds of a higher rating, collapsed across follow-up periods) suggested that the HF treatment odds ratio (OR) also decreased as a function of *sd* of residuals, though this interaction was not statistically significance.

Under two exclusion criteria based on interviewer-assessed validity and sincerity of observations, sensitivity analyses yielded similar estimates as the inclusive QOLI-20 total score analysis, though the interaction of treatment x *sd* of residuals was no longer significant. The global item analysis however showed a stronger and statistically significant estimate of the adjustment effect estimate under the more stringent of the two exclusion criteria. Under the less stringent criterion, the analysis of the global item showed similar estimates as the inclusive analysis without statistical significance.

6.2 How the results compare to the literature

The study was the first to examine and adjust for response shift in QOL outcomes in homeless individuals with mental illness. Previous research by Evans and Huxley [7] found similar patterns in both healthy and mentally ill individuals who were however stably housed. More specifically, their findings indicate that response shift, inferred from an indicator of adaptation to change in circumstance, is related to either higher or lower QOL ratings compared to individuals showing no response shift, depending on the domain of QOL analyzed or direction of the observed response shift. The effects of the two response shift directions aggregated to a negative but small net effect of response shift in QOL ratings. Without the capacity for direct assessment of such adaptation, the current analysis found a similar, though perhaps larger, negative net response shift effect. Given the nature of the current study, the negative effects of response shift represent a diminished treatment effect, whereas this effect consisted of a decrease in average ratings in Evans and Huxley[7].

These results help explain some of the smaller than expected QOL improvements in the At Home/Chez Soi project, but may also shed light on equivocal QOL findings in other HF interventions. To date, only studies of HF interventions without concurrent comparisons to controls report any positive effects on QOL [1,3,85]. Only one other study by Mares and colleagues [2] compared changes in QOL among HF recipients to QOL changes in controls and did not find any evidence for a positive HF treatment effect. What the current response shift adjustment achieves in terms of explaining unexpectedly modest results may indeed be relevant to other studies with unexpected null QOL findings, as seen in Mares and colleagues [2].

These findings also highlight the value of secondary analysis of response shift at the individual level in comparison to the group level. An analysis of schizophrenia patients identified different reprioritization in mental and physical determinants of self-reported health, using a random forest

method (bootstrapping of classification and regression trees)[92]. However, such response shift is inferred from changes in mean squared error (MSE) of all predictions from random forests. The iterative nature of random forests partly controls for random error in predictions, but also limits response shift differentiation, thereby limiting adjustment, among individuals.

Furthermore, the results as well as the methodology of this study add to the current body of research investigating response shift through residual values from an explanatory random effects model. The only other response shift study to also examine residuals in a randomized controlled trial (with stroke patients) did not however perform any adjustment for response shift in estimates of treatment effects [97]. The authors did report similarities in the patterns of residuals in either treatment or control groups, consistent with the equivalent between-group distributions of *sd* of residuals in the current study.

As in previous response shift research examining residuals, the current analysis identified a predominance of no or little response shift among participants. The majority of participants show only small fluctuation of residual values reflected in a median *sd* of 0.88. In other studies using LTA, participants fitting the stable residual group made up between 67% of 387 stroke patients[10], 82% of 359 IBD patients[96], and 99.7% of 1566 multiple sclerosis patients[95].

6.3 Strengths and Limitations of Methodology

This identification of individuals with minimal residual fluctuation, or no response shift, may be the most essential result of response shift analysis within a randomized controlled trial. Groups that do show response shift may reflect significant heterogeneity in response shift catalysts and time of onset, directions and magnitude, or response shift mechanisms. Stable residuals, on the other hand, have a less ambiguous interpretation: they suggest true change in self-reported outcomes, or at least change

that only deviates from expectations in a consistent manner. Accurately identifying such true change can be valuable for estimating treatment effects while safeguarding from concerns of reconceptualization, reprioritization, or recalibration of outcome measures. The *sd* of residuals, though a cruder metric that does not differentiate between forms of response shift trajectories, is more sensitive than LTA to any discrepancy from true change. For example, latent grouping introduces some uncertainty in gathering participants into a stable residual group based on relative probability of fit. Sorting by *sd* of residuals instead assigns a specific measure of fluctuation for each individual, rather than a probable latent measure of fluctuation.

Additionally, the potential heterogeneity behind various response shift trajectories would require comparisons of treatment effects within large numbers of response shift subgroups, incurring both problems associated with power (type II errors) and multiple comparisons (type I error) [109,110]. As a common metric of residual fluctuation across participants, the *sd* of residuals allows for formal tests of interaction that more closely follow subgroup analysis guidelines for trials [110].

Despite model constraints that facilitate the emergence of response shift in residual patterns (the absence of a time variable, response shift-prone variables fixed at baseline, and interviewer-assessed covariates), residual patterns will invariably represent other unexplained changes in QOL beyond response shift. The advantage of LTA methods is that common patterns of residuals shared among many participants suggest these patterns are less likely to results from random error. In addition, unlike the *sd* of residuals, LTA can make conclusion about the patterns of response shift groups that do not show excessive variation (positive, negative, “rebounding” response shift, etc.)[96]. In fact, Mayo and colleagues [10] avoided conclusion about an additional trajectory group representing excessive variation of residuals that did not follow an intelligible response shift pattern (a highly non-monotonic pattern).

Response shift inferences from the *sd* of residuals is then likely to also represent some of the additional variation not attributed to response shift in an LTA method. This variation may represent the effects of either random error or erratic responses. On the other hand, more variation in response shift patterns may be expected in a sample with mental illness and substance use problems. The cyclical pattern and affective or cognitive impact of their symptoms may initiate response shift processes without the same monotonicity as in physical illnesses.

However, a complementary literature on QOL provides another perspective on the unexplained variance captured in residual patterns. While the deviation from the stable residuals is inferred as response shift in this study (while acknowledging the potential impact of random error), fluctuation in residuals may reflect a volatility in QOL appraisal outside of the response shift framework. Several experiments present results that suggest QOL appraisal does not reliably follow the theorized mechanisms of models in QOL research [5,6,40]. One study has demonstrated that participants' responses to a health-related QOL measure, the EQ-VAS, can be manipulated by framing appraisal with written descriptions of extreme health states (both extremely good and extremely poor)[111]. Other experiments in social psychology have demonstrated that global QOL ratings (or "overall subjective well-being") are impacted by seemingly more transient factors than those in theoretical models of QOL. In one experiment, global QOL ratings were significantly higher for participants whose interviews were conducted in a pleasant room than for those randomized to interviews taking place in a small, poorly-lit storage room[112]. This condition did not however significantly impact domain-specific QOL ratings (satisfaction with housing). In another experiment, higher global QOL ratings were given by participants who were interviewed on sunny rather than rainy days [113]. Despite the fact that none of these experiments impacted the "overall" life experience of these participants, their ratings of QOL were influenced by the way questions were asked (i.e., health state framing), as well as where and when they were asked.

If such volatility is common in responses of participants in the At Home/Chez Soi study, the methodology of current analysis may also be capturing this effect on QOL ratings. Participants whose responses are subject to such transient contexts are also likely to show fluctuation in residuals derived from an explanatory model of QOL. Though a similar concept, response shift is more narrowly defined by a process in which a catalyst triggers a change in appraisal, involving the different types of response shift mechanisms. Instead, these changes in appraisal, and their impacts on QOL ratings, may be occurring as a results of several determinants not traditionally considered as response shift catalysts.

Each one of these alternative interpretations of residual patterns, random error or erratic responses, represents a limitation of any methodology that seeks to infer response shift from residuals. However, the ability to isolate participants with stable residual remains a likely method to more accurately assess true change in QOL ratings, regardless of what the deviation from this true change represents.

6.4 Explanatory Model of QOL

Whatever the comparisons generated from residual patterns, the validity of how residuals are interpreted is also contingent on a well-developed explanatory model. The current analysis chose covariates that cover several domains of QOL determinants, following a broad review of theoretical models of QOL in populations with mental illness. Beyond the large sample size and multiple follow-up periods, the study benefitted from the extensive variety of data collected for the At Home/Chez Soi project, both in terms of interviewer-assessed measures and self-reported measures that are not vulnerable to response shift. Furthermore, some improvements to the fit of the explanatory model resulted from adaptations to the original methodology in Mayo et al. [10]. Namely, the model included covariates whose effects are modified by time, as the contribution of their effect across time is

preferable to ignoring their effect altogether. Similarly, response shift-prone covariates were included, but fixed at their baseline values. This circumvents the trade-off between response shift influencing estimates of expected QOL ratings and excessive unexplained variance from ignoring these covariates.

The explanatory model step of the analysis also contributes additional findings to the literature on QOL in mental illness. However, any comparison to other published QOL analyses should be done with caution, given this model did not have a conventional aim of simply explaining QOL changes, but rather sought to identify response shift in the residuals. This meant that many important covariates were not included or the effects of their baseline values were estimated.

The nature of covariates in the multivariate model bear similarities to those that are reliably identified in the literature. In terms of personal characteristics, only one demographic variable was significantly associated with QOL (age), while far more measures of psychiatric symptoms were included as significant predictors (self-reported anxiety, cognitive trouble, difficulty fitting in, feelings of self-harm/suicide, number of substance problems as well as interviewer-assessed psychosis/thought disorder and mood abnormality). The importance of such psychiatric factors is a common finding in almost all QOL studies in a population with mental illness, whether homeless [37,73] or housed[41,47–49,51].

Given the similarity of the perspective from which they are reported, subjective measures were consistently strong predictors of QOL (baseline measures of social, family, leisure, and finances QOL domains, recovery assessment scale total score, and self-reported general health as well time-varying measures of unmet health need and personal safety). For several of the covariates restricted to baseline, associations across time could not be estimated given the assumption that response shift would occur similarly between the outcome and the covariate. Nonetheless, these associations support the emphasis reported in the QOL literature on such subjective measures[28,36,43,60,73].

Furthermore, objective indicators of QOL were also identified as predictors of QOL (chronic medical conditions at baseline, stable housing, victimization, employment, income, and the absence a

close confidante). As in previous literature, these covariates were however generally less significant in explaining QOL than psychiatric symptoms and subjective measures.

Another important element of QOL modelling in the literature is the mediational role of self-related constructs (autonomy, self-esteem, mastery) between other covariates and QOL. However, few measures in the study captured these constructs. Were there more such instruments available, the current model would have excluded them anyway to avoid explaining response shift in the outcome, rather than enabling response shift to appear in residual patterns.

6.5 Validity of QOL measurement

A concern of the current study is that the assessment of QOL in a homeless population with mental illness may not validly measure QOL, due to affective and cognitive biases in the sample. Although the ample literature in this population suggests otherwise, it is understandable that one would be apprehensive to giving too much weight to QOL ratings. Homelessness often means that the effects of psychotic episode and heavy intoxication are visible to the general public outside of psychiatric settings. For this reason, the validity of QOLI-20 global item ratings was assessed through spearman correlations at baseline with measures that were similar, though not equivalent, to the construct of QOL.

The item's strongest correlation was seen with the QOLI-20 total score, validating the global item's aim of capturing life satisfaction in all relevant domains. More moderate results were seen in terms of correlations with items that capture similar self-reported constructs (physical, mental, overall/general health from EQ-5D and the SF-12), with the RAS item of "I like myself" showing the highest association. These latter associations were not as strong as expected, but considering their similarity to previous literature [28,36,40] and their consistent statistical significance, they are still reassuring in answering questions of QOL measurement validity.

6.6 Positive Effects of Psychosis on QOL

Despite seemingly acceptable validity of the QOL measure in this population, one finding of the explanatory model is worth addressing. In the final multivariate model (as well as associations examined bivariately), greater severity in interviewer-assessed psychosis/thought disorder seemed to be a predictor of *higher* QOL.

A similar example of unexpected QOL results was seen in a study by Van der Plas and colleagues[76], where schizophrenia patients who were homeless rated their QOL in the health domain higher than those who were housed. In other research participants with schizophrenia reported higher QOL ratings in comparison to those of other diagnostic groups, such as depression and borderline personality disorder[41]. Similar findings in research by Atkinson and colleagues[54], lead the other to conclude that higher QOL ratings among individuals with schizophrenia likely reflects poor insight and should lead to questions of whether QOL measures are appropriate in this population. On the other hand, it is possible that higher QOL is observed in patients with schizophrenia mainly because they are often compared to individuals with affective disorders such as depression, which is a strong determinant of lower QOL.

Regardless of any consistencies with other literature, several studies' results still suggest that greater psychosis symptoms are a predictor of lower QOL ratings in housed [49] and homeless participants with schizophrenia[81]. It is unclear whether the method in which psychosis symptoms is assessed explains the direction of its effect on QOL. While self-reported symptoms of psychosis were used in Lam and Rosenheck[81], the Eack and Newhill study[49] was a meta-analysis across papers with different methods. It is likely that the type of psychosis capture though an interviewer assessment of symptoms is not the same as that of a self-reported assessment, with each suggesting opposite QOL effects.

Fortunately, the current study's sensitivity analysis confirm that the positive psychosis effects on QOL have some value, considering that they remain positive following the exclusion of responses deemed to have little validity or sincerity. Preliminary tests were also performed to ensure that the effects of other covariates on QOL were not meaningfully different as a function of psychosis level, revealing that psychosis did not largely impact these associations (based on visual inspection of tested interactions).

6.7 Future Studies

Some of the concern about the residual values representing random error, and thereby random patterns, may be mitigated by other methods that would produce expected QOL values averaging over a large sample of possible models. Future studies may expand on the modeling methods developed in LTA studies and replicated here by using bootstrapping methods, Bayesian model averaging (BMA), or a random forest methods. These methods would produce multiple residual values per observation, requiring some form of averaging in order to measure residual patterns over time within individuals. Though perhaps adding more levels of uncertainty in the response shift metric, such residual patterns are less likely to have resulted from chance due to the iterative estimates. For example, participants whose residuals are consistent over a varying set of models, rather than showing such a pattern from a single estimated model, may allow stronger inferences that they are showing true change.

Another complementary avenue for future research would be the comparison of these response shift findings with qualitative data collected in the narrative interviews of a subsample of 197 participants[24]. Qualitative analyses may identify various forms of response shift (namely reconceptualization or reprioritization) as well as evidence of the adaptation to circumstance described in Evans and Huxley[7]. The content of these interviews may reveal evidence that some participants have

changed how they appraised their QOL or how they evaluate the importance of new QOL determinants. Response shift identified qualitatively would complement some inferences made from the same participants using the current statistical methods.

Finally, given that the current results suggest that response shift did not explain the lack of group differences at the 24-month assessment, further analyses should investigate what explains this null effect. That is, the HF treatment was not associated with higher QOLI-20 total scores at the study's endpoint, regardless of whether participants showed true change or their residuals deviated drastically from a consistent pattern over time. If indeed this finding is due to the adaptation of HF participants to their new environment, or also the resignation of TAU participants to their circumstance, the current method fails to identify this pattern. The other possibility is that the HF treatment is only effective enough to provide an early QOL improvement that is obtained over time by TAU participants. In this case, the analysis suggests that this early boost is perhaps greater than estimated under the unadjusted analysis, but that TAU improvements do reflect true QOL change.

6.8 Conclusion

Research on individuals with mental illness suggest that the influence of various determinants on QOL may be mediated by how these individuals perceive and understand aspects of themselves, others, and the world around them. Though less extensive, the research on QOL in individuals with mental illness who are also homeless suggests that this population's QOL is determined by similar factors, with housing generally improving their QOL over time.

The results of the At Home/Chez Soi project have already demonstrated that an intervention focused on immediate access to housing with clinical services is a feasible, effective, a cost-offsetting method to improve the lives of homeless individuals suffering from mental illness. However, the

improvements demonstrated in the project's results were mainly seen for the primary outcome of housing stability, whereas QOL measures showed more modest treatment effects.

The findings of the current study suggest that response shift may partly explain why larger treatment effects were not seen. A more conservative conclusion from the results is that QOL improvements from a HF intervention are more evident in individuals for whom reported and expected QOL differ by an amount that is relatively consistent over time. HF treatment effects are smaller in participants whose residual patterns fluctuate over time. This fluctuation is interpreted as response shift, in accordance with previous literature, though alternative interpretations may also explain some of these patterns.

What this implies for the future interventions for homeless individuals with mental illness is unclear. It is not necessarily undesirable for a treatment to initiate response shift, especially if response shift improves QOL ratings beyond what is expect from circumstance. These findings do support other authors in encouraging future researchers to rely on QOL as a measure of a treatment effects, especially in mental illness and homeless. However, the results also highlight the importance of adopting direct methods of assessing response shift during interviews with QOL ratings.

In conclusion, the current analysis is a novel adjustment for analyzing changes in QOL that deviate from expected QOL change, inferred in this study as response shift. Though limited in providing confident interpretations as a response shift metric, the study nonetheless provides a simple tool for refining comparisons of self-reported outcomes for other researchers. The value of such adjustments is greater now than it ever has been, given the increasing importance of outcomes like QOL in various fields and the growing adoption and eventual evaluation of Housing First interventions in different countries

Bibliography

1. Gilmer, T. P., Stefancic, A., Ettner, S. L., Manning, W. G. & Tsemberis, S. Effect of full-service partnerships on homelessness, use and costs of mental health services, and quality of life among adults with serious mental illness. *Arch. Gen. Psychiatry* **67,** 645–652 (2010).
2. Mares, A. S., Rosenheck, R. A. & Director, V. A. HUD/HHS/VA Collaborative Initiative to Help End Chronic Homelessness National Performance Outcomes Assessment Preliminary Client Outcomes Report February 26, 2007. (2007). at <http://158.74.242.66/hsp/homelessness/CICH07/outcomes07/report.pdf>
3. Perlman, J. & Parvensky, J. Denver Housing First Collaborative. *Cost Benefit Anal. Program Outcomes Rep. Colo. Colo. Coalit. Homeless* (2006). at <http://denverbwww.denversroadhome.org/files/FinalDHFCCostStudy_1.pdf>
4. Tabol, C., Drebing, C. & Rosenheck, R. Studies of 'supported' and 'supportive' housing: A comprehensive review of model descriptions and measurement. *Eval. Program Plann.* **33,** 446–456 (2010).
5. Sprangers, M. A. & Schwartz, C. E. Integrating response shift into health-related quality of life research: a theoretical model. *Soc. Sci. Med. 1982* **48,** 1507–1515 (1999).
6. Rapkin, B. D. & Schwartz, C. E. Toward a theoretical model of quality-of-life appraisal: Implications of findings from studies of response shift. *Health Qual. Life Outcomes* **2,** 14 (2004).
7. Evans, S. & Huxley, P. Adaptation, response-shift and quality of life ratings in mentally well and unwell groups. *Qual. Life Res. Int. J. Qual. Life Asp. Treat. Care Rehabil.* **14,** 1719–1732 (2005).
8. Barclay-Goddard, R., Epstein, J. D. & Mayo, N. E. Response shift: a brief overview and proposed research priorities. *Qual. Life Res. Int. J. Qual. Life Asp. Treat. Care Rehabil.* **18,** 335–346 (2009).

9. Schwartz, C. E. *et al.* Guidelines for secondary analysis in search of response shift. *Qual. Life Res. Int. J. Qual. Life Asp. Treat. Care Rehabil.* **22,** 2663–2673 (2013).

10. Mayo, N. E., Scott, S. C., Dendukuri, N., Ahmed, S. & Wood-Dauphinee, S. Identifying response shift statistically at the individual level. *Qual. Life Res. Int. J. Qual. Life Asp. Treat. Care Rehabil.* **17,** 627–639 (2008).

11. Schwartz, C. E. *et al.* Response shift in patients with multiple sclerosis: an application of three statistical techniques. *Qual. Life Res. Int. J. Qual. Life Asp. Treat. Care Rehabil.* **20,** 1561–1572 (2011).

12. Gaetz, S. A., Gullier, T., Richter, T., Marsolais, A. & Homeless Hub. *The state of homelessness in Canada 2014*. (2015). at <http://site.ebrary.com/id/10993043>

13. City of Toronto. *2013 Street Needs Assessment*. (City of Toronto, 2013).

14. Eberle, M. *Vancouver Homeless Count : Final Report*. (Eberle Planning and Research, 2013). at <http://www.homelesshub.ca/resource/vancouver-homeless-count-2013-final-report>

15. University of Winnipeg. Homelessness in Winnipeg. *News - Homelessness Facts* (2014). at <http://www.uwinnipeg.ca/index/news-homeless-facts>

16. Greater Moncton Homelessness Steering Committee. *7th Report Card on Homelessness in Greater Moncton*. (Greater Moncton Homelessness Steering Committee, 2014). at <http://www.monctonhomelessness.org/>

17. Gaetz, S. A. *The state of homelessness in Canada 2013*. (2013). at <http://www.homelesshub.ca/ResourceFiles/SOHC2103.pdf>

18. Laird, G. *Shelter: homelessness in a growth economy : Canada's 21st century paradox : a report for the Sheldon Chumir Foundation for Ethnics in Leadership*. (Sheldon Chumir Foundation for Ethics in Leadership, 2007).

19. Goering, P., Tolomiczenko, G., Sheldon, T., Boydell, K. & Wasylenki, D. Characteristics of persons who are homeless for the first time. *Psychiatr. Serv.* **53,** 1472–1474 (2002).

20. Frankish, C. J., Hwang, S. W. & Quantz, D. Homelessness and health in Canada: research lessons and priorities. *Can. J. Public Health Rev. Can. Santé Publique* **96 Suppl 2,** S23–29 (2005).

21. Nelson, G., Aubry, T. & Hutchison, J. Housing and Mental Health. *International Encyclopedia of Rehabilitation* at <http://cirrie.buffalo.edu/encyclopedia/en/article/132/>

22. Gaetz, S. A., Scott, F. & Gulliver, T. *Housing First in Canada: Supporting Communities to End Homelessness.* (Canadian Homelessness Research Network Press, 2013).

23. Goering, P. N. *et al.* The At Home/Chez Soi trial protocol: a pragmatic, multi-site, randomised controlled trial of a Housing First intervention for homeless individuals with mental illness in five Canadian cities. *BMJ Open* **1,** e000323–e000323 (2011).

24. Goering, P. *et al. National At Home/Chez Soi Final Report*. (Mental Health Commission of Canada, 2014). at <http://www.mentalhealthcommission.ca>

25. Aubry, T. *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.* (2015).

26. Stergiopoulos, V. *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 Trial. *JAMA* **313,** 905 (2015).

27. Polvere, L., Macnaughton, E. & Piat, M. Participant perspectives on housing first and recovery: Early findings from the At Home/Chez Soi project. *Psychiatr. Rehabil. J.* **36,** 110–112 (2013).

28. Lehman, A. F. The well-being of chronic mental patients. *Arch. Gen. Psychiatry* **40,** 369–373 (1983).

29. Uttaro, T. & Lehman, A. Graded response modeling of the Quality of Life Interview. *Eval. Program Plann.* **22,** 41–52 (1999).

30. Evans, S. in *Quality of life impairment in schizophrenia, mood and anxiety disorders new perspectives on research and treatment* (eds. Ritsner, M. S. & Awad, A. G.) (Springer, 2007). at <http://public.eblib.com/choice/publicfullrecord.aspx?p=372606>

31. Mental Health Commission of Canada. *Changing directions, changing lives: the mental health strategy for Canada*. (Mental Health Commission of Canada, 2012).

32. Van Nieuwenhuizen, C., Schene, A. H., Koeter, M. W. J. & Huxley, P. J. The Lancashire Quality of Life Profile: modification and psychometric evaluation. *Soc. Psychiatry Psychiatr. Epidemiol.* **36,** 36–44 (2001).

33. Lehman, A. F. Measures of quality of life among persons with severe and persistent mental disorders. *Soc. Psychiatry Psychiatr. Epidemiol.* **31,** 78–88 (1996).

34. Lehman, A. F. A quality of life interview for the chronically mentally ill. *Eval. Program Plann.* **11,** 51–62 (1988).

35. Van Nieuwenhuizen, C. & Wolf, J. R. L. M. Measuring the quality of life of clients with severe mental illness: a review of instruments. *Br. J. Haematol. - BRIT J HAEMATOL* **20,** 33–42 (1997).

36. Trauer, T., Duckmanton, R. A. & Chiu, E. A Study of the Quality of Life of the Severely Mentally Ill. *Int. J. Soc. Psychiatry* **44,** 79–91 (1998).

37. Lehman, A. F., Rachuba, L. T. & Postrado, L. T. Demographic influences on quality of life among persons with chronic mental illnesses. *Eval. Program Plann.* **18,** 155–164 (1995).

38. Hansson, L. Determinants of quality of life in people with severe mental illness. *Acta Psychiatr. Scand.* **113,** 46–50 (2006).

39. Corrigan, P. W. & Buican, B. The construct validity of subjective quality of life for the severely mentally ill. *J. Nerv. Ment. Dis.* **183,** 281–285 (1995).

40. Zissi, A., Barry, M. M. & Cochrane, R. A mediational model of quality of life for individuals with severe mental health problems. *Psychol. Med.* **28,** 1221–1230 (1998).

41. Schaar, I. & Ojehagen, A. Predictors of improvement in quality of life of severely mentally ill substance abusers during 18 months of co-operation between psychiatric and social services. *Soc. Psychiatry Psychiatr. Epidemiol.* **38,** 83–87 (2003).

42. Holloway, F. & Carson, J. Quality of life in severe mental illness. *Int. Rev. Psychiatry* **14,** 175–184 (2002).

43. Mares, A. S., Young, A. S., Mcguire, J. F. & Rosenheck, R. A. Residential environment and quality of life among seriously mentally ill residents of board and care homes. *Community Ment. Health J.* **38,** 447–458 (2002).

44. Pinikahana, J., Happell, B., Hope, J. & Keks, N. A. Quality of life in schizophrenia: a review of the literature from 1995 to 2000. *Int. J. Ment. Health Nurs.* **11,** 103–111 (2002).

45. Browne, S. *et al.* Determinants of quality of life at first presentation with schizophrenia. *Br. J. Psychiatry J. Ment. Sci.* **176,** 173–176 (2000).

46. Röder-Wanner, U. U., Oliver, J. P. & Priebe, S. Does quality of life differ in schizophrenic women and men? An empirical study. *Int. J. Soc. Psychiatry* **43,** 129–143 (1997).

47. Holloway, F. & Carson, J. Subjective quality of life, psychopathology, satisfaction with care and insight: an exploratory study. *Int. J. Soc. Psychiatry* **45,** 259–267 (1999).

48. Kaiser, W. *et al.* Profiles of subjective quality of life in schizophrenic in- and out-patient samples. *Psychiatry Res.* **66,** 153–166 (1997).

49. Eack, S. M. & Newhill, C. E. Psychiatric Symptoms and Quality of Life in Schizophrenia: A Meta-Analysis. *Schizophr. Bull.* **33,** 1225–1237 (2007).

50. Awad, A. G., Voruganti, L. N. P. & Heslegrave, R. J. A conceptual model of quality of life in schizophrenia: Description and preliminary clinical validation. *Qual. Life Res.* **6,** 0–0 (1997).

51. Sullivan, G., Wells, K. B. & Leake, B. Clinical factors associated with better quality of life in a seriously mentally ill population. *Psychiatr. Serv.* **43,** 794–798 (1992).

52. Urbanoski, K. A., Cairney, J., Adlaf, E. & Rush, B. Substance abuse and quality of life among severely mentally ill consumers: A longitudinal modelling analysis. *Soc. Psychiatry Psychiatr. Epidemiol.* **42,** 810–818 (2007).

53. Connell, J., Brazier, J., O'Cathain, A., Lloyd-Jones, M. & Paisley, S. Quality of life of people with mental health problems: a synthesis of qualitative research. *Health Qual. Life Outcomes* **10,** 1–16 (2012).

54. Atkinson, M., Zibin, S., Chuang, H. & others. Characterizing quality of life among patients with chronic mental illness: a critical examination of the self-report methodology. *Am. J. Psychiatry* **154,** 99–105 (1997).

55. Lehman, A. F. The effects of psychiatric symptoms on quality of life assessments among the chronic mentally ill. *Eval. Program Plann.* **6,** 143–151 (1983).

56. Oliver, J. P., Huxley, P. J., Priebe, S. & Kaiser, W. Measuring the quality of life of severely mentally ill people using the Lancashire Quality of Life Profile. *Soc. Psychiatry Psychiatr. Epidemiol.* **32,** 76–83 (1997).

57. Becker, T. *et al.* Links between social networks and quality of life: an epidemiologically representative study of psychotic patients in South London. *Soc. Psychiatry Psychiatr. Epidemiol.* **33,** 299–304 (1998).

58. Aubry, T., Flynn, R. J., Virley, B. & Neri, J. Social role valorization in communiy mental health housing: does it contribute to the community integration and life satisfaction of people iwth psychiatric disabilities? *J. Community Psychol.* **41,** 218–235 (2013).

59. UK700 Group. Predictors of quality of life in people with severe mental illness. Study methodology with baseline analysis in the UK700 trial. *Br. J. Psychiatry* **175,** 426–432 (1999).

60. Skinner, E. A. *et al.* Met and Unmet Needs for Assistance and Quality of Life for People with Severe and Persistent Mental Disorders. *Ment. Health Serv. Res.* **1,** 109–118 (1999).

61. Schneider, J., Wooff, D., Carpenter, J., Brandon, T. & McNiven, F. Community mental healthcare in England: associations between service organisation and quality of life. *Health Soc. Care Community* **10,** 423–434 (2002).

62. Taylor, R. E., Leese, M., Clarkson, P., Holloway, F. & Thornicroft, G. Quality of life outcomes for intensive versus standard community mental health services. PRiSM Psychosis Study. 9. *Br. J. Psychiatry* **173,** 416–422 (1998).

63. Barry, M. M. & Crosby, C. Quality of life as an evaluative measure in assessing the impact of community care on people with long-term psychiatric disorders. *Br. J. Psychiatry* **168,** 210–216 (1996).

64. Ruggeri, M. *et al.* Changes and predictors of change in objective and subjective quality of life: multiwave follow-up study in community psychiatric practice. *Br. J. Psychiatry J. Ment. Sci.* **187,** 121–130 (2005).

65. Huxley, P., Evans, S., Burns, T., Fahy, T. & Green, J. Quality of life outcome in a randomized controlled trial of case management. *Soc. Psychiatry Psychiatr. Epidemiol.* **36,** 249–255 (2001).

66. Graf, J. *et al.* Perceived Stigmatization of Mentally Ill People and Its Consequences for the Quality of Life in a Swiss Population: *J. Nerv. Ment. Dis.* **192,** 542–547 (2004).

67. Eklund, M. Work status, daily activities and quality of life among people with severe mental illness. *Qual. Life Res.* **18,** 163–170 (2009).

68. Rosenfield, S. Factors Contributing to the Subjective Quality of Life of the Chronic Mentally Ill. *J. Health Soc. Behav.* **33,** 299–315 (1992).

69. Boyd, A. S. & Bentley, K. J. The Relationship Between the Level of Personal Empowerment and Quality of Life Among Psychosocial Clubhouse Members and Consumer-Operated Drop-In Center Participants. *Soc. Work Ment. Health* **4,** 67–93 (2006).

70. Markowitz, F. E. Modeling Processes in Recovery from Mental Illness: Relationships between Symptoms, Life Satisfaction, and Self-Concept. *J. Health Soc. Behav.* **42,** 64 (2001).

71. Baker, F., Jodrey, D. & Intagliata, J. Social support and quality of life of community support clients. *Community Ment. Health J.* **28,** 397–411 (1992).

72. Srebnik, D., Livingston, J., Gordon, L. & King, D. Housing choice and community success for individuals with serious and persistent mental illness. *Community Ment. Health J.* **31,** 139–152 (1995).

73. Calsyn, R. J., Morse, G. A., Tempelhoff, B., Smith, R. & Allen, G. Homeless mentally ill clients and the quality of life. *Eval. Program Plann.* **18,** 219–225 (1995).

74. Calsyn, R. J. & Morse, G. Homeless men and women: Commonalities and a service gender gap. *Am. J. Community Psychol.* **18,** 597–608 (1990).

75. Stein, J. A. & Gelberg, L. Comparability and representativeness of clinical homeless, community homeless, and domiciled clinic samples: Physical and mental health, substance use, and health services utilization. *Health Psychol.* **16,** 155 (1997).

76. Van der Plas, A. G. M., Hoek, H. W., van Hoeken, D., Valencia, E. & van Hemert, A. M. Perceptions of quality of life and disability in homeless persons with schizophrenia and persons with schizophrenia living in non-institutional housing. *Int. J. Soc. Psychiatry* **58,** 629–634 (2012).

77. Wolf, J., Burnam, A., Koegel, P., Sullivan, G. & Morton, S. Changes in subjective quality of life among homeless adults who obtain housing: a prospective examination. *Soc. Psychiatry Psychiatr. Epidemiol.* **36,** 391–398 (2001).

78. Kertesz, S. G. *et al.* Homeless chronicity and health-related quality of life trajectories among adults with addictions. *Med. Care* **43,** 574–585 (2005).

79. Marshall, M., Lockwood, A. & Gath, D. Social services case-management for long-term mental disorders: a randomised controlled trial. *Lancet* **345,** 409–412 (1995).

80. Lehman, A. F., Dixon, L. B., Kernan, E., DeForge, B. R. & Postrado, L. T. A randomized trial of assertive community treatment for homeless persons with severe mental illness. *Arch. Gen. Psychiatry* **54,** 1038–1043 (1997).

81. Lam, J. A. & Rosenheck, R. A. Correlates of improvement in quality of life among homeless persons with serious mental illness. *Psychiatr. Serv. Wash. DC* **51,** 116–118 (2000).

82. Bebout, R. R., Drake, R. E., Xie, H., McHugo, G. J. & Harris, M. Housing status among formerly homeless dually diagnosed adults. *Psychiatr. Serv. Wash. DC* **48,** 936–941 (1997).

83. Camasso, M. J. Quality of Life Perception in Transitional Housing Demonstration Projects: An Examination of Psychosocial Impact. *Brief Treat. Crisis Interv.* **3,** 99 (2003).

84. Schutt, R. K., Goldfinger, S. M. & Penk, W. E. Satisfaction with residence and with life: When homeless mentally ill persons are housed. *Eval. Program Plann.* **20,** 185–194 (1997).

85. Toronto Shelter Support & Housing Administration,. *What Housing First means for people: results of streets to homes 2007 post-occupancy research*. (Toronto Shelter Support & Housing Administration, 2007).

86. Howard, G. S. *et al.* Internal Invalidity in Pretest-Posttest Self-Report Evaluations and a Re-evaluation of Retrospective Pretests. *Appl. Psychol. Meas.* **3,** 1–23 (1979).

87. Golembiewski, R. T., Billingsley, K. & Yeager, S. Measuring Change and Persistence in Human Affairs: Types of Change Generated by OD Designs. *J. Appl. Behav. Sci.* **12,** 133–157 (1976).

88. Schwartz, C. E., Andresen, E. M., Nosek, M. A., Krahn, G. L. & RRTC Expert Panel on Health Status Measurement. Response shift theory: important implications for measuring quality of life in people with disability. *Arch. Phys. Med. Rehabil.* **88,** 529–536 (2007).

89. Schwartz, C. E., Feinberg, R. G., Jilinskaia, E. & Applegate, J. C. An evaluation of a psychosocial intervention for survivors of childhood cancer: paradoxical effects of response shift over time. *Psychooncology.* **8,** 344–354 (1999).

90. Barry, M. M. & Zissi, A. Quality of life as an outcome measure in evaluating mental health services: a review of the empirical evidence. *Soc. Psychiatry Psychiatr. Epidemiol.* **32,** 38–47 (1997).

91. Chamberlain, K. & Zika, S. Stability and change in subjective well-being over short time periods. *Soc. Indic. Res.* **26,** 101–117 (1992).

92. Boucekine, M. *et al.* Exploring the response shift effect on the quality of life of patients with schizophrenia: an application of the random forest method. *Med. Decis. Mak. Int. J. Soc. Med. Decis. Mak.* **35,** 388–397 (2015).

93. Schwartz, C. E. & Sprangers, M. A. Methodological approaches for assessing response shift in longitudinal health-related quality-of-life research. *Soc. Sci. Med. 1982* **48,** 1531–1548 (1999).

94. Nagin, D. S. & Odgers, C. L. Group-Based Trajectory Modeling (Nearly) Two Decades Later. *J. Quant. Criminol.* **26,** 445–453 (2010).

95. Ahmed, S., Mayo, N., Scott, S., Kuspinar, A. & Schwartz, C. Using latent trajectory analysis of residuals to detect response shift in general health among patients with multiple sclerosis. [corrected]. *Qual. Life Res. Int. J. Qual. Life Asp. Treat. Care Rehabil.* **20,** 1555–1560 (2011).

96. Mayo, N. E., Scott, S. C., Bernstein, C. N. & Lix, L. M. How are you? Do people with inflammatory bowel disease experience response shift on this question? *Health Qual. Life Outcomes* **13,** 52 (2015).

97. Mayo, N. E., Scott, S. C. & Ahmed, S. Case management poststroke did not induce response shift: the value of residuals. *J. Clin. Epidemiol.* **62,** 1148–1156 (2009).

98. Ware, J., Kosinski, M. & Keller, S. D. A 12-Item Short-Form Health Survey: construction of scales and preliminary tests of reliability and validity. *Med. Care* **34,** 220–233 (1996).

99. Rabin, R. & Charro, F. de. EQ-SD: a measure of health status from the EuroQol Group. *Ann. Med.* **33,** 337–343 (2001).

100. Corrigan, P. W., Salzer, M., Ralph, R. O., Sangster, Y. & Keck, L. Examining the Factor Structure of the Recovery Assessment Scale. *Schizophr. Bull.* **30,** 1035–1041 (2004).

101. Reavis, J. A., Looman, J., Franco, K. A. & Rojas, B. Adverse Childhood Experiences and Adult Criminality: How Long Must We Live before We Possess Our Own Lives? *Perm. J.* **17,** 44–48 (2013).

102. Sam Tsemberis, G. M. Measuring Homelessness and Residential Stability: The Residential Time-Line Follow-Back Inventory. *Sam Tsemberis* **35,** (2007).

103. Barker, S., Barron, N., McFarland, B. H. & Bigelow, D. A. A community ability scale for chronically mentally ill consumers: Part I. Reliability and validity. *Community Ment. Health J.* **30,** 363–383 (1994).

104. Boothroyd, R. A. & Chen, H. J. The psychometric properties of the Colorado Symptom Index. *Adm. Policy Ment. Health* **35,** 370–378 (2008).

105. Latimer, E. A. *et al.* Generalisability of the individual placement and support model of supported employment: results of a Canadian randomised controlled trial. *Br. J. Psychiatry J. Ment. Sci.* **189,** 65–73 (2006).

106. Dennis, M. L., Chan, Y.-F. & Funk, R. R. Development and validation of the GAIN Short Screener (GSS) for internalizing, externalizing and substance use disorders and crime/violence problems among adolescents and adults. *Am. J. Addict. Am. Acad. Psychiatr. Alcohol. Addict.* **15 Suppl 1,** 80–91 (2006).

107. Nakagawa, S. & Schielzeth, H. A general and simple method for obtaining R2 from generalized linear mixed-effects models. *Methods Ecol. Evol.* **4,** 133–142 (2013).

108. StataCorp. *Stata Statistical Software : Release 13*. (StataCorp LP, 2013).

109. Altman, D. G. & Bland, J. M. Interaction revisited: the difference between two estimates. *BMJ* **326,** 219 (2003).

110. Schulz, K. F., Altman, D. G., Moher, D. & $author firstName. CONSORT 2010 Statement: updated guidelines for reporting parallel group randomised trials. *BMC Med.* **8,** 18 (2010).

111. McPhail, S., Beller, E. & Haines, T. Reference bias: presentation of extreme health states prior to EQ-VAS improves health-related quality of life scores. a randomised cross-over trial. *Health Qual. Life Outcomes* **8,** 146 (2010).

112. Schwarz, N., Strack, F., Kommer, D. & Wagner, D. Soccer, rooms, and the quality of your life: Mood effects on judgments of satisfaction with life in general and with specific domains. *Eur. J. Soc. Psychol.* **17,** 69–79 (1987).

113. Schwarz, N. & Clore, G. L. Mood, misattribution, and judgments of well-being: informative and directive functions of affective states. *J. Pers. Soc. Psychol.* **45,** 513 (1983).

Appendix A: QOLI- 20 instrument[29]

Quality of Life Inventory 20

ID | | | | | | Y 2 0 | | M | | D | |

Now I'll read a list of things about your life overall. I recognize that some of these things may be hard to talk about for some people so I appreciate your patience. For each item, I'd like you to tell me how you feel, on a scale of 1 to 7, where 1=terrible and 7=delighted.

1. How do you feel about your family in general?
   ○1 ○2 ○3 ○4 ○5 ○6 ○7 ○Don't know ○Declined
2. How do you feel about how often you have contact with your family?
   ○1 ○2 ○3 ○4 ○5 ○6 ○7 ○Don't know ○Declined
3. How do you feel about the way you and your family act toward each other?
   ○1 ○2 ○3 ○4 ○5 ○6 ○7 ○Don't know ○Declined
4. How do you feel about the way things are in general between you and your family?
   ○1 ○2 ○3 ○4 ○5 ○6 ○7 ○Don't know ○Declined
5. How do you feel about how comfortable and well off you are financially?
   ○1 ○2 ○3 ○4 ○5 ○6 ○7 ○Don't know ○Declined
6. How do you feel about the amount of money you have available to spend for fun?
   ○1 ○2 ○3 ○4 ○5 ○6 ○7 ○Don't know ○Declined
7. How do you feel about the way you spend your spare time?
   ○1 ○2 ○3 ○4 ○5 ○6 ○7 ○Don't know ○Declined
8. How do you feel about the amount of time you have to do the things you want to do?
   ○1 ○2 ○3 ○4 ○5 ○6 ○7 ○Don't know ○Declined
9. How do you feel about the chance you have to enjoy pleasant or beautiful things?
   ○1 ○2 ○3 ○4 ○5 ○6 ○7 ○Don't know ○Declined
10. How do you feel about the amount of fun you have?
    ○1 ○2 ○3 ○4 ○5 ○6 ○7 ○Don't know ○Declined
11. How do you feel about the amount of relaxation in your life?
    ○1 ○2 ○3 ○4 ○5 ○6 ○7 ○Don't know ○Declined
12. How do you feel about the living arrangements where you live?
    ○1 ○2 ○3 ○4 ○5 ○6 ○7 ○Don't know ○Declined
13. How do you feel about how safe you are in your neighbourhood?
    ○1 ○2 ○3 ○4 ○5 ○6 ○7 ○Don't know ○Declined
14. How do you feel about how safe you are where you live?
    ○1 ○2 ○3 ○4 ○5 ○6 ○7 ○Don't know ○Declined
15. How do you feel about the chance of finding someone to help in an emergency?
    ○1 ○2 ○3 ○4 ○5 ○6 ○7 ○Don't know ○Declined
16. How do you feel about your personal safety?
    ○1 ○2 ○3 ○4 ○5 ○6 ○7 ○Don't know ○Declined
17. How do you feel about the things you do with other people?
    ○1 ○2 ○3 ○4 ○5 ○6 ○7 ○Don't know ○Declined

Quality of Life Inventory 20

ID | | | | | | Y 2 0 | | M | | D | |

18. How do you feel about the amount of time you spend with other people?

○1 ○2 ○3 ○4 ○5 ○6 ○7 ○Don't know ○Declined

19. How do you feel about the people you see socially?

○1 ○2 ○3 ○4 ○5 ○6 ○7 ○Don't know ○Declined

20. How do you feel about your life overall (as a whole)?

○1 ○2 ○3 ○4 ○5 ○6 ○7 ○Don't know ○Declined

Now two questions about relationships.

21. Do you have a close confidante, that is, someone that you can share sensitive personal information with?

Health, social service or other providers do not count as close confidantes.

○Yes ○No ○Don't know ○Declined

21.a Do you see this person at least once a month? 'See' can include having contact on the phone or on-line as well as seeing the confidante in person.

○Yes ○No ○Don't know ○Declined

Appendix B. At Home/Chez Soi Final Report Analysis

Table B-1. QOLI-20 Total Score Analysis (At Home/Chez Soi Final Report Equivalent)[24]

Random Effects Model		Coefficient	*SE*	*p*-val	95% Conf. Int.
HF treatment (vs TAU group)		-0.15	0.94	.875	-1.98, 1.69
Time (vs baseline)					
	6 month	7.98 **	0.71	.000	6.59, 9.38
	12 month	12.26**	0.76	.000	10.77, 13.76
	18 month	12.61**	0.81	.000	11.03, 14.20
	24 month	14.48**	0.83	.000	12.85, 16.11
HF treatment x time					
	HF treatment x 6month	6.47**	0.93	.000	4.65, 8.29
	HF treatment x 12 month	4.42**	1.00	.000	2.46, 6.37
	HF treatment x 18 month	2.85**	1.06	.007	0.78, 4.91
	HF treatment x 24 month	2.91**	1.09	.007	0.78, 5.04
	Intercept	71.88	0.70		70.51, 73.26
	var(time)	1.81	0.38		1.20, 2.73
	var(intercept)	261.76	14.91		234.10, 292.69
	cov(time, intercept)	-4.26	1.79		-7.77, -0.76
	var(residual)	220.155	7.50		205.93, 235.37
Observations: 9,291	# of Individuals: 2,252				

[a] *p*< .1, **p*< .05, ** *p*< .01

Appendix C Methodology Diagram of Response Shift Identification by Mayo et al.[10]

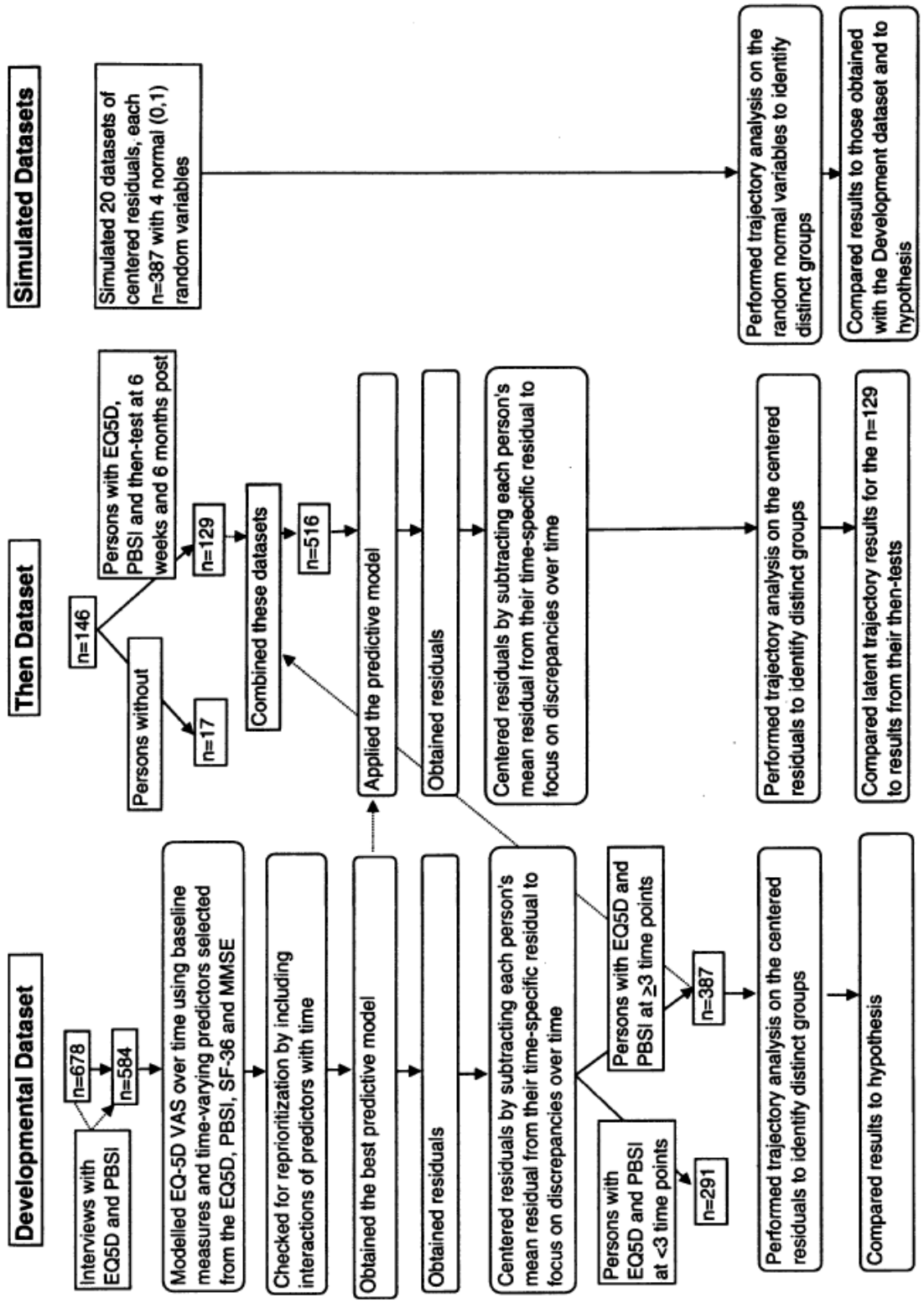


Fig. 1 Methodological approach

Appendix D

Table D-1. Response-Shift-Adjusted QOLI-20 Total Score Analysis

Random Effects Model	Coefficient	*SE*	*p*-val	95% Conf. Int.
Main effects				
HF treatment (vs TAU group)	3.06	2.15	.153	-1.14, 7.27
Time (vs baseline)				
6 month	9.41 **	1.63	.000	6.22, 12.60
12 month	13.71 **	1.64	.000	10.49, 16.92
18 month	15.16 **	1.73	.000	11.76, 18.55
24 month	18.16 **	1.84	.000	14.54, 21.77
sd of residuals (1 *sd* increase)	0.49	1.47	.737	-2.39, 3.38
2-way interactions (vs main effects)				
(at baseline)				
HF treatment x *sd* of residuals	-3.89 *	1.96	.047	-7.74, -0.05
(at *sd* of residuals = 0)				
HF treatment x 6month	6.43 **	2.10	.002	2.30, 10.55
HF treatment x 12 month	3.63 [a]	2.14	.089	-0.56, 7.82
HF treatment x 18 month	3.11	2.27	.169	-1.33, 7.55
HF treatment x 24 month	-1.03	2.40	.668	-5.74, 3.68
(at TAU group)				
sd of residuals x 6 month	-1.01	1.49	.500	-3.92, 1.91
sd of residuals x 12 month	-1.05	1.50	.481	-3.98, 1.88
sd of residuals x 18 month	-2.17	1.59	.171	-5.29, 0.94
sd of residuals x 24 month	-3.39 *	1.67	.043	-6.66, -0.11
3-way interactions (vs 2-way interactions)				
HF treatment x 6 month x *sd* of residuals	-0.08	1.94	.969	-3.87, 3.72
HF treatment x 12 month x *sd* of residuals	0.70	1.96	.722	-3.14, 4.54
HF treatment x 18 month x *sd* of residuals	-0.26	2.09	.902	-4.35, 3.83
HF treatment x 24 month x *sd* of residuals	4.04 [a]	2.20	.066	-0.27, 8.35
Effect of sd of residuals on HF treatment (simple effects within 3-way interaction)				
HF treatment x 6month	-3.97 [a]	2.04	.052	-7.97, 0.03
HF treatment x 12 month	-3.20	1.60	.109	-7.10, .711
HF treatment x 18 month	-4.15 *	2.02	0.043	-8.17, 0.13
HF treatment x 24 month	0.15	2.10	.946	-3.97, 4.26
Intercept	71.6	1.63		68.3, 74.7
var(time)	2.85	0.29		2.33, 3.47
var(intercept)	307.0	16.26		276.7, 340.6
cov(time, intercept)	-11.67	1.78		-15.16, -8.17
var(residual)	192.7	3.98		185.0, 200.6
Observations: 7,483, # of Individuals: 1,876 AIC: 23746.4				

[a] $p < .1$, *$p < .05$, ** $p < .01$

Appendix E

Table E-1. Response-Shift-Adjusted QOLI-20 Global Item Analysis

Partial-Proportional Odds Generalized Ordered Logistic Regression	Coefficient	*SE*	*p*-val	95% Conf. Int.
Main effects				
Baseline QOL rating (poor, fair, good)	2.66 **	1.05	.000	2.43, 2.92
HF treatment (vs TAU group)	1.58 **	1.16	.002	1.17, 2.11
sd of residuals (1 *sd* increase) †				
Better than *poor*	0.69**	1.12	.001	0.55, 0.85
Better than *fair*	1.39 **	1.12	.003	1.12, 1.73
2-way interactions (vs main effects)				
HF treatment x *sd* of residuals	0.83	1.15	.175	0.63, 1.09
Intercepts				
Odds contrast 1	0.77	1.16		0.58, 1.02
Odds contrast 2	0.04	1.16		0.03, 0.06
Observations: 6,470, # of Individuals: 1,798				

[a] *p*< .1, **p*< .05, ** *p*< .01. † Odds Ratios not proportional at both levels of ordinal logit equation
Estimates represent Odds Ratios of higher QOL Global Item category (poor: 1-3, fair: 4-5, good: 6-7)

Appendix F: Results of Sensitivity Analysis

Table F-1. Response-Shift-Adjusted QOLI-20 Total Score Analysis under *Exclusion Criterion A*

Random Effects Model	Coefficient	*SE*	*p*-val	95% Conf. Int.
Main effects				
HF treatment (vs TAU group)	0.89	2.20	.686	-3.42, 5.20
Time (vs baseline)				
6 month	9.74**	1.77	.000	6.28, 13.21
12 month	14.64**	1.74	.000	11.24, 18.05
18 month	16.25**	1.84	.000	12.64, 19.87
24 month	17.57**	1.97	.000	13.70, 21.44
sd of residuals (1 *sd* increase)	-0.50	1.55	.748	-3.53, 2.54
2-way interactions (vs main effects)				
(at baseline)				
HF treatment x *sd* of residuals	-2.40	2.04	.240	-6.40, 1.60
(at *sd* of residuals = 0)				
HF treatment x 6month	7.36**	2.28	.001	2.89, 11.84
HF treatment x 12 month	4.71*	2.27	.039	0.25, 9.16
HF treatment x 18 month	3.65	2.39	.127	-1.04, 8.34
HF treatment x 24 month	1.76	2.57	.493	-3.27, 6.79
(at TAU group)				
sd of residuals x 6 month	-0.81	1.68	.632	-4.10, 2.49
sd of residuals x 12 month	-0.85	1.62	.598	-4.02, 2.32
sd of residuals x 18 month	-2.61	1.73	.131	-6.00, 0.78
sd of residuals x 24 month	-2.46	1.84	.181	-6.06, 1.14
3-way interactions (vs 2-way interactions)				
HF treatment x 6 month x *sd* of residuals	-0.32	2.16	.884	-4.55, 3.91
HF treatment x 12 month x *sd* of residuals	0.02	2.11	.993	-4.12, 4.16
HF treatment x 18 month x *sd* of residuals	0.37	2.23	.869	-4.01, 4.75
HF treatment x 24 month x *sd* of residuals	2.64	2.39	.269	-2.04, 7.32
Effect of sd of residuals on HF treatment (simple effects within 3-way interaction)				
HF treatment x 6month	-2.40	2.04	.240	-6.40, 1.60
HF treatment x 12 month	-2.71	2.15	.206	-6.92, 1.49
HF treatment x 18 month	-2.38	2.02	.240	-6.35, 1.59
HF treatment x 24 month	-2.03	2.09	.332	-6.13, 2.07
Intercept	71.8	1.67		68.57, 75.11
var(time)	2.71	0.32		2.15, 3.40
var(intercept)	289.1	17.10		257.4, 324.6
cov(time, intercept)	-10.63	1.92		-14.40, -6.86
var(residual)	182.3	4.24		174.2, 190.8
Observations: 6,951 # of Individuals: 1,692				

[a] *p*< .1, **p*< .05, ** *p*< .01

Table F-2. Response-Shift-Adjusted QOLI-20 Total Score Analysis under *Exclusion Criterion B*

Random Effects Model	Coefficient	*SE*	*p*-val	95% Conf. Int.
Main effects				
HF treatment (vs TAU group)	2.18	2.86	.446	-3.43, 7.80
Time (vs baseline)				
6 month	14.50**	2.34	.000	9.92, 19.08
12 month	15.57**	2.35	.000	10.95, 20.18
18 month	18.79**	2.44	.000	14.00, 23.58
24 month	20.65**	2.61	.000	15.54, 25.76
sd of residuals (1 *sd* increase)	-0.54	2.05	.793	-4.56, 3.48
2-way interactions (vs main effects)				
(at baseline)				
HF treatment x *sd* of residuals	-4.62[a]	2.74	.092	-9.99, 0.75
(at *sd* of residuals = 0)				
HF treatment x 6month	5.94[a]	3.04	.051	-0.02, 11.90
HF treatment x 12 month	5.92[a]	3.05	.052	-0.06, 11.91
HF treatment x 18 month	4.12	3.18	.196	-2.12, 10.36
HF treatment x 24 month	2.19	3.40	.519	-4.47, 8.85
(at TAU group)				
sd of residuals x 6 month	-3.24	2.23	.146	-7.61, 1.12
sd of residuals x 12 month	-0.72	2.26	.751	-5.14, 3.71
sd of residuals x 18 month	-2.66	2.30	.248	-7.17, 1.85
sd of residuals x 24 month	-2.49	2.47	.315	-7.33, 2.36
3-way interactions (vs 2-way interactions)				
HF treatment x 6 month x *sd* of residuals	2.81	2.92	.336	-2.92, 8.54
HF treatment x 12 month x *sd* of residuals	1.82	2.95	.539	-3.97, 7.61
HF treatment x 18 month x *sd* of residuals	1.48	3.06	.628	-4.51, 7.48
HF treatment x 24 month x *sd* of residuals	3.55	3.28	.279	-2.88, 9.97
Effect of sd of residuals on HF treatment (simple effects within 3-way interaction)				
HF treatment x 6month	-4.62[a]	2.74	.092	-9.99, 0.75
HF treatment x 12 month	-1.81	2.52	.472	-6.74, 3.12
HF treatment x 18 month	-2.80	2.44	.251	-7.59, 1.98
HF treatment x 24 month	-3.14	2.46	.203	-7.96, 1.69
Intercept	69.30	2.18		65.02, 73.58
var(time)	2.88	0.45		2.11, 3.92
var(intercept)	261.36	21.79		221.96, 307.76
cov(time, intercept)	-10.52	2.66		-15.73, -5.30
var(residual)	158.22	5.29		148.18, 168.93
Observations: 6,951 # of Individuals: 1,692				

[a] $p < .1$, *$p < .05$, ** $p < .01$

Table F-3. Response-Shift-Adjusted QOLI-20 Global Item Analysis *under Exclusion Criterion A*

Partial-Proportional Odds Generalized Ordered Logistic Regression	Coefficient	*SE*	*p*-val	95% Conf. Int.
Main effects				
Baseline QOL rating (poor, fair, good)	2.75	1.05	.000	2.49, 3.04
HF treatment (vs TAU group)	1.61	1.17	.002	1.18, 2.19
sd of residuals (1 *sd* increase) †				
Better than *poor*	0.66	1.13	.001	0.52, 0.84
Better than *fair*	1.46	1.13	.002	1.15, 1.86
2-way interactions (vs main effects)				
HF treatment x *sd* of residuals	0.83	1.16	.210	0.63, 1.11
Intercepts				
Odds contrast 1	0.78	1.17	.129	0.57, 1.07
Odds contrast 2	0.04	1.18	.000	0.03, 0.06
Observations: 5,490, # of Individuals: 1,680				

[a] *p*< .1, **p*< .05, ** *p*< .01. † Odds Ratios not proportional at both levels of ordinal logit equation
Estimates represent Odds Ratios of higher QOL Global Item category (poor: 1-3, fair: 4-5, good: 6-7)

Table F-4. Response-Shift-Adjusted QOLI-20 Global Item Analysis under *Exclusion Criterion B*

Partial-Proportional Odds Generalized Ordered Logistic Regression	Coefficient	*SE*	*p*-val	95% Conf. Int.
Main effects				
Baseline QOL rating (poor, fair, good)	2.67	1.07	.000	2.33, 3.05
HF treatment (vs TAU group)	1.93	1.20	.000	1.35, 2.75
sd of residuals (1 *sd* increase) †				
Better than *poor*	0.78	1.15	.065	0.52, 0.84
Better than *fair*	1.49	1.16	.006	1.15, 1.86
2-way interactions (vs main effects)				
HF treatment x *sd* of residuals	0.69	1.18	.027	0.49, 0.96
Intercepts				
Odds contrast 1	0.73	1.21	.093	0.50, 1.05
Odds contrast 2	0.04	1.22	.000	0.03, 0.06
Observations: 3,204 # of Individuals: 1,158				

[a] *p*< .1, **p*< .05, ** *p*< .01. † Odds Ratios not proportional at both levels of ordinal logit equation
Estimates represent Odds Ratios of higher QOL Global Item category (poor: 1-3, fair: 4-5, good: 6-7)