SCUOLA DI DOTTORATO
UNIVERSITÀ DEGLI STUDI DI MILANO-BICOCCA

Department of

MEDICINE AND SURGERY

PhD program PUBLIC HEALTH Cycle XXXV

Curriculum in BIOSTATISTICS AND EPIDEMIOLOGY

The efficacy of a motivational interview intervention in patients with heart failure and their caregivers: a dyadic analysis

Surname OCCHINO Name GIUSEPPE

Registration number 771380

Tutor: REBORA PAOLA

Co-tutor: AUSILI DAVIDE LUIGI LINO

Coordinator: GRASSI GUIDO

ACADEMIC YEAR 2021/2022

Contents

List of abbreviations

APIM: Actor-Partner Interdependence Model

BLUP: Best Linear Unbiased Predictor

CC: Caregiver Contribution

CCI: Charlson Comorbidity Index

CI: Confidence Interval

CPS: Caregiver Preparedness Scale

CS: Compound Symmetric

HADS: Hospital Anxiety and Depression Scale

HAS: Hospital Anxiety Scale

HDS: Hospital Depression Scale

HF: Heart Failure

HFSPS: Heart Failure Somatic Perception Scale

HR: Hazard Ratio

ICC: Intra-Class Correlation

KCCQ: Kansas City Cardiomyopathy Questionnaire

LMM: Linear Mixed Model

MCS: Mental Component Summary

MI: Motivational Interview

MLM: Multilevel Modelling

MoCA: Montreal Cognitive Assessment

MS: Mutuality Scale

NYHA: New York Heart Association

OR: Odds Ratio

PCS: Physical Component Summary

PSQI: Pittsburg Sleep Quality Index

QOL: Quality of Life

RCT: Randomized Clinical Trial

REML: Restricted Maximum Likelihood

SCHFI: Self-Care Heart Failure Index

SD: Standard Deviation

SF: Short Form

Aim of this project

In this thesis I will focus on the use of multilevel models on dyadic data. This topic was driven by a randomized clinical trial (RCT) on the efficacy of a motivational interview (MI) in improving self-care in patients with heart failure (HF). Longitudinal mixed models were used to evaluate the effect of MI on a battery of primary and secondary outcomes in patients and caregivers, including self-care, caregiver contribution to self-care, HF symptoms, generic and specific quality of life, anxiety, depression, sleep quality, caregiver preparedness, use of emergency services and mortality.

Secondary outcomes included also the mutuality between patient and caregiver, which evaluates the positive quality of the care relationship by separate questionnaires to the patients and to the caregivers. Using the dyad as the unit of analysis, rather than the individuals, multilevel modelling is an appropriate and effective methodology for considering the dyadic context. In this framework it is possible to investigate the impact of health and behaviour changes on both the patient and caregiver's perceived relationship quality. The problem of interdependence in dyadic data is further complicated when measuring dyadic outcomes repeatedly over time. Then, it is necessary not only to account for the non-independence of the members within a dyad, but also for the dependence of the longitudinal measures within one member. Therefore, this thesis studies the application of multilevel models to longitudinal dyadic data in healthcare research, evaluating advantages and disadvantages of this approach and commenting on the implications of these models.

1. Introduction to clinical context

1.1 Self-care in heart failure patients

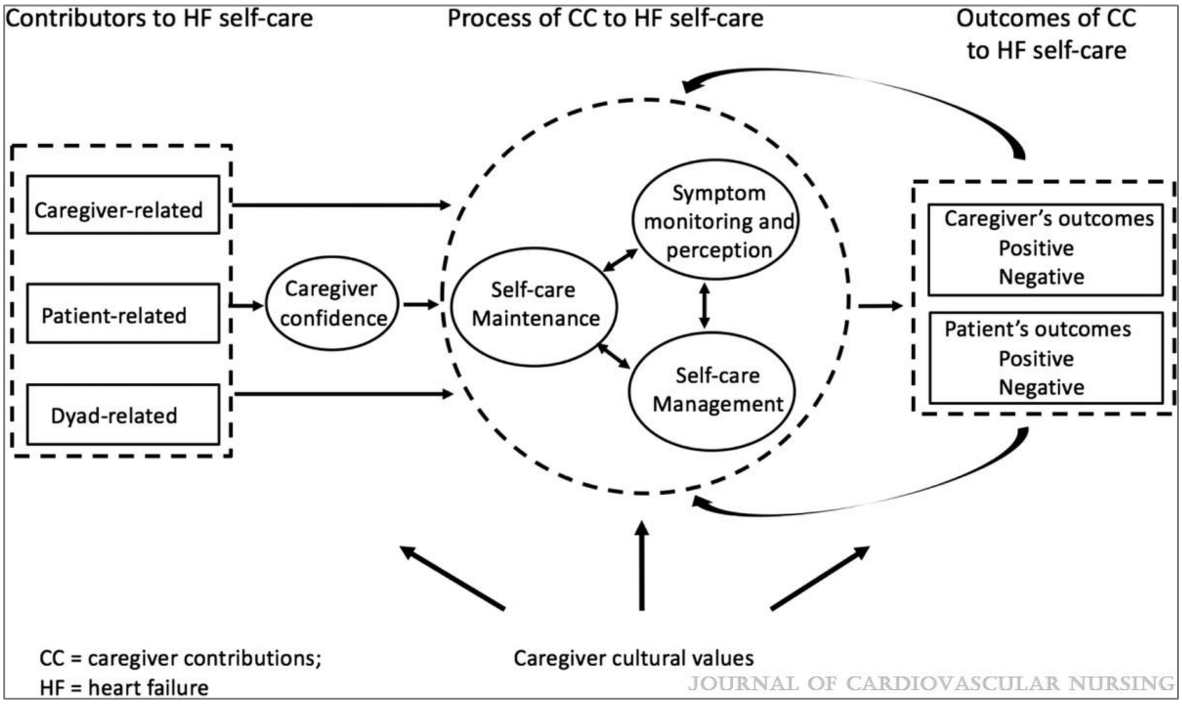
Over 15 million people in Europe and 5.7 million in the USA are chronically affected with heart failure (HF) [1]. Because of population aging, HF is becoming more common worldwide as the ultimate stage of a wide range of heart disorders. According to projections, HF will afflict 46% of the population who are aged at least 18 years old by 2030 [2].

HF has a significant effect on patients' and their families' quality of life (QOL) with very high symptom loads and depression levels [3]. More unmet needs are experienced by family caregivers of HF patients than by other caregiver populations [4].

Outcomes for HF patients, such as QOL and hospitalizations, have been shown to improve with self-care [5], defined as "a naturalistic decision-making process that promotes behaviours that maintain physiologic stability (maintenance), assist the perception of symptoms (symptom perception), and direct the management of those symptoms (management)" [6]. The ability to maintain self-care, perceive symptoms, and manage them are all influenced by task-specific self-efficacy in the self-care process.

The theory of caregiver contribution to heart failure self-care was illustrated by Riegel et al. (2016) [6]. Contributors to HF self-care are included in the first block of Figure 1, which might be related to the caregiver, patient, or dyad [6]. Caregiver confidence mediates, completely or in part, between the majority of the contributors and the process by which caregivers contribute to HF self-care, but it might not be able to influence all contributors. Caregivers contribute to the maintenance, symptom monitoring and perception, and management of HF self-care. These processes interact with each another. The outcomes are related to the caregiver and the patient and might be either positive or negative (Figure 1).

Figure 1. Illustration of the theory of caregiver contribution to heart failure self-care.



Note. Figure published in *Riegel B, Dickson VV, Faulkner KM. The Situation-Specific Theory of Heart Failure Self-Care: Revised and Updated. J Cardiovasc Nurs. 2016;31(3):226-235.*

Trials aimed at improving HF self-care have produced contradictory findings on the efficacy of educational interventions [7,8]. A recent systematic review and meta-analysis that included 20 randomized controlled trials (RCT) with a total of 5624 HF patients revealed that interventions to improve self-care were effective in reducing mortality and hospitalization. However, the authors reached the conclusion that it was unclear how self-care could be improved. In fact, this systematic review indicated that interventions with standardized training for interventionists, peer interaction, record-keeping, or goal-setting abilities were even less successful than those without these features [9]. Another recent systematic review of 33 trials examined at the main mechanisms through which HF self-care programs are effective [10].

Findings indicate that interventions may be more successful if these programs:

1. mobilize significant support from providers (e.g., developing a good supportive relationship with patients);

2. promote self-efficacy (e.g., promoting personal confidence and hope);
3. increase insight and understanding of HF and self-care (e.g., linking symptoms to HF self-care tasks);
4. increase the involvement of other people (e.g., including a family caregiver);
5. are tailored to the needs of each individual (e.g., basing the intervention on patient need and preferences).

Motivational interviewing places a strong emphasis on each of these mechanical components.

1.2 Motivational interviewing

Using a collaborative and evocative approach that respects the patient's autonomy to elicit his or her own motivation to change behaviours in the interest of health, motivational interviewing (MI), a counselling technique, is described as a "person-cantered method of guiding to elicit and reinforce personal motivation for change". In fact, MI is opposed with delivering unsolicited advice. MI examines and removes behavioural ambivalence in people (such as someone who believes exercise is necessary but does not exercise), which increases intrinsic motivation to improve. Develop discrepancy, show empathy, avoid argument and direct confrontation, roll with resistance, support self-efficacy and optimism are the guiding principles of MI [11].

Although the MI method was developed in psychology, healthcare professionals have started using it. According to several randomized controlled trials, MI enhances medication compliance, dietary adherence and weight reduction in diabetic patients, as well as avoiding smoking, reducing depression, and improving quality of life in people with cardiovascular illnesses [12–14].

Masterson Creber et al. have shown in a sample of 67 HF patients that those randomly assigned to a personalized MI intervention had better self-care maintenance [15] and a decrease in hospital readmissions at the three-month follow-up [16] than those in the control group. In an RCT involving 30 HF patients, Paradis et al. found that MI significantly increased self-care confidence, while Brodie et al. found that MI improved both general and disease-specific QOL in HF patients [17,18]. The use of reflective listening, empathy, acknowledgment of cultural beliefs, an effort to overcome barriers and constraints, the facilitation of an action plan, the development of skills, and the activation of support resources were all factors that made MI effective in this population, according to a study by Riegel et al. on 41 HF patients [19].

1.3 The role of caregiver and the dyadic context

A number of obstacles prevent patients from performing effective self-care, including older age, low self-care self-efficacy, cognitive impairment, comorbidities, and depression [20]. Because of this, the presence of an informal caregiver may be beneficial in removing the above-mentioned obstacles and encouraging patients to take care of themselves. The "supply of time, effort, and assistance on behalf of another person who needs to execute HF self-care" has been defined as a caregiver contribution (CC) to HF self-care [21]. Thus CC also includes CC to self-care management, CC to symptom monitoring and perception, and CC to self-care maintenance [22].

Supporting the patient in maintaining HF stability is what is meant by a caregiver's contribution to self-care maintenance, such as through encouraging exercise or providing medications; CC to self-care management relates to responding to HF symptoms when they occur, for example, by contacting the healthcare provider in case the patient experiences swelling. CC to symptom monitoring and perception includes supporting patients in monitoring and perceiving symptoms, for example, by observing for HF symptoms such as swelling or weight gain. The situation-specific theory of CC to HF self-care states that caregivers may experience both positive and negative effects from CC to self-care. Indeed, providing care for another person fosters personal development and boosts the reward, happiness, and self-worth of the carer [23].

In the studies found in literature, few caregivers received assistance on how better contribute to the self-care of HF patients. Even though caregivers are recognised to play a significant part in HF patients' efforts to improve their own care, a psychometrically reliable measure of this contribution was only created in 2013 [21]. Therefore, whether increasing this contribution through an intervention will have an impact on patients' self-care is still unclear.

Caregiver burden can, however, also be an issue when it comes to self-care. A patient's response to a caregiver's contribution to self-care may be both favourable and unfavourable. In fact, higher CC to self-care in HF may improve patient compliance with treatment, exercise, diet, and flu vaccination, and reduce clinical event risks such mortality and rehospitalisation [24,25]. However, patients could become distressed if they are continually reminded to act in a particular way. As a result, caregivers need to be prepared in order to have adequate levels of CC for self-care. Studies have shown that the HF population's CC to self-care is below the required standards [26].

1.4 **Importance of mutuality**

Many close relationships throughout life involve providing care for a loved one, however there are occasions when this care scenario might take over and dominate other interactions [27]. An illness experience can affect family relations and change how the patient and caregiver address their well-being and the care situation, with the consequent relationship reciprocity considerably reduced [28].

Caregivers are at risk for poor physical and mental health, as well as, in the case of spouses who feel severe strain, at greater risk for mortality, according to a large amount of research on family care [29]. However, caregivers who report having a positive relationship with the patient have been found to experience significantly less role captivity (a sense of being enveloped in the role of caregiver) and overload, as well as less stress from providing direct care, tension, and feelings of manipulation, and role conflict, compared to caregivers who report a negative relationship [30]. High mutuality, defined as the positive quality of the care relationship, is what enables the caregiver to continue providing care even in challenging situations. Indeed, high mutuality has actually been linked to less instances of engaging in potentially harmful activities and a lower chance of being admitted to a nursing home [31].

High levels of mutuality with the patient are known to be a protective factor for the caregiver, but research findings also suggest that the mutuality deteriorates over the course of care and is negatively impacted by the presence of one member's presence of cognitive impairment.

As described in the literature, it is not a novel idea that mutuality could have a positive impact on self-care in the HF population. Indeed, authors have observed that mutuality has a direct influence on homecare workers' contribution to self-care [32] and a positive impact on HF patients' self-care maintenance and on patients and caregivers' self-care confidence [33]. Similarly, it was shown that HF patients and caregivers with higher mutuality were more confident in the patient's self-care [34].

Knowledge of whether a motivational interview can improve mutuality in HF patient–caregiver dyads is crucial because mutuality has a positive impact on outcomes for both HF patients and caregivers, such as HF patients' lower mortality, anxiety and depression, and better health status, and caregivers' lower burden and depressive symptomatology [35–37]. Because it has increasingly often been highlighted how important it is to include both patient and caregiver in the care process, [38] studying the effect of MI on the patient–caregiver mutuality can be fundamental. Indeed, both patient and caregiver can benefit from this effect during the treatment process and, consequently, experience better outcomes, especially after the patient's discharge.

2. Introduction to dyadic context

2.1 Dyads framework

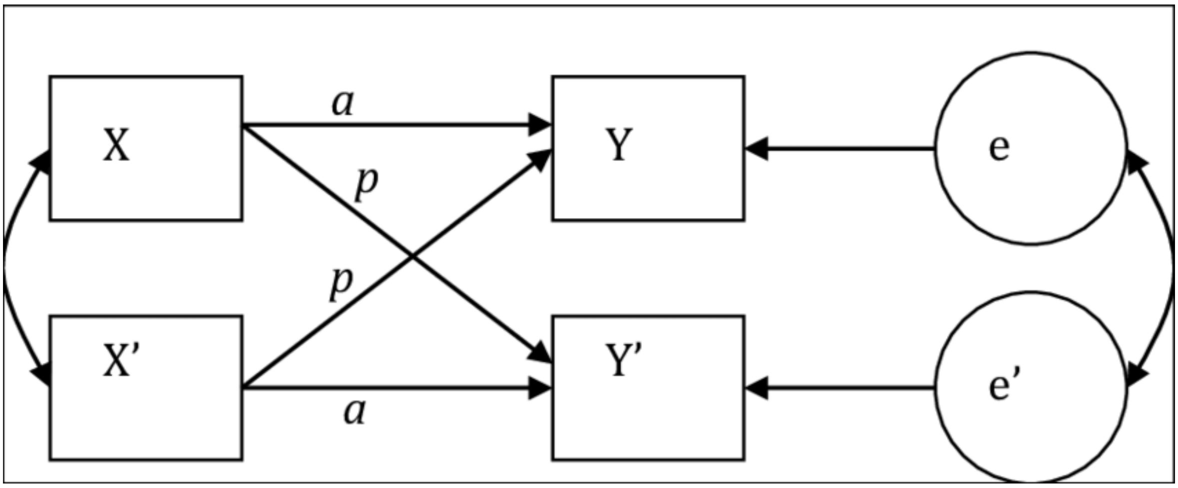
The wellbeing of one member in a care dyad is strongly associated with the wellbeing of the other member [39,40]. The interactive and dynamic processes of the care situation are being emphasized more frequently in recent family care research. In order to better focus on the context of the care dyad, many research questions are multilevel in nature, requiring more than one unit of analysis. However, the majority of family care research have been carried out using techniques that ignore the care dyad's hierarchical structure and, as a result, the variation in both within- and between-dyad processes. Using the dyad as the unit of analysis, it is important to modify this framework to investigate the impact of health changes on both the patient and caregiver's perceived relationship quality.

The outcome of each individual in a relationship can therefore be affected by both his or her own cognitions, emotions, and behaviours as well as those of the other person. In various study fields, the Actor-Partner Interdependence Model (APIM) [41] offers an appealing method to model such dyadic data. When each individual in the study is connected to only one other person and both member in the dyad have measurements for the same variables, this method is frequently utilized. When a person's score on an outcome variable is affected by that person's score on a predictor variable, it is said to have an actor effect; while, when it is affected by his or her partner's score, it is said to have a partner effect [42]. By means of partner effects, interdependence or reciprocity (e.g., the extent to which a person's response is dependent on some characteristics of his or her partner) is measured. Evidence for bidirectionality can only be found when both members of a dyad have significant partner effects.

Figure 2 shows the APIM for cross-sectional dyadic data. X and X' denote scores of person A and B on predictor variables, respectively; while Y and Y' denote scores of person A and

B on the outcome variable, respectively. The single-headed arrows represent predictive relationships, while the double-headed arrows indicate dependencies or correlations. Intrapersonal actor effects are symbolized by *a* and interpersonal partner effects by *p*. The double–headed arrow between *e* and *e'* represents the residual non-independence in the outcome scores that can't be explained by the Actor-Partner Interdependence Model, thus by the predictors included in the model. It's sometimes referred to as the intra-class correlation (ICC) and indicates a correlation between the unexplained variances of the dependent variables, even after the covariance due to partner effects has been removed [42].

Figure 2. The actor–partner interdependence model (APIM)



Note. Figure published in Cook WL, Kenny DA. The Actor–Partner Interdependence Model: A model of bidirectional effects in developmental studies. Int J Behav Dev. 2005;29(2):101-109.

A process known as multilevel modelling (MLM) allows for the simultaneous examination of data gathered at many units of analysis (e.g., children in schools, individuals in families, repeated assessments in individuals). Using multilevel modelling, the researcher can explore how different settings and systems, such as a community, a school, a family, or a dyad, affect individual change. The analysis of the APIM using multilevel modelling procedures requires thinking about the organisation of data and the estimation of effects. MLM estimates all the parameters of the model within a single equation and so implies a very different data structure compared to when there are two equations, one for each member of the dyad [42].

There are several ways to model dyadic processes within MLM. We will illustrate what is called the ‘‘two-intercept’’ approach that was introduced by Raudenbush, Brennan, and Barnett (1995) [43]. Many cross-sectional studies on married couples try to use covariates measured on people and families to explain outcomes evaluated at the person level. A suitable model should take into account the dependence that develops as a result of individuals nesting within couples. A traditional multivariate regression model with estimates by ordinary least squares can be utilized if the same covariates are used to predict the female and male outcomes. But if data are missing for some couples' members or if different variables are used to predict the female outcome than the male outcome, the model is not appropriate.

This approach makes sure that when estimating covariation between outcomes for the same couple, these estimates will be corrected for measurement error [43]. Similar corrections will be made to the estimated proportions of variance explained. The Level 1 model, often known as the within-dyads model, in this situation simply assumes that the outcome score equals the sum of the true score and the measurement error:

$$Y_{tp} = (male)\left[\beta_{mp}\right] + (female)\left[\beta_{fp}\right] + r_{tp}$$

where *male* is an indicator variable that assumes a value of 1 for men and 0 for women; vice versa *female* is equal to 1 for women and 0 for men. Therefore, the prediction equation for men is $Y_{tp} = \beta_{mp} + r_{tp}$ and for women is $Y_{tp} = \beta_{fp} + r_{tp}$.

The pair of true scores (β_{mp} and β_{fp}) for each couple become latent multivariate outcome variables at Level 2 (between-dyads). The population average of the βs reflect the male and female population means, the variation of the βs about their averages denotes individual variation in the outcome, and the correlation between the βs denotes the strength of linear association between one’s own outcome and that of one’s partner. The quantities were estimated and thus explanatory models were formulated to better understand the correlates of the outcome in both men and women [44]. The analysis allowed the computation of proportions of true-score variation explained for men and women. Randomly missing data do not prevent a valid statistical inference (for instance, some couples may only have the data from one partner).

2.2 Longitudinal dyad modelling

When dyads are studied using a repeated measures design, the issue of interdependence in dyadic data is made much more challenging. When measuring dyadic outcomes repeatedly over time, it is necessary to take into account both the dependence of the longitudinal measures within one member in addition to the non-independence of the members of the dyad.

The literature describes a number of potential methods to model longitudinal data on dyads for outcomes measured at the interval level. Because of the two types of dependency, within-subject and within-dyad, modelling longitudinal dyadic data is difficult, and methodologists have not yet reached a consensus on which data-analytic approaches are most suitable for handling this type of data [45].

Data collected from dyads have a multilevel structure and the units are organized in a hierarchy, with one set of units nesting inside another. There are three levels of analysis in this situation, where dyadic outcomes are measured repeatedly over time: the dyad, the individuals within the dyad, and the observations within the individuals. The dyad serves as the independent sample unit. In other words, dyads are sampled independently, but, the individuals and observations within each dyad are dependent.

So far we saw that, when responses from dyad members are thought of as Level 1 units nested inside the Level 2 unit, the dyad, multilevel modelling for dyads extends multiple regression to the case. The model can be modified to represent the cross-sectional model for matched couples [44] described in the previous paragraph, in which the pair serves as the unit of analysis rather than an individual. Combining elements of the cross-sectional model for matched pairs and the longitudinal model for individual change we get the longitudinal matched-pairs model [43], which compares the patterns of change in trajectories for both members and is fitted to the repeated assessments of the outcome for both members. This multivariate model estimates each dyad member's latent trajectory, which can vary between members in terms of pattern and magnitude.

Now, we highlight some further benefits of multilevel modelling when estimating longitudinal trajectory models for dyads [46]. The main benefit of within-dyad design is that each member can have a unique trajectory that can differ in both pattern (i.e., change can be linear for patients and nonlinear for caregivers) and magnitude (i.e., the rate of change can be steeply negative for patients and flat for caregivers). A generalized multivariate hypothesis test compares the average trajectories for each member for significant differences at the intercept (the predicted score at a particular measurement occasion), the slope (rate of change), or both. This is not possible when trajectories are estimated in separate models. A second benefit of this model is the control for the autocorrelation among the repeated measures and the adjustment of the error variance for the interdependence of partner outcomes within the same dyad. Standard errors and their associated hypothesis tests are improved by this adjustment. The third benefit is the ability to allow for unbalanced designs. For example, different dyads may have different numbers and spacings between measurement periods. A fourth benefit is the possibility to have missing responses under the assumption that they are missing at random. So it is possible for just one member of the dyad to provide information or for the pattern of missing responses to be different for each partner.

3. The MOTIVATE-HF randomized controlled trial

The motivating context of this project is a three-arm randomized controlled trial named MOTIVATional intErviewing to improve self-care in Heart Failure patients (MOTIVATE-HF) [47].

The aims of the MOTIVATE randomized controlled trial are:

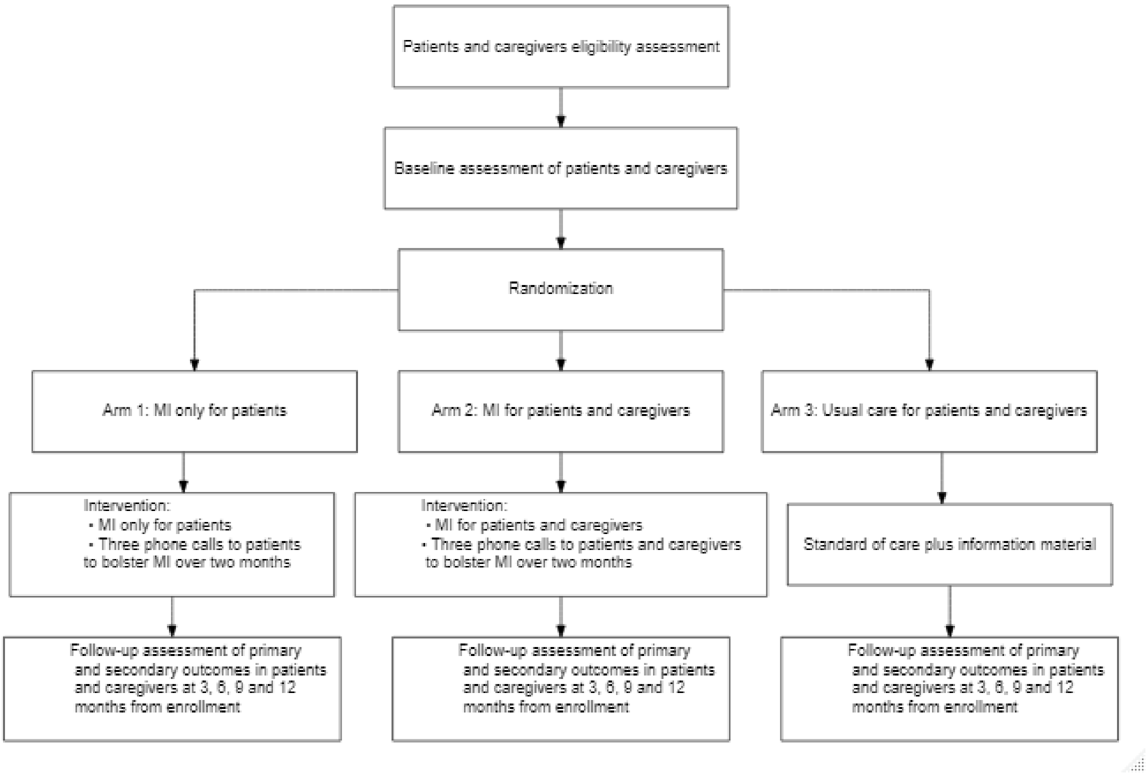
1. to evaluate the effect of motivational interviewing in heart failure patients and caregivers in improving patient self-care maintenance (primary outcome);
2. to evaluate the effect of motivational interviewing in caregivers in improving patient self-care in addition to motivational interviewing performed only on patients;
3. to evaluate the effect of motivational interviewing on the following secondary outcomes:
    a. in HF patients: self-care management, self-care confidence, anxiety and depression, generic and disease-specific quality of life, sleep quality, hospitalizations, use of emergency services, and mortality, HF somatic symptom perception, and mutuality with the caregiver;
    b. in caregivers: caregiver contribution to self-care, preparedness, anxiety and depression, generic quality of life, sleep quality, and mutuality with the patient.

3.1 Study design and participants

Patients in this trial were randomly assigned to one of the following three arms (Figure 3):

1. Motivational interviewing intervention for patients only;
2. Motivational interviewing intervention for patients and caregivers;
3. Standard of care for patients and caregivers.

Figure 3. Flow-chart of the MOTIVATE-HF randomized controlled trial.



Note. Figure published in *Vellone E, Paturzo M, D'Agostino F, et al. MOTIVATional intErviewing to improve self-care in Heart Failure patients (MOTIVATE-HF): Study protocol of a three-arm multicenter randomized controlled trial. Contemp Clin Trials. 2017;55:34-38.*

Patients with HF and their caregivers were recruited at three centres in the Lazio region of Italy: one hospital, one outpatient, and one community setting. The following inclusion and exclusion criteria were used to determine if HF patients and their caregivers were eligible for the study. Patients had to meet the following requirements in order to be included in the study:

1. a confirmed diagnosis of heart failure (HF) in accordance with international guidelines [1];
2. New York Heart Association (NYHA) functional class II–IV;
3. insufficient self-care as measured by a score of 0, 1, or 2 in at least two items of the self-care maintenance or self-care management scales of the Self-Care Heart Failure Index (SCHFI);

4. willingness to participate in the study and to sign the informed consent form.

Patients had to meet the following criteria in order to be excluded from the study:

1. severe cognitive impairment, as measured by a score of 0–4 on the Six-item Screener [48];
2. an episode of acute coronary syndrome within the previous three months;
3. residence in a residential setting (such as a nursing home);
4. caregivers who refused to take part in the study.

Caregivers had to meet the following requirement in order to be included in the study:

1. providing the majority of care for the HF patient, whether they were family members or not, as indicated by the patient as the primary informal caregiver.

This was the only caregiver exclusion criteria:

1. the patient was unwilling to take part in the research.

Both were excluded to enrol if either the patient or the caregiver refused to take part in the trial. However, if one participant left the study after enrolment, the other one continued. The trial has been registered at ClinicalTrials.gov and complies with the Declaration of Helsinki. It has also received approval from the Institutional Review Board of the University of Rome "Tor Vergata".

3.2 Intervention and control group

Eighteen registered nurses, six in each centre, who have completed a 40-hour training about motivational interviewing administered the intervention. Both patients and caregivers received this intervention in arms 1 and 2, respectively. The interventionist covered one or two aspects of self-care that the participants desired to address during the first session, which lasted about 60 minutes. According to the MI's guiding principles, the interventionist:

1. developed discrepancy (e.g., by assisting the patient or caregiver in realizing that current behaviours would make it difficult to achieve health goals);
2. expressed empathy (e.g., through active listening and an attitude of acceptance);
3. avoided arguing and direct confrontation (e.g., by respecting the patient's or caregiver's choices or preferences);
4. rolled with resistance (e.g., by involving the patient or caregiver in problem-solving);
5. supported self-efficacy and optimism (e.g., by verbal persuasion and encouraging a focus on past successes).

The same interventionist called the participant again after this initial intervention to reinforce it and offered further support as necessary. Following the initial intervention, these phone calls were made three times at intervals of two weeks (for a total of two months). In addition to receiving the intervention, patients and caregivers received also educational materials on HF management that adhered to global standards. All interventions were audio recorded in order to evaluate the effectiveness of MI for treatment fidelity purposes. In the control group (arm 3), patients and caregivers received standard care, which often includes giving patients and their family members oral education on the disease and its treatment as well as a medical check-up every 6–12 months depending on the patient's health.

3.3 Baseline and follow-up assessment

At baseline, sociodemographic characteristics on patients and caregivers were evaluated. Patients' clinical features (such as their NYHA functional class, the SCHFI score and the Six-item Screener) were also evaluated. Patients and caregivers were assessed using a battery of psychometrically sound measures (Table 1) as part of a follow-up assessment that was carried out at 3, 6, 9 and 12 months after enrolment in order to assess the primary and secondary outcomes. Trained nurse research assistants who were blinded to group and distinct from those who perform MI conducted the baseline and follow-up assessments.

Table 1. Battery of primary and secondary outcomes

Variable	Instrument	Administered to	Times of data collection (months)				
			Baseline	3	6	9	12
Self-care maintenance, management and confidence	SCHFI	P	X	X	X	X	X
Caregiver contribution to HF self-care	CC-SCHFI	C	X	X	X	X	X
HF symptoms	HFSPS	P	X	X	X	X	X
Generic QOL	SF-12	P-C	X	X	X	X	X
Specific QOL	KCCQ	P	X	X	X	X	X
Anxiety/depression	HADS	P-C	X	X	X	X	X
Sleep quality	PSQI	P-C	X	X	X	X	X
Mutuality	MS	P-C	X	X	X	X	X
Caregiver Preparedness	CPS	C	X	X	X	X	X
Hospitalizations	Questionnaire	C		X	X	X	X
Use of emergency services	Questionnaire	C		X	X	X	X
Mortality	Questionnaire	C		X	X	X	X

Note. P: patient; C: caregiver; SCHFI: Self-Care of Heart Failure Index; CC-SCHFI: Caregiver contribution to self-care of HF index; HFSPS: Heart Failure Somatic Perception Scale; QOL: quality of life; SF-12: Short Form 12; KCCQ: Kansas City Cardiomyopathy Questionnaire; HADS: Hospital Anxiety and Depression Scale; PSQI: Pittsburgh Sleep Quality Index; MS: Mutuality scale; CPS: Caregiver Preparedness Scale.

3.4 Outcome measures

The Self-Care Maintenance Scale of the Self-Care of HF Index version 6.2 (SCHFI) was used to assess HF self-care maintenance in patients as the primary outcome of this RCT [49]. The SCHFI is a tool used globally to measure the self-care aspects of management and maintenance. The Self-Care Maintenance Scale, which was evaluated for validity and reliability, records HF symptom monitoring (e.g., daily weighing) and treatment adherence (e.g., taking meds as directed). Higher scores indicate greater self-care maintenance on the self-care maintenance scale, which produces a score between 0 and 100. Three months after enrolment, the primary outcome of patients maintaining their self-care was assessed as well as at 6, 9 and 12 months following enrolment.

A variety of secondary outcomes were assessed using a battery of instruments, all of which have been shown to be valid and reliable (Table 1). In patients, the following specifically were used:

1. the Self-Care Management and Self-Care Confidence scales of the SCHFI measured reactions to signs and symptoms of HF worsening and the confidence in managing all self-care processes, respectively [49]. As the primary outcome, higher scores indicate better self-care and each scale generates a standardized score ranging from 0 to 100. Only if the patient had previously reported experiencing HF symptoms, such as dyspnea, the self-care management scale was completed.
2. the HF somatic perception scale (HFSPS) measures the burden of symptoms [50]. This instrument consists of an 18-item questionnaire that asks patients how concerned they were by their HF symptoms the previous week. All HFSPS items are categorized into four subscales: edema, dyspnea, early and subtle symptoms, and chest discomfort. Scores for each response vary from "I did not have the symptom" (scoring 0) to "Extremely bothersome symptoms" (score 5). The scale has standardized 0–100 scores for each subscale, with higher scores indicating more symptom burden.
3. the SF-12 measures general physical and mental QOL [51]. It is a multipurpose, short-form general health status assessment tool that measures physical functioning, role limitation (due to physical or emotional health issues), bodily pain, general health, vitality (energy and fatigue), social functioning, and mental health (psychological distress and psychological well-being). A mental component score (MCS-12) and a physical component score (PCS-12) are both presented as two summary scores. All

scores are standardized in the range of 0-100 where higher ratings indicate greater mental and physical QOL.

4. the Kansas City Cardiomyopathy Questionnaire (KCCQ) measures HF specific QOL [52]. It is a 23-item, self-administered questionnaire to evaluate the patient's assessment of their health status. Questionnaire asked about HF symptoms, how they affect physical and social function, and how HF has affected their quality of life during the past two weeks. Six different areas (symptoms, physical function, QOL, social limitation, self-efficacy, and symptom stability) are quantified by the KCCQ, and two summary scores (clinical summary score and overall summary score) are also provided.
5. the Hospital Anxiety and Depression Scale (HADS) measures anxiety and depression [53]. It is a 14-item scale including 7 items related to anxiety and 7 items related to depression that provide 2 scores. Each item is given a value between 0 and 3, so the anxiety (HAS) and depression (HDS) summary scores range from 0 to 21. An anxiety or depression score of 8 to 10 suggests a moderate degree of symptoms, and a score of 11 or more indicates a considerable level of symptoms.
6. the Pittsburg Sleep Quality Index (PSQI) measures sleep quality [54]. It consists of 19 self-reported questions and assesses daytime dysfunction over the past month as well as subjective sleep quality, sleep latency, sleep length, habitual sleep efficiency, sleep disturbances, and use of sleep medications. A subscale score of 0 to 3 is generated for each component and these seven scores are added together to provide a single overall score that ranges from 0 to 21 points (0 = no difficulty, 21 = severe difficulty), with a score of less than 5 indicating poor sleep quality.
7. the Mutuality scale (MS) evaluates the relationship between the patient and caregiver [55]. It can be utilized to assess mutuality from the viewpoint of the patient or the caregiver. The MS consists of 15 items divided into four dimensions: love and

affection, shared pleasurable activities, shared values, and reciprocity. Each item is scored on a five-point Likert scale ranging from 0 (“not at all”) to 4 (“a great deal”), and the total scale score is calculated by averaging all item scores. Higher scores indicate higher mutuality.

In caregivers, the following secondary outcomes were assessed:

1. the Caregiver Contribution to Self-Care of HF Index (CC-SCHFI) focuses into how much caregivers encourage patients to practice self-care or perform self-care on behalf of the patients when they are unable to do [26]. It consists of 22 items divided into three scales: the CC to self-care maintenance scale, which assesses how much caregivers assist patients in following pharmacological and behavioural prescriptions and monitoring symptoms; the CC to self-care management scale, which evaluates how likely caregivers are to assist patients in responding to symptoms when they occur; and the CC to self-care confidence scale, which assesses caregiver self-efficacy in supporting patients. A standardized total score between 0 and 100 is provided for each measure and better CC to self-care and caregiver self-efficacy are indicated by higher ratings.
2. the SF-12, measuring general physical and mental QOL, as for the patients [51].
3. the Hospital Anxiety and Depression Scale (HADS) as for the patients [53].
4. the Pittsburgh Sleep Quality Index (PSQI) as for the patients [54].
5. the Mutuality scale (MS) to evaluate the relationship with the patient [55].
6. the Caregiver Preparedness Scale (CPS) assesses how well-prepared a caregiver is to meet the needs of the patient both physically and mentally [56]. The CPS consists of eight items using a 5-point Likert scale ranging from 0 (not well prepared) to 4 (very well prepared). The total score ranges from 0 to 4, with higher score indicating better preparedness.

7. Caregivers were questioned about the patient's use of healthcare services (such as hospitalizations and usage of emergency services) as well as all-cause mortality at each follow-up. In particular, research assistants blinded to the study arm assignment, called the caregiver of each patient and asked about the patient's use of healthcare care (emergency services and hospitalizations) related to HF causes (such as dyspnea) in the previous three months as well as whether the patient had eventually died, regardless of the cause. Emergency services utilization for causes other than HF was not taken into account (e.g. use of emergency services for a bone fracture).

All secondary outcomes were evaluated at baseline and 3, 6, 9 and 12 months after the enrolment (Table 1).

3.5 Sample size and randomization

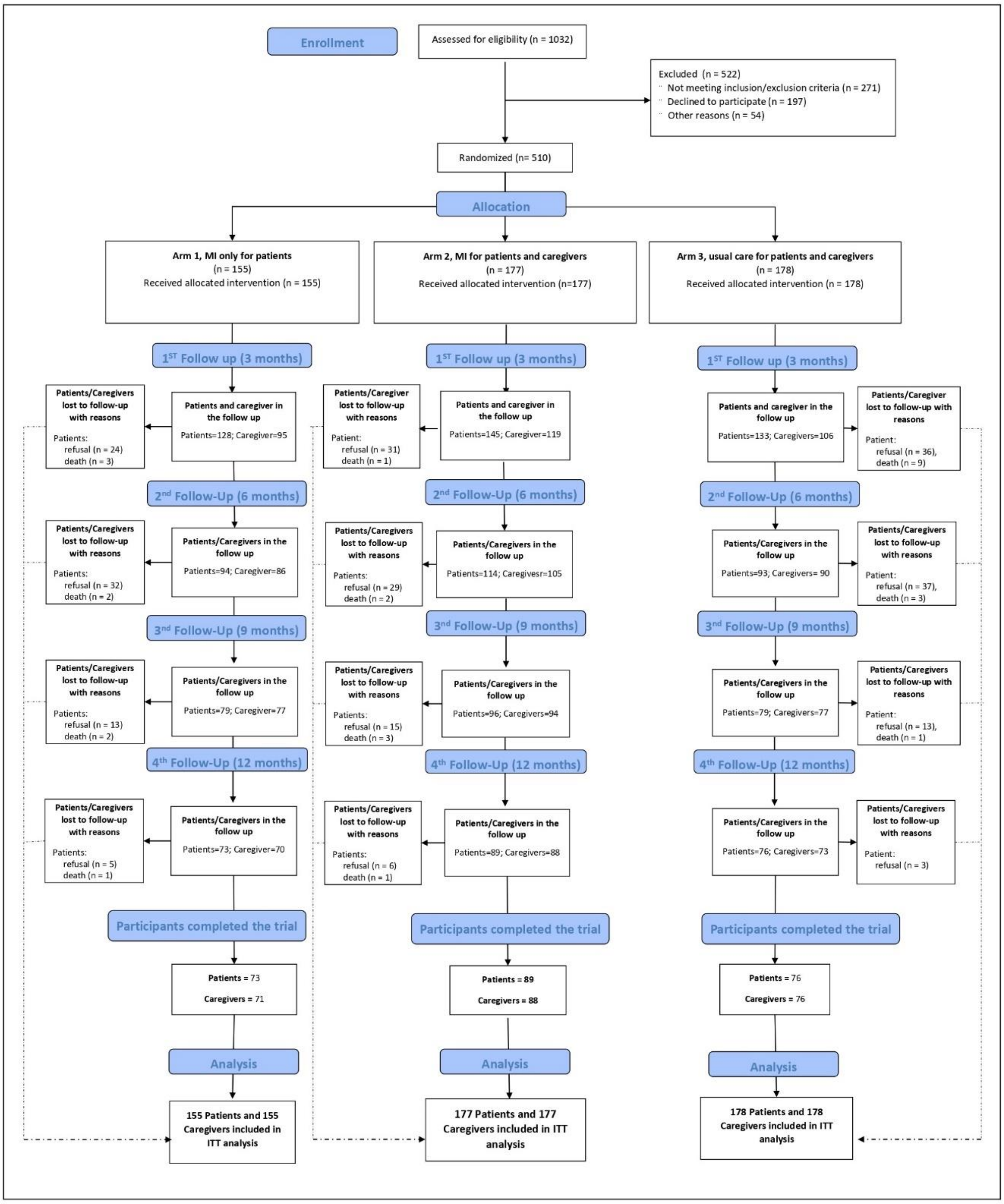
In order to detect an 8% difference in patients' self-care maintenance at three months following the MI intervention (Arms 1 and 2) compared to patients receiving standard care (Arm 3), an overall sample of 240 patients (80 in each arm) was calculated to achieve 83% power. Accounting for an expected 50% attrition rate, 480 patients were planned to be recruited. When considering the evaluation of MI on caregivers (Arm 2), a two-sided two-sample t-test was used to predict that group sample sizes of 80 and 80 would have reached 71% power to detect a difference of 8 points in self-care with a standard deviation of 20 and a significance level of 0.05.

Participants were randomly assigned with a 1:1:1 ratio in the three arms of the study. A block randomization scheme consisting of 15 patient and caregiver dyads was created in order to ensure balance among the three arms. These blocks were randomly included into a list that had 400 random assignments for each enrolment centre. A study assistant then created three containers, one for each centre, placed 400 envelopes in each and put one group assignment in each package (i.e. Arm 1, Arm 2, or Arm 3). The assistant who prepared the containers was no longer working on the project when the three containers were transferred to a different research assistant. The second research assistant opened an envelope each time a patient and caregiver dyad was enrolled in order to reveal the assignment arm. To let the interventionist know which intervention to perform, he then phoned the enrolment centre. The research assistant did not contact the interventionist if the dyad was placed in Arm 3 (the control group). Blinding was successful because no dyads were known to the first research assistant who prepared the envelopes for the three centres, the second research assistant was unable to affect the group assignment, and the interventionists did not collect any data.

3.6 Participant flow

The eligibility of 1032 patient and caregiver dyads was evaluated between June 2014 and October 2018. Following evaluation, 522 of these dyads were excluded for the following reasons: n = 271 patient and caregiver dyads met the exclusion criteria (n = 48 patients were in NYHA I; n = 28 patients did not exhibit inadequate self-care; n = 47 patients had experienced a myocardial infarction in the previous three months; n = 5 patients had severe cognitive impairment; n = 97 caregivers refused to participate despite patients' requests; n = 46 patients resided in residential settings); n = 197 patients declined to participate; n = 54 declined for unknown reasons. Intention-to-treat analyses comprised each of the 510 HF patient and caregiver dyads that had been enrolled and randomly assigned [57]. The analysis of the primary endpoint included a total of 406 individuals who completed T1 (3 months after enrolment), resulting in a patient attrition of 20.4% and a caregiver attrition of 37.2%. 238 HF patients and 235 caregivers completed the full 12-month trial (Figure 4).

Figure 4. Participant flow of the MOTIVATE-HF randomized controlled trial



Note. Figure published in *Vellone E, Rebora P, Ausili D, et al. Motivational interviewing to improve self-care in heart failure patients (MOTIVATE-HF): a randomized controlled trial. ESC Heart Fail. 2020;7(3):1309-1318.*

4. Statistical methods

4.1 Longitudinal data analysis

A longitudinal study is designed to measure intra-individual change over time. Individual subjects are observed repeatedly, typically at a series of regular intervals defined by the study protocol. One of the fundamental goals of longitudinal research is to relate change over time in individuals to an experimental condition (such as drug treatment arm, or motivational interview, etc.), or to their characteristics (exposure, sex, etc.). Each subject is observed in repeated measures studies; variations in response resulting from conditions are the main focus. Measurements, counts, or dichotomous indicators can be used as outcomes. Multivariate outcomes can also be measured at several times [58].

All subjects should be measured at the same times in the ideal situation because this makes analysis and interpretation much easier. Design studies can be distinguished between balanced and unbalanced. A design is said to be balanced when N individuals are measured at the same n occasions and unbalanced otherwise. In some situations, getting measurements of every subject at the same time may be exceedingly challenging. This is particularly true when studying individuals for a long time of follow-up, especially in case of clinic populations where patient availability is significantly impacted by illness. The mistiming and/or missingness of observations is possible, and subjects may withdraw or stop being available for observation.

4.1.1 The correlated data

Considering a sample of N selected subjects with n_i measurements of response on each subject, $i = 1, \ldots, N,$ each longitudinal response can be expressed as Y_i independent vectors. Associated with the jth occurrence on the ith subject, $j = 1, \ldots, n_i$, p represents the number of covariates. In the design matrix X_i, the rows correspond to the predictors associated with the different times of measurements, and the columns correspond to the different variables [58]. Thus each individual has a vector of outcomes Y_i and a matrix of covariates X_i. Since j

represents occasions of measurements, the covariate X_{ij} may include functions of explanatory variables measured at time j.

1. The kth covariate, $1 \leq k \leq p$, is between-cluster or time-invariant if, for all $i = 1, \ldots, N$,

   $X_{i1k} = X_{i2k} = \cdots = X_{in_ik}$

   Examples include sex and fixed experimental conditions such as treatment assignment in a longitudinal clinical trial.

2. The kth covariate, $1 \leq k \leq p$, is within-cluster or time-varying if, for some $i = 1, \ldots, N$,

   $X_{ijk} \neq X_{ij'k}$ for at least one pair j, j' and $j \neq j'$.

   Examples include current smoking status or weight in a longitudinal study. In some cases (pure repeated measures designs, or longitudinal studies with fixed time points), these covariates vary systematically in the same way for each subject so that $X_{ijk} = X_{i'jk}$ for all i and j, for fixed k.

4.1.2 Model formulation

The matrix of the data can be written as follows:

$$\begin{pmatrix} Y_{i1} \\ Y_{i2} \\ \vdots \\ Y_{in_i} \end{pmatrix} = \begin{pmatrix} X_{i11} & X_{i12} & \cdots & X_{i1P} \\ X_{i21} & X_{i22} & \cdots & X_{i2P} \\ \vdots & \vdots & \ddots & \vdots \\ X_{in_i1} & X_{in_i2} & \cdots & X_{in_iP} \end{pmatrix} \begin{pmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_P \end{pmatrix} + \begin{pmatrix} e_{i1} \\ e_{i2} \\ \vdots \\ e_{in_i} \end{pmatrix}$$

The model relating the response Y_i to the covariates X_i has the following assumptions [58]:

1. $(Y_1, X_1), \ldots, (Y_N, X_N)$ are independently distributed.
2. Given X_i, $E(Y_i) = X_i \beta$.
3. Given X_i, $Cov(Y_i) = \Sigma_i$, where Σ_i is some known function of the covariates X_i.

Assumption 1 states that N independently selected units make up the sample, while assumption 2 states that the conditional mean of the jth outcome of unit i given $X_{i1}, \ldots, X_{in_i}$ is a linear function of X_{ij} only, i.e.,

$$E(Y_{ij} | X_i) = X_{ij}^T \beta = \beta_1 X_{ij1} + \cdots + \beta_p X_{ijp}.$$

Assumption 3 allows for dependencies among measurements on the same unit. The covariance may differ with covariates, e.g., across groups (such as sex), or entries may be functions of time. Table 2 shows various Σ structure types for n times observation. Each diagonal term presents the variance of repeated observations each time, while elements in the above diagonal term present the covariance of repeated observations between times [58]. In the Unstructured (UN) matrix each variance-covariance term will be estimated separately. This is the most used Σ structure for longitudinal data, because it doesn't make any assumptions about the error terms and allows for a unique pattern of correlation between observations. Both CS and AR(1) make the assumption that all variances are constant over time, with CS assuming constant correlation over time and AR(1) assuming exponentially decreasing correlation. The CSH and ARH(1) models are possible expansions of the CS and AR(1) covariance models, respectively. Since the structural models AR(1), CS, ARH(1), and CSH have a constant parameter, they can be more effective and powerful Σ structures for identifying treatment effects [58].

Table 2. Structures of the covariance matrix for *n* times observation

Type	Matrix R	Number of Parameters
Unstructured (UN)	$UN = \begin{bmatrix} \sigma_1^2 & \cdots & \sigma_{1n} \\ \vdots & \ddots & \vdots \\ \sigma_{n1} & \cdots & \sigma_n^2 \end{bmatrix}$	$n\ (n + 1) / 2$
Compound Symmetric (CS)	$CS = \sigma^2 \begin{bmatrix} 1 & \cdots & \rho \\ \vdots & \ddots & \vdots \\ \rho & \cdots & 1 \end{bmatrix}$	2
Heterogenous Compound Symmetric (CSH)	$CSH = \begin{bmatrix} \sigma_1^2 & \cdots & \sigma_1\sigma_n\rho \\ \vdots & \ddots & \vdots \\ \sigma_1\sigma_n\rho & \cdots & \sigma_n^2 \end{bmatrix}$	$n + 1$
First-order Autoregressive (AR(1))	$AR(1) = \sigma^2 \begin{bmatrix} 1 & \cdots & \rho^{n-1} \\ \vdots & \ddots & \vdots \\ \rho^{n-1} & \cdots & 1 \end{bmatrix}$	2
Heterogenous First-order Autoregressive (ARH(1))	$ARH(1) = \begin{bmatrix} \sigma_1^2 & \cdots & \sigma_1\sigma_n\rho^{n-1} \\ \vdots & \ddots & \vdots \\ \sigma_1\sigma_n\rho^{n-1} & \cdots & \sigma_n^2 \end{bmatrix}$	$n + 1$

In clinical trials, the key objective is to characterize the patterns of change in the mean response over time in the intervention and control groups and to determine if the shapes of the mean response profiles differ between the groups. Concerns about whether change patterns are the same across groups may be translated into hypotheses about the interaction between the group factor and time [58].

4.1.3 Missing data and dropout

A frequent and challenging issue in the analysis of longitudinal data is missing data [58]. Even though the majority of longitudinal studies are set up to collect information on every member in the sample at each point of follow-up, many of them have some missing observations. This is also the case of the MOTIVATE-HF trial. A vast variety of diverse missingness patterns are produced when a person's response is missing at one follow-up time and then measured at a subsequent follow-up time. A different issue that frequently affects longitudinal research is attrition, or "dropout", which occurs when some participants leave the study before it is finished. The word "missing data" is used in both scenarios to describe the inability to achieve an intended measurement.

Three important consequences on the statistical analysis of longitudinal data arise from missing data. First, when there are missing data, the data set is unbalanced over time since not every person has the same amount of repeated measurements at the same set of events. However, the regression techniques that will be discussed in the following paragraphs do not have a problem with the imbalance brought on by missingness. Second, there must inevitably be some information loss when there are missing data. In other words, the precision of estimated changes in the mean response over time is reduced by missing data. It results that the precision loss is directly proportional to the amount of missing data; that is, the precision loss increases as the amount of missing data increases. Third, in some cases missing data can create bias and hence result in false conclusions about changes in the mean response, leading a complication in the analysis of partially missing longitudinal data. The causes of any missing data, often known as the missing data mechanism, must therefore be carefully taken into account [58].

As mentioned before, the majority of longitudinal studies are designed to collect information on every member of the sample at predetermined intervals. However, attrition is a persistent issue in longitudinal research; as a result, some participants "drop out" of the study in advance. Dropout describes a specific situation in which if Y_{ij} is missing, then $Y_{ij+1}, ..., Y_{in}$ are missing as well. This results in the monotone missing data pattern, in contrast to the non-monotone patterns that can develop when data are missing intermittently, giving rise to a significantly greater number of potential missing data patterns [58].

When a longitudinal study experiences dropout, the fundamental concern is whether individuals who "drop out" and those who continue to participate in the study differ in any other pertinent way. If not, studies that are limited to the remaining participants produce valid inferences. In contrast, if they differ, such "complete-case" analyses could be biased. As well as missing data, dropout can occur *completely at random*, *at random*, or *not at random*. When dropout is completely at random, the probability of dropout on each occasion is independent to any outcomes from the past, current, or future (given the covariates). On the other hand, when dropout is at random, the probability of dropout at each occasion can vary depending on the outcomes that have already been recorded up to, but excluding the current occasion. Finally, when dropout is not at random, the probability of dropping out on any given occasion can depend on current and future unobserved outcomes. In other words, the process is dependent on the unrecorded values of the outcome variable that would have been observed if the participant remained in the study [58].

The proper handling of mortality as a dropout factor is one of the trickier problems in a longitudinal data analysis [59]. Mortality cannot really be considered a dropout when the outcome is self-care or mutuality. Dealing with quality of life in EQ-5D-3L1 scale [60] was proposed to give score 0 to dead subjects, in an index score that generally ranges from 0 (where 0 is a health state equivalent to death) to 1 (perfect health), with higher scores indicating higher health utility. Both Rubin and Frangakis [61] and Robins et al. [62] presented two potential strategies, which essentially entailed inferences about the subpopulation of people who would survive or who have a non-zero probability of surviving to a specific time *t*.

4.2 Linear mixed models

4.2.1 Fixed and random effects

A statistical method known as a linear mixed model (LMM) analysis offers a flexible approach for statistical analyses of correlated longitudinal data. Response variables that are repeatedly measured for each unit analysis and time as the repeated factor are used to characterize longitudinal data. As described before, for each unit analysis in a longitudinal data structure, the number of repeated measurements may not be equal, and the time of observation may not also occur at the same intervals or spacings. In other words, the features of longitudinal data allow for observations with missing values and unequal space.

The fundamental premise of linear mixed effects models is that some subsets of the regression parameters vary randomly from one subject to another, effectively accounting for the sources of population heterogeneity. That is, it is assumed that each subject of the population has their own subject-specific mean response trajectories over time and a subset of the regression parameters are now regarded to be random. The term *mixed* in LMM refers to the simultaneous use of and analysis of the fixed and random effects in the model. *Fixed effects* can represent systematic mean patterns, such as treatment approaches. In contrast, *random effects* can account for correlation patterns between repeated measurements within subjects, heterogeneities between subjects, or both [58].

4.2.2 Within-group and between-group levels

Multilevel modelling is the integration of the within-group (or within-subject) model with the between-group (or between-subject) model for longitudinal data. At first level, the mean response trajectory of each individual is modelled as a function of the same set of covariates where linear regression of the observations is defined. The response variable is typically modelled as a function of time and a within subject residual component [58]. The level-1 within-subject model is written as follows in matrix notation:

$$Y_i = Z_i\beta_i + e_i, \quad e_i \sim N\left(0,\, \sigma^2 I_{n_i}\right).$$

Regardless of the number of longitudinal responses n_i, the number of subject-specific regression coefficients β_i (fixed effects) is the same and can be interpreted as the "true" regression coefficients for the i^{th} subject. Alternatively, $Z_i\beta_i$ might be viewed as the "true" underlying mean response trajectory for the i^{th} subject. When regarded in this way, the longitudinal responses on the i^{th} subject are assumed to follow the subject-specific response trajectory provided by $Z_i\beta_i$, with the addition of measurement or sampling errors e_i.

At the second level, the subject-specific regression coefficients β_i are modelled in terms of population averages and subject deviations from these parameters. The coefficients β_i are used as random dependent variables in the level-2 between-subject model in the longitudinal context. Specifically, variation in β_i from one subject to another is modelled as a function of a set of between-subject (or time-invariant) covariates (e.g., gender, treatment group) [58]. The model can be written as follows using matrix notation:

$$\beta_i = A_i\beta + b_i, \;\; b_i \sim N(0, G)$$

where b_i (random effects) has a multivariate normal distribution with mean zero and covariance matrix G. Here, the b_i yield the regression coefficients from a subject's residual trajectory over time, after the effects of the covariates have been taken into account. In other words, the b_i represent the deviation from the population mean response of the i^{th} subject.

We obtain the following formula as the matrix notation of LMM by replacing the formula of level-2 between-subject model in the formula of level-1 within-subject model:

$$\begin{aligned} Y_i &= Z_i\beta_i + e_i \\ &= Z_i(A_i\beta + b_i) + e_i \\ &= (Z_iA_i)\beta + Z_ib_i + e_i \\ &= X_i\beta + Z_ib_i + e_i, \end{aligned}$$

where $X_i = Z_i A_i$. When averaged over the random effects b_i,

$$E(Y_i) = (Z_i A_i)\beta = X_i\beta, \text{ and}$$

$$Cov(Y_i) = Z_i G Z_i' + \sigma^2 I_{n_i}$$

A longitudinal model with only random intercepts (or random subject-effects) implies a compound symmetry covariance matrix with constant variance over time, and constant correlation among pairs of repeated measurements [58]. In presence of a more complex random effects covariance structure (e.g., as random intercepts and slopes), linear transformations of components of Z_i produce equivalent mixed effects models only when the covariance matrix, G, has been left unstructured. When G is unstructured, the appropriate changes to the variances and covariances of the random effects can be produced [58].

4.2.3 Prediction of random effects and goodness of fit

The focus of inference in longitudinal data is on the fixed effects β. Since the linear mixed effects model distinguishes between fixed and random effects, we can also estimate (or predict) subject-specific response trajectories over time. Therefore, predictions of the subject-specific effects, b_i, or the subject-specific response trajectories, $X_i\beta + Z_i b_i$, are possible. Technically, we "predict" the random effects rather than "estimate" them because the b_i are random variables rather than fixed population parameters [58].

In general, the problem of predicting a random variable reflects predicting its conditional mean, given the available data. Thus the best predictor of b_i is the conditional mean of b_i, given the vector of responses Y_i (and $\hat{\beta}$),

$$E(b_i|Y_i) = G Z_i' \Sigma_i^{-1}\left(Y_i - X_i\hat{\beta}\right),$$

where $\Sigma_i = \text{Cov}(Y_i) = Z_i G Z'_i + R_i$. This is known as the “Best Linear Unbiased Predictor” (BLUP). This predictor of b_i depends on the unknown covariance among the Y_i. When the unknown covariance parameters are replaced by their Restricted Maximum Likelihood (REML) estimates, the resulting predictor

$$\widehat{b_\iota} = \hat{G} Z_i' \hat{\Sigma}_i^{-1}\left(Y_i - X_i\hat{\beta}\right),$$

is referred to as the “empirical BLUP”. Given $\widehat{b_\iota}$, we can also obtain the i^{th} subject’s predicted response profile:

$$\widehat{Y_\iota} = X_i\hat{\beta} + Z_i\widehat{b_\iota}.$$

Transformed residuals from the linear mixed effects model can be obtained by taking the Cholesky decomposition of $\widehat{\Sigma_\iota}$, so that they have constant variance and zero correlation. The adequacy of the random effects covariance structure can be assessed from the plot of the empirical semi-variogram for the transformed residuals. For longitudinal data the semi-variogram is defined as one-half the expected squared difference between residuals obtained on the same individual [58]. Furthermore, by using linear mixed effects models, we can predict the random effects (empirical BLUPs) and look at their distribution for any evidence of outliers, which may reflect subjects with slightly atypical subject-specific response profiles. Using histograms and normal quantile plots of the empirical BLUPs to evaluate the adequacy of the normal distribution assumption for the random effects should be done with caution because the empirical BLUPs are known to be strongly influenced by the normal distribution assumption for the random effects [58].

4.3 Longitudinal dyad models

When longitudinal measures are taken from a pair on dyadic outcomes, the multivariate outcomes model described by Raudenbush et al. [43] gives estimates of the average partner trajectories as well as the heterogeneity across dyads around the average trajectories. Because the equations are fitted concomitantly, and the within-dyad dependence of the observations is taken into account, it differs from analysis on separate samples of partners[46].

4.3.1 Level 1 model: within-dyad model

The repeated measures for both dyad members are shown in the Level 1 or within-dyad model as functions of time, together with a residual term (r) that accounts for measurement error, or the time-specific error of prediction. Within and across dyads, the variance of these measurement errors is thought to remain constant. The model is written as follows for care dyads composed, as in the MOTIVATE-HF RCT, of a HF patient and a caregiver:

$$Y_{tp} = (HF)\left[\beta_{1p} + \beta_{2p}\left(LIN_{tp}\right)\right] + (CG)\left[\beta_{3p} + \beta_{4p}\left(LIN_{tp}\right)\right] + r_{tp}$$

where Y_{tp} is the outcome score ($t = 1, ..., K$ responses per dyad and time measurement) for dyad p. HF is an indicator variable that has a value of 1 if the response came from an HF patient and a value of 0 if it came from a caregiver. CG is an indicator variable that has a value of 1 when a caregiver provided the response, and a value of 0 when a HF patient provided the response. The coefficients and variables are multiplied by the corresponding indicator variable, as shown by the brackets. Therefore, the trajectory of the HF patient is represented by the latent growth parameters β_{1p} and β_{2p} (intercept and linear component of time), whereas the trajectory of the caregiver is represented by the latent growth parameters β_{3p} and β_{4p}. As a result, each dyad includes four coefficients (β_{1p}, β_{2p}, β_{3p}, and β_{4p}), which stand for the dyad's actual growth parameters. The within-dyad residuals, or level 1 random effects, are known as the r_{tp}. It is presumed that they are normally distributed, with a mean of 0 and a variance of σ^2.

The repeated measures situation might make the assumption of constant variance σ^2 implausible, and it should be assessed using a test for homogeneity of variance. This assumption is typically tested by fitting a model with a different structure for the Level 1 error component, such as heterogeneous variance, which enables the estimation of a different variance at each time point. The researcher can assess the reliability of the assumption of constant variance using a likelihood ratio test that compares the deviance statistic of the alternative model against the fit of the more restrictive model (i.e., homogenous variance). The alternative model is preferable in all consequent studies if it considerably improves the fit to the data to prevent misspecification of the Level 1 model [63].

4.3.2 Unconditional Level 2 model: between-dyad model

The Level 2 outcomes are represented by the Level 1 coefficients. These outcomes are allowed to vary across all Level 2 units, which are the dyads, so each dyad may have a different value. Fitting an unconditional model, or a model without any predictor variables, at Level 2 is the initial stage in any longitudinal analysis. The details of the model are as follows:

$$\beta_{1p} = \gamma_{10} + u_{1p}$$

$$\beta_{2p} = \gamma_{20} + u_{2p}$$

$$\beta_{3p} = \gamma_{30} + u_{3p}$$

$$\beta_{4p} = \gamma_{40} + u_{4p}$$

The unconditional between-dyad model offers estimates of the population averages for each growth parameter related to the HF patient (γ_{10} and γ_{20}), as well as for the caregiver (γ_{30} and γ_{40}). These estimates, which are referred to in the model as the fixed effects, represent the means of the distribution of the corresponding coefficients across dyads. The deviation of each dyad from the corresponding population average growth parameter is represented by the Level 2 random effects (u_{1p}, u_{2p}, u_{3p} and u_{4p}). The variances of these random effects can be estimated and serve to illustrate the heterogeneity between dyads. Whether any variance in the population differs significantly from zero can be tested and if any are, significant variability between dyads exists, and predictors can be introduced to explain this variability.

4.3.3 Conditional Level 2 model: between-dyad model

Predictors are added to the conditional model to explain the variance in the trajectories across dyads. The predictors could be variables with specific values for each member (e.g., age) or with common values for each dyad member. For example, the caregiver in the dyad could live with the HF patient or not. The details are as follows in the level 1:

$$Y_{tp} = (HF)\left[\beta_{1p} + \beta_{2p}\left(LIN_{tp}\right) + \beta_{11p}\, Predictor_1 + \cdots + \beta_{1np}\, Predictor_n\right] + (CG)\left[\beta_{3p} + \beta_{4p}\left(LIN_{tp}\right) + \beta_{21p}\, Predictor_1 + \cdots + \beta_{2np}\, Predictor_n\right] + r_{tp}$$

while in level 2:

$$\beta_{1p} = \gamma_{10} + u_{1p} \qquad \beta_{11p} = \gamma_{11}$$

$$\beta_{2p} = \gamma_{20} + u_{2p} \qquad \beta_{1np} = \gamma_{1n}$$

$$\beta_{3p} = \gamma_{30} + u_{3p} \qquad \beta_{21p} = \gamma_{21}$$

$$\beta_{4p} = \gamma_{40} + u_{4p} \qquad \beta_{2np} = \gamma_{2n}$$

The estimates of the population averages for each growth parameter for the HF patient (γ_{10} and γ_{20}) and for the caregiver (γ_{30} and γ_{40}) are provided by the conditional between-dyad model and have been adjusted for the effects of the predictors in each equation. The fixed effects of each predictor are reflected by the corresponding regression coefficient $\gamma_{11}, \ldots, \gamma_{1n}$ and $\gamma_{21}, \ldots, \gamma_{2n}$ which show how each predictor and the corresponding growth parameter are related. The conditional deviation of each dyad from the corresponding population average growth parameter, or the unexplained residual variance in each parameter, is now represented by the Level 2 random effects u_{1p}, u_{2p}, u_{3p} and u_{4p}

4.4 Statistical analyses in the MOTIVATE-HF trial

4.4.1 Baseline characteristics and change in scores during follow-up

The characteristics of the HF patients and caregivers in the MOTIVATE-HF trial, as well as the primary and secondary outcomes, were described using measures of central tendency and variability. For continuous data, baseline characteristics were presented as medians and quartiles or as means and standard deviations (SDs), while for categorical data, they were expressed as absolute numbers and their frequencies (%). The characteristics of the HF patients and caregivers were presented overall and by treatment arm.

The difference (Δ) between the scores related to primary and secondary outcomes at each follow-up period (T1, T2, T3, and T4) and the same score at baseline (T0) was used to calculate the change in scores during follow-up. For patients, a two-sample t-test was used to compare the change of each score in intervention Arms 1 and 2 compared to the control Arm 3. For caregivers' endpoints, the change in Arm 2 (the only arm in which the caregivers received the intervention) was compared to Arms 1 and to Arm 3 using two independent two-sample t-tests.

Regarding all-cause mortality and health-care service use (hospitalizations and emergency service use) among patients, they were summarized as absolute numbers and frequencies (%) among the three study arms at each follow-up interval (3, 6, 9, and 12 months after enrolment). Fisher's exact test was used to compare statistical differences in health-care service use and all-cause death between the three arms at each follow-up.

4.4.2 Change in scores over time using linear mixed models

Longitudinal linear mixed models were used to analyse changes over time (from baseline to T4) in primary and secondary outcomes in patients and caregivers to account for dropout and missing values. The outcome from T0 to T4 was included as response variable. The dependence within subject (patient or caregiver) was accounted for by including a random intercept and slope. The visit number, the randomization arm and their interaction were included in the models as independent variables. Visit number was included as a continuous variable t_{ij}, with value from 0, at baseline, to 4 at 12 months, because it express the rate of change per 3 months in the outcomes that were assumed to be linear over the 12-month period:

$$Y_{ij} = \beta_0 + \beta_1 t_{ij} + \beta_2\, Arm1_i + \beta_3\, Arm2_i + \beta_4 t_{ij} * Arm1_i + \beta_5 t_{ij} * Arm2_i + b_{0i} + b_{1i} t_{ij} + e_{ij}$$

where Y_{ij} represents the outcome for patients (self-care maintenance, management, confidence, anxiety and depression, generic and disease-specific quality of life, sleep quality, HF symptom perception) and caregivers (caregiver contribution to self-care, preparedness, anxiety and depression, generic quality of life, sleep quality). Furthermore, $Arm1_i$ = 1 if the

i^{th} subject was assigned to MI only for patients, $Arm2_i = 1$ if the i^{th} subject was assigned to MI for patients and caregivers, and, simultaneously, $Arm1_i = 0$ and $Arm2_i = 0$ if the i^{th} subject was assigned to standard of care. The interaction between arm and visit number was used to test the efficacy of MI.

When analysing patients, Arm 3 was the reference of the randomization arm. When analysing caregivers, Arm 2 was the reference of the randomization arm, and the cohabitation with the patient (as a dichotomous variable) was also included to account for baseline imbalances in this variable among the three arms.

A longitudinal generalized linear mixed model with logit link was applied to assess whether the three groups differed in their use of health-care services during follow-up. Once again, the dependence of health-care service use among different visits on the same patient was accounted for by the inclusion of a random intercept and slope in the model. It included the visit number as a categorical variable with five levels (to account for non-linearity), the randomization arm, and the interaction between the study arm and visit number.

The association between treatment arm and all-cause mortality was examined using an unadjusted Cox proportional-hazards regression model. Schoenfeld residuals and a graphic evaluation of proportionality of hazard were also used. In the case of non-proportionality, time was split at follow-up time determined by graphical evaluation in a time-dependent Cox model, providing hazard ratios (HRs) with 95% confidence intervals (CIs) for each time interval.

4.4.3 Change in mutuality over time using dyadic models

Multilevel modelling was used to analyse the Mutuality Scale (MS) at the level of the patient and caregiver dyad to control for interdependencies in the data. A longitudinal dyad model was tested for each domain of the MS (i.e., love and affection, shared pleasurable activities, shared values, and reciprocity) and for the total score. These were linear models of MS within dyads over time, which estimated the population averages of the MS for both patients and caregivers (fixed effects), the interdependence between the MS of the members of the patient–caregiver dyad (tau correlations), and the variability around the average trajectories of the MS for both members (random effects). The models included, as for the models analysing changes over time of the individual scores, the treatment arm, its interaction with the visit number and the patient–caregiver living together condition as covariates, for both patients and caregivers.

The details are as follows in the level 1 model:

$$\begin{aligned} MS_{ij} = (HF)\big[\beta_1 + \beta_2 t_{ij} + \beta_{11}\, Arm1_i + \beta_{12}\, Arm2_i + \beta_{13} t_{ij} * Arm1_i + \beta_{14} t_{ij} * Arm2_i \\ + \beta_{15}\, LivTog_i\big] \\ + (CG)\big[\beta_3 + \beta_4 t_{ij} + \beta_{21}\, Arm1_i + \beta_{22}\, Arm2_i + \beta_{23} t_{ij} * Arm1_i + \beta_{24} t_{ij} \\ * Arm2_i + \beta_{25}\, LivTog_i\big] + r_{tp} \end{aligned}$$

while in level 2 model:

$$\begin{aligned} &\beta_1 = \gamma_{10} + u_1 && \beta_{11} = \gamma_{11}, \dots, \beta_{15} = \gamma_{15} \\ &\beta_2 = \gamma_{20} + u_2 && \\ &\beta_3 = \gamma_{30} + u_3 && \beta_{21} = \gamma_{21}, \dots, \beta_{25} = \gamma_{25} \\ &\beta_4 = \gamma_{40} + u_4 && \end{aligned}$$

where MS_{ij} is the outcome score, mutuality, for dyad $i = 1, \dots, 510$ and $j = 1, ..., 10$ responses per dyad and time measurement. As mentioned above, HF is an indicator variable taking a value of 1 if the response came from an HF patient, and 0 from a caregiver, and CG is an indicator variable taking a value of 1 when a caregiver provided the response, and 0 when a HF patient did it. $LivTog_i$ is a dichotomous variable taking a value of 1 if the patient and the caregiver live together, and 0 otherwise.

5. Results

5.1 Participants' characteristics

Table 3 displays the baseline characteristics of patients and caregivers, categorized by study arm.

Patients were generally male (58%), with a median age of 74. They were mostly retired (76.2%). The majority of patients were in NYHA Class II (61.9%) and 33.6% had an ischemic HF aetiology. At baseline, all the sociodemographic and clinical characteristics of each participant in the three Arms were comparable, as well as the self-care levels [47]. At baseline, the mean self-care maintenance, management, and confidence scale scores were 45.5, 39.7, and 51.4, respectively, indicating that the scores on the self-care scale were mainly inadequate, since adequate self-care is defined as a score ≥ 70. At baseline, the levels of anxiety, depression, sleep quality (PSQI), and quality of life (QOL) in the three groups were comparable [64]. In particular, poor sleep quality and moderate levels of anxiety and depression were found. The assessment of both general and disease-specific QOL revealed a moderate symptom burden [65] (Table 3).

The median age of caregivers was 55, and 75.5% of them were female. Among them, 72% were married, 73.5% employed and 60% living with the patients. The only difference between the randomized caregivers in the three arms was the variable "caregivers living with patients", which was more prevalent in Arm 2 than in Arms 1 and 3 ($p = 0.001$). With the exception of the CC to self-care maintenance scores, which were significantly higher in Arm 2 than Arm 3 ($p = 0.006$), the mean scores of the CC to self-care scores ranged between 48.1 and 60.0 and were comparable across the three arms, as well as the CPS scores [66]. The mean anxiety scores ranged of 7.3 to 7.7 among the three arms, while the mean depression scores ranged between 5.7 and 6.1, indicating low levels for both scales. Physical and mental QOL mean scores ranged between 48.3 and 49.2, and between 48.0 and 49.6, respectively (Table 3). All the above scores relative to the caregivers were comparable among the three Arms.

Table 3. Baseline characteristics of HF patients and caregivers (n = 510)

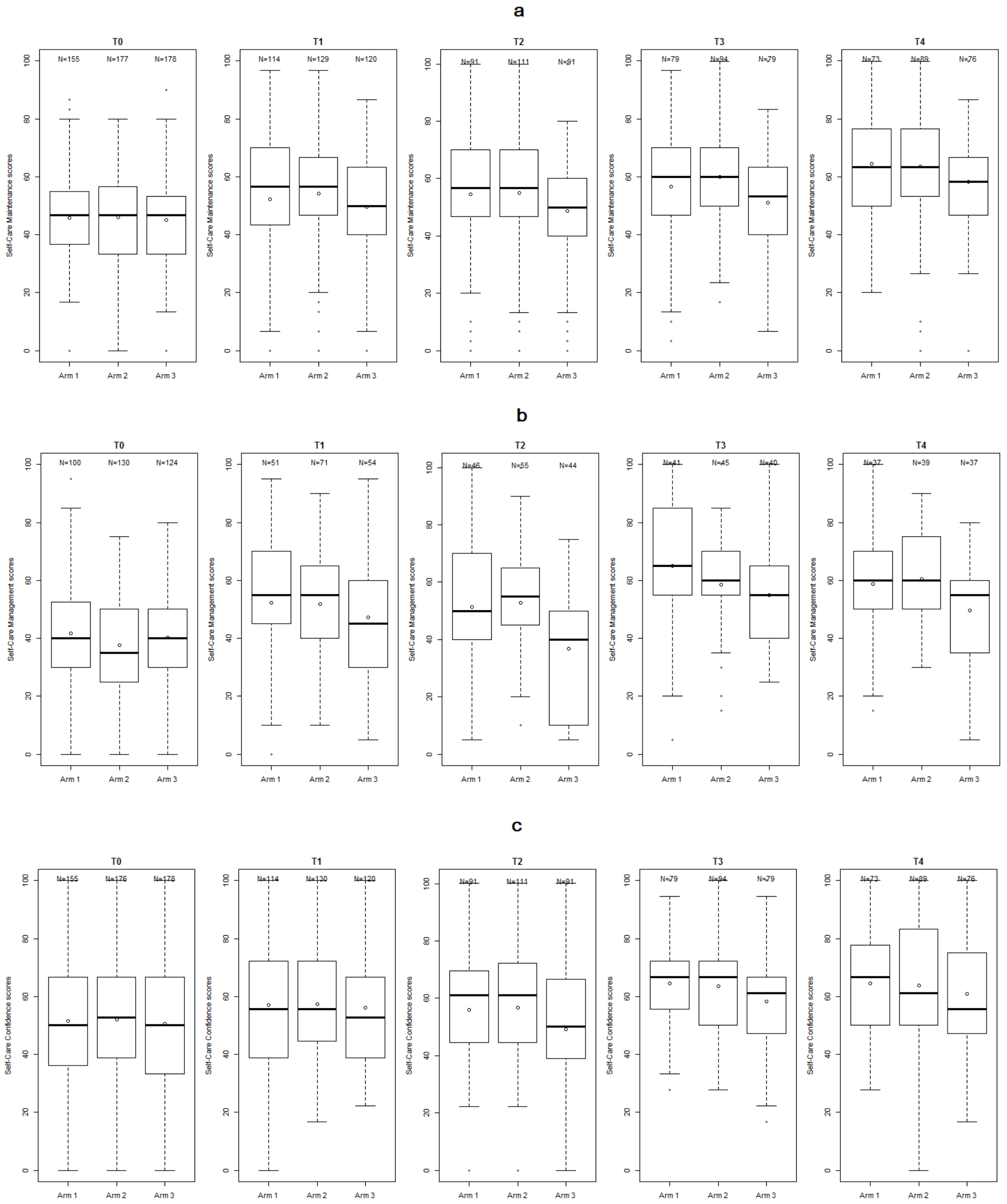
	Arm 1: MI only for patients (*n* = 155)		**Arm 2: MI for patients and caregivers (*n* = 177)**		**Arm 3: Usual care for patients and caregivers (*n* = 178)**	
	Patients	**Caregivers**	**Patients**	**Caregivers**	**Patients**	**Caregivers**
Characteristics	**Median [Q1-Q3]**	**Median [Q1-Q3]**	**Median [Q1-Q3]**	**Median [Q1-Q3]**	**Median [Q1-Q3]**	**Median [Q1-Q3]**
Age	74 [65–82]	54 [44–64]	73 [64–81]	57 [44–68]	75 [64–83]	53 [42–64]
Time with HF (months)	36 [24–72]	——	36 [15–84]	——	48 [20–96]	——
Caregiving duration (months)	——	2 [2-4]	——	2 [1-4]	——	3 [1-4]
N. of medications	6 [4–8]	——	7 [5–9]	——	6 [4–8]	——
CCI scores	2 [2–4]	——	2 [2–4]	——	2 [1–4]	——
MoCA Scores	25 [21–27]	——	26 [19–28]	——	24 [18–27]	——
	***n* (%)**	***n* (%)**	***n* (%)**	***n* (%)**	***n* (%)**	***n* (%)**
Sex (male)	80 (51.6)	36 (24.0)	107 (60.5)	42 (23.9)	109 (61.2)	45 (25.4)
Marital status						
Married	81 (52.3)	108 (72.5)	123 (69.5)	124 (70.5)	112 (62.9)	129 (72.9)
Widower	55 (35.5)	6 (4.0)	44 (24.9)	3 (1.7)	51 (28.7)	3 (1.7)
Divorced	10 (6.5)	10 (6.7)	4 (2.3)	14 (8.0)	6 (3.4)	12 (6.8)
Single	9 (5.8)	25 (16.8)	6 (3.4)	35 (19.9)	9 (5.1)	33 (18.6)
Education (high schools or higher)	41 (26.4)	90 (59.6)	44 (24.8)	86 (49.4)	47 (26.4)	99 (56.2)
Employment (retired)	119 (76.8)	33 (22.0)	137 (77.8)	50 (28.4)	131 (74.0)	52 (29.4)
Income						
Not the necessary to live	7 (4.5)	——	7 (4.0)	——	8 (4.5)	——
The necessary to live	131 (84.5)	——	138 (78.0)	——	141 (79.2)	——
More than the necessary to live	17 (11.0)	——	32 (18.1)	——	29 (16.3)	——
Caregiver living with patient	——	76 (51.0)	——	126 (71.6)	——	104 (58.8)
Relationship with the patient						
Spouse	——	43 (28.7)	——	82 (46.6)	——	64 (36.2)
Child	——	70 (46.7)	——	62 (35.2)	——	64 (36.2)
Other	——	37 (24.7)	——	32 (18.2)	——	49 (27.7)
NYHA Class						
II	98 (63.2)	——	108 (61.7)	——	107 (60.8)	——
III	49 (31.6)	——	55 (31.4)	——	56 (31.8)	——
IV	8 (5.2)	——	12 (6.9)	——	13 (7.4)	——

	Mean (SD)	Mean (SD)	Mean (SD)	Mean (SD)	Mean (SD)	Mean (SD)
N. of medications	6.3 (2.6)	— —	7.0 (2.9)	— —	6.4 (3.0)	— —
Self-care maintenance scores	45.7 (15.2)	— —	45.9 (16.3)	— —	44.9 (14.6)	— —
Self-care management scores [a]	41.7 (17.9)	— —	37.6 (18.4)	— —	40.3 (16.4)	— —
Self-care confidence scores	51.5 (20.9)	— —	52.0 (21.2)	— —	50.6 (22.5)	— —
Hospital Anxiety Scale (HAS)	7.8 (4.5)	7.7 (4.5)	7.7 (4.3)	7.3 (4.4)	8.0 (4.4)	7.5 (4.7)
Hospital Depression Scale (HDS)	7.9 (4.5)	5.7 (4.5)	7.6 (4.1)	5.9 (4.1)	8.4 (4.6)	6.1 (4.6)
Global PSQI Score	12.4 (3.9)	9.8 (3.6)	12.3 (3.5)	10.1 (3.5)	12.2 (3.7)	9.7 (3.0)
SF-12 Physical Component Summary (PCS)	36 (10.3)	49.2 (8.0)	35 (9.2)	48.8 (8.5)	35.4 (9.3)	48.3 (8.3)
SF-12 Mental Component Summary (MCS)	44.4 (10.5)	48.1 (9.1)	45.2 (9.3)	49.6 (9.4)	44.5 (10.7)	48 (9.1)
KCCQ Overall Summary Score	48.6 (23.4)	— —	50.3 (22.1)	— —	50 (23.3)	— —
HFSPS Total Score	30.6 (18.2)	— —	30.2 (18.2)	— —	31.8 (19.0)	— —
CC to Self-care maintenance score	— —	51.6 (19.9)	— —	54.8 (19.0)	— —	48.1 (19.7)
CC to Self-care management score [a]	— —	49.4 (19.9)	— —	53 (20.5)	— —	51 (20.7)
Caregiver self-efficacy score	— —	56.9 (22.8)	— —	60 (21.6)	— —	57.3 (23.3)
Caregiver Preparedness Scale (CPS)	— —	2.1 (0.8)	— —	2.2 (0.8)	— —	2.1 (0.8)
Mutuality Scale (MS)						
love and affection	3.3 (0.7)	3.2 (0.7)	3.4 (0.7)	3.4 (0.7)	3.2 (0.7)	3.4 (0.6)
shared pleasurable activities	2.8 (0.7)	2.6 (0.8)	2.9 (0.7)	2.7 (0.8)	2.7 (0.8)	2.7 (0.8)
shared values	2.6 (0.8)	2.7 (0.9)	2.8 (0.8)	2.8 (0.9)	2.7 (0.9)	2.7 (0.9)
reciprocity	2.9 (0.7)	2.6 (0.8)	3 (0.7)	2.7 (0.8)	2.8 (0.7)	2.6 (0.8)
total score	2.9 (0.6)	2.8 (0.7)	3 (0.6)	2.8 (0.7)	2.9 (0.6)	2.8 (0.6)

Note. MI, Motivational Interviewing; HF, heart failure; CCI, Charlson Comorbidity Index; MoCA, Montreal Cognitive Assessment; PSQI Pittsburgh Sleep Quality Index, SF-12 Short Form (12), KCCQ Kansas City Cardiomyopathy Questionnaire; HFSPS, Heart Failure Somatic Perception Scale; CC, Caregiver Contribution. All patient and caregiver characteristics were not statistically different among the three Arms as well as between the composite group (Arms 1 and 2) and the control group (Arm 3).

[a] Self-care management score can be computed only if patients have had HF symptoms in the last month (n = 354): symptomatic patients were n = 100 in Arm 1, n = 130 in Arm 2, and n = 124 in Arm 3; all corresponding percentages are referred to the number of symptomatic participants in each Arm.

Figure 5. Box-plot on Self-Care maintenance (a), management (b) and confidence (c) scale scores (randomized set with available data)



Note. The box represents the first and third quartile, the thick line the median, the empty dot the mean, the whiskers the minimum and maximum and the small dots outliers (defined as patients with self-care that is more than 1.5 times the interquartile range above the third quartile (or more than 1.5 times the interquartile range below the first quartile). Arm 1: Motivational Interviewing (MI) only for patients; Arm 2: MI for patients and caregivers; Arm 3: standard of care.

5.2 Change in scores during follow-up in patients

The improvement in the self-care maintenance scores (the primary endpoint) from baseline to 3 months following enrolment (T1) was greater in the intervention Arms 1 and 2 than in the control Arm 3 (Table 4). The mean improvement was 6.9, 7.4, and 2.5 points, in Arms 1, 2, and 3, respectively ($p = 0.0282$; Table 4, Figure 5) and we generally found a higher improvement in Arms 1 and 2 with respect to Arm 3 at the subsequent follow-ups [47]. The mean self-care management scores considerably increased at T1 by 12.3, 15.2 and 7.7 points, in Arms 1, 2, and 3, respectively ($p = 0.0284$; Table 4, Figure 5). The improvement in the three arms was significantly greater in Arms 1 and 2 than Arm 3 also at T2 (6 months after enrolment) but not at T3 and T4 (9 and 12 months after enrolment, respectively). Regarding the self-care confidence score, there were no significant differences between the three arms at T1 and T4; however, at T2 and T3, patients in Arms 1 and 2 improved their self-care confidence more than patients in Arm 3 ($p = 0.037$ and $p = 0.031$, respectively; Table 4, Figure 5).

The scores for anxiety, depression and sleep quality did not significantly change over time in any of the three research groups [64]. Although the generic QOL (SF-12) scores increased with time, no appreciable differences were seen between the three arms. However, there was different improvement in the disease-specific QOL (KCCQ overall summary score), with intervention Arms 1 and 2 having significantly higher scores than Arm 3 at T3 and T4 (differences of 6.7, 95% CI: 1.7; 11.7, and 8.4 points, 95% CI: 2.9; 13.8, respectively; Table 4), but not at T1 and T2. In terms of the HFSPS total score, Arms 1 and 2 compared to Arm 3 showed a substantial improvement at T3 (difference: -4.5, 95% CI: -9.1; -0.1) [65]. Regarding MS, there were no significant differences between the three arms during follow-up, except for a decline in the mean MS reciprocity from baseline to T2 in Arms 1 and 2 compared to Arm 3 (difference: -0.2, 95% CI: -0.3; 0, $p = 0.0475$), but the trend changed through the follow-up.

Table 4. Changes in scores during follow-up in patients

Variable	***N***	**Arm 1: MI only for patients (n = 155)**	**Arm 2: MI for patients and caregivers (n = 177)**	**Arm 3: Standard of care (n = 178)**	**Difference (95% CI) [a]**	***P* value**
Δ in Self-care maintenance scores		**Mean (SD)**	**Mean (SD)**	**Mean (SD)**		
T1	363	6.9 (19.6)	7.4 (20.1)	2.5 (18.2)	4.6 (0.5;8.7)	0.0282
T2	293	9.6 (18.9)	10.1 (22.2)	4.6 (21.7)	5.2 (−0.1;10.5)	0.0558
T3	252	13.8 (16.5)	15.9 (16.9)	7.8 (20.9)	7.1 (1.8;12.4)	0.0083
T4	238	21.1 (16.7)	18.8 (20.7)	14.6 (18.8)	5.2 (0.1;10.4)	0.0480
Δ in Self-care management scores		**Mean (SD)**	**Mean (SD)**	**Mean (SD)**		
T1	152	12.3 (15.2)	15.2 (16.9)	7.7 (15.8)	6.2 (0.6;11.9)	0.0284
T2	110	13.5 (23.1)	18.8 (24.1)	4.8 (18.6)	11.6 (3.2;20.1)	0.0076
T3	90	24.1 (23.1)	21.9 (22.1)	14.4 (16.5)	8.4 (0.0;16.9)	0.0503
T4	89	18.1 (19.5)	26.7 (23.1)	15 (18.7)	7.5 (−1.5;16.6)	0.1009
Δ Self-care confidence scores		**Mean (SD)**	**Mean (SD)**	**Mean (SD)**		
T1	363	6.4 (21.5)	5.8 (24.1)	3.4 (19.9)	2.7 (−1.9;7.3)	0.2495
T2	292	7.3 (23.5)	6.6 (28.6)	0.1 (25.9)	6.9 (0.4;13.4)	0.0374
T3	251	17.4 (24.0)	16.0 (22.9)	9.8 (22.9)	6.8 (0.6;13)	0.031
T4	237	17.6 (20.7)	15.4 (27.2)	12.5 (27.5)	3.9 (−3.3;11.2)	0.2865
Δ in Hospital Anxiety Scale (HAS)		**Mean (SD)**	**Mean (SD)**	**Mean (SD)**		
T1	364	0.1 (4.9)	-0.1 (3.9)	-0.1 (4.1)	0.1 (-0.8;1.0)	0.8504
T2	292	-0.7 (4.2)	-1.6 (4.4)	-1.4 (4.1)	0.1 (-0.9;1.2)	0.7580
T3	252	0.7 (4.2)	0.1 (4.3)	0.3 (4.4)	0.1 (-1.1;1.2)	0.9075
T4	238	-0.4 (4.8)	-1.7 (4.3)	-0.4 (4.5)	-0.7 (-1.9;0.5)	0.2594
Δ in Hospital Depression Scale (HDS)		**Mean (SD)**	**Mean (SD)**	**Mean (SD)**		
T1	364	0.1 (4.7)	0.2 (4)	-0.5 (4.4)	0.7 (-0.2;1.6)	0.1485
T2	292	-0.5 (3.5)	-0.9 (3.8)	-0.7 (4.2)	0 (-0.9;0.9)	0.9511
T3	252	0.9 (3.6)	0.8 (4)	0.4 (4.1)	0.4 (-0.6;1.4)	0.4318
T4	238	-0.6 (3.7)	-0.8 (4.2)	-0.7 (4.4)	0 (-1.1;1.1)	0.9850
Δ in Global PSQI Score		**Mean (SD)**	**Mean (SD)**	**Mean (SD)**		
T1	363	-0.5 (2.9)	-0.3 (3.2)	-0.7 (3.5)	0.3 (-0.3;1)	0.3566
T2	291	-0.4 (2.9)	-0.4 (3.3)	-0.4 (3.6)	0 (-0.7;0.8)	0.9283
T3	249	-0.1 (2.8)	-0.9 (3.2)	-0.3 (3.1)	-0.2 (-1;0.6)	0.6115
T4	234	-0.5 (2.6)	-0.7 (3.4)	-0.1 (3.2)	-0.5 (-1.4;0.2)	0.1828
SF-12 Physical Component Summary (PCS)		**Mean (SD)**	**Mean (SD)**	**Mean (SD)**		
T1	364	0.6 (10)	2.1 (10)	1.1 (8.1)	0.3 (-1.5;2.2)	0.7238
T2	292	2.9 (8)	2.6 (9.2)	1.1 (7.9)	1.6 (-0.4;3.8)	0.1167
T3	252	1.3 (7.4)	2.8 (9.5)	0 (8.3)	2.1 (-0.1;4.4)	0.0634
T4	238	1.3 (8.4)	3.2 (10.7)	0.7 (8.6)	1.6 (-0.9;4.2)	0.2189
SF-12 Mental Component Summary (MCS)		**Mean (SD)**	**Mean (SD)**	**Mean (SD)**		
T1	364	1.3 (11.6)	1.2 (10.1)	1.5 (10.7)	-0.2 (-2.6;2)	0.8116
T2	292	0.3 (11.9)	-0.4 (11.7)	-0.7 (10.6)	0.6 (-2.2;3.4)	0.6671
T3	252	3.0 (10.2)	4.4 (10.9)	2.2 (9.3)	1.4 (-1.2;4.2)	0.2840
T4	238	3.13 (10.46)	3.35 (11.69)	2.61 (10.54)	0.6 (-2.3;3.6)	0.6754

Variable	***N***	**Arm 1: MI only for patients (n = 155)**	**Arm 2: MI for patients and caregivers (n = 177)**	**Arm 3: Standard of care (n = 178)**	**Difference (95% CI) [a]**	***P* value**
Δ in KCCQ Overall Summary Score		**Mean (SD)**	**Mean (SD)**	**Mean (SD)**		
T1	364	1.3 (21.2)	0.7 (22.9)	0.8 (21.5)	0.1 (-4.6;4.9)	0.9613
T2	293	7.5 (17.7)	7.6 (21.0)	3.0 (16.5)	4.5 (-0.1;9.2)	0.0528
T3	252	11.0 (17.9)	13.2 (20.1)	5.5 (17.4)	6.7 (1.7;11.7)	0.0082
T4	238	11.4 (18.6)	13.4 (22.3)	4.1 (17.9)	8.4 (2.9;13.8)	0.0025
Δ in HFSPS Total Score		**Mean (SD)**	**Mean (SD)**	**Mean (SD)**		
T1	364	-1.5 (14.8)	-4.0 (19.5)	-3.3 (19.5)	0.4 (-3.5;4.4)	0.8295
T2	293	-2.7 (16.2)	-7.0 (20.9)	-5.1 (15.5)	0 (-4.1;4.1)	0.9891
T3	252	-5.0 (16.1)	-11.6 (17.7)	-4.1 (15.9)	-4.5 (-9.1;-0.1)	0.0475
T4	238	-4.9 (15.6)	-11.1 (18.3)	-4.3 (16.5)	-4 (-8.7;0.6)	0.0912
Δ in MS Love and affection		**Mean (SD)**	**Mean (SD)**	**Mean (SD)**		
T1	364	-0.3 (0.9)	-0.3 (0.9)	-0.2 (1.0)	-0.1 (-0.3;0.1)	0.2869
T2	291	0.0 (0.7)	0.0 (0.8)	0.0 (0.6)	0.0 (-0.2;0.1)	0.7378
T3	252	0.1 (0.6)	0.0 (0.7)	0.0 (0.5)	0.0 (-0.1;0.1)	0.9692
T4	238	0.1 (0.6)	0.0 (0.6)	-0.1 (0.7)	0.1 (0.0;0.3)	0.0950
Δ in MS Shared pleasurable activities		**Mean (SD)**	**Mean (SD)**	**Mean (SD)**		
T1	364	-0.2 (0.9)	-0.2 (0.8)	-0.1 (0.9)	-0.1 (-0.3;0.1)	0.1425
T2	291	0.1 (0.6)	0.0 (0.7)	0.2 (0.5)	-0.1 (-0.3;0.0)	0.0508
T3	252	0.2 (0.6)	0.2 (0.7)	0.2 (0.5)	0.0 (-0.1;0.1)	0.8770
T4	238	0.2 (0.5)	0.2 (0.7)	0.2 (0.7)	0.1 (-0.1;0.3)	0.4767
Δ in MS Shared values		**Mean (SD)**	**Mean (SD)**	**Mean (SD)**		
T1	364	-0.2 (0.9)	-0.3 (0.9)	-0.2 (1.0)	0.0 (-0.2;0.2)	0.9468
T2	291	0.1 (0.7)	0.0 (0.7)	0.2 (0.8)	-0.1 (-0.3;0.1)	0.1641
T3	252	0.2 (0.6)	0.2 (0.7)	0.2 (0.8)	0.0 (-0.2;0.2)	0.9710
T4	238	0.2 (0.7)	0.1 (0.7)	0.1 (0.9)	0.1 (-0.1;0.3)	0.5213
Δ in MS Reciprocity		**Mean (SD)**	**Mean (SD)**	**Mean (SD)**		
T1	364	-0.2 (0.9)	-0.2 (0.8)	-0.1 (1.0)	-0.1 (-0.3;0.1)	0.1891
T2	291	0.2 (0.6)	0.0 (0.7)	0.2 (0.5)	-0.2 (-0.3;0.0)	0.0475
T3	252	0.2 (0.6)	0.2 (0.6)	0.2 (0.5)	0.0 (-0.1;0.1)	0.8942
T4	237	0.2 (0.6)	0.2 (0.6)	0.1 (0.7)	0.1 (-0.1;0.3)	0.4264
Δ in MS Total score		**Mean (SD)**	**Mean (SD)**	**Mean (SD)**		
T1	364	-0.2 (0.8)	-0.2 (0.7)	-0.1 (0.9)	-0.1 (-0.3;0.1)	0.2197
T2	291	0.1 (0.5)	0.0 (0.6)	0.2 (0.4)	-0.1 (-0.3;0.0)	0.0760
T3	252	0.2 (0.5)	0.1 (0.5)	0.2 (0.4)	0.0 (-0.1;0.1)	0.9890
T4	237	0.2 (0.4)	0.2 (0.5)	0.1 (0.6)	0.1 (-0.1;0.2)	0.2747

Note. T1, T2, T3 and T4 correspond to 3, 6, 9, and 12 months from enrolment.
CI, confidence interval; MI, motivational interviewing; SD, standard deviation; PSQI, Pittsburgh Sleep Quality Index; KCCQ: Kansas City Cardiomyopathy Questionnaire HFSPS: Heart Failure Somatic Perception Scale; MS, Mutuality scale.
Δ scores. The columns for each arm report the delta (Δ) of the scores computed subtracting the score at baseline from the score at each follow-up time (T1, T2, T3, and T4).
[a] The difference is between Arms 1 and Arm 2 vs. Arm 3.

5.3 Change in scores during follow-up in caregivers

CC to self-care management and CC to self-care maintenance both considerably improved over the course of the 12-month follow-up, but there were no significant differences across the three arms (Table 5) [66]. Caregiver self-efficacy in the three arms also increased over time, however at the 9-month follow-up (T3), Arm 2 had a higher score than Arm 3 (difference: 8.3 points, 95% CI: 3.1; 13.5, p = 0.002). It was similarly higher in Arm 2 than in Arm 3 at the 12-month follow-up (T4), although this difference was only marginally statistically significant (difference: 6.5, 95% CI: 0.4; 13.6, p = 0.064). Over time, there was no significant differences in the Caregiver Preparedness Scale (CPS), anxiety, depression, physical and mental caregiver quality of life (QOL) scores among the three arms, although they slightly improved over the 12 months of the study. Over the course of the follow-up, caregiver sleep disturbances decreased in each of the three arms, but Arms 1 and 3 saw the greatest reductions. Caregivers' mean sleep disturbances decreased more in Arm 2 than in Arm 3 from baseline to T1 (difference: -0.9, 95% CI: -1.6; -0.2, p = 0.008), but the difference faded with time (Table 5). Among caregivers, the mean Mutuality Scale (MS) total score increased more in Arm 2 than in Arm 3 from baseline to T3 (9 months; difference: 0.2, 95% CI: 0; 0.3, p = 0.0314); however, the difference between the two arms shrank at T4 (12 months).

Table 5. Changes in scores during follow-up in caregivers

Variable	*N*	Arm 1: MI only for patients (n = 155)	Arm 2: MI for patients and caregivers (n = 177)	Arm 3: Standard of care (n = 178)	Arm 2 vs Arm 1		Arm 2 vs Arm 3	
					Difference (95%CI)	***P* value**	**Difference (95%CI)**	***P* value**
Δ in CC to self-care Maintenance		**Mean (SD)**	**Mean (SD)**	**Mean (SD)**				
T1	319	2.6 (19.1)	2.5 (18.1)	4.9 (21.1)	-0.1 (-5.1; 4.8)	0.9545	-2.3 (-7.5; 2.7)	0.3616
T2	278	1.6 (18.2)	0.9 (23.4)	6.3 (22.4)	-0.7 (-6.7; 5.2)	0.8161	-5.3 (-11.9;1.1)	0.1080
T3	246	5.9 (21.4)	7.1 (21.8)	5.4 (20.5)	1.1 (-5.4; 7.7)	0.7260	1.6 (-4.7; 8.1)	0.6077
T4	229	6.4 (21.1)	9.3 (22.4)	7.1 (20.6)	2.8 (-4.1; 9.8)	0.4165	2.1 (-4.6; 8.9)	0.5270
Δ in CC to self-care Management		**Mean (SD)**	**Mean (SD)**	**Mean (SD)**				
T1	144	5.5 (10.8)	7.6 (14.9)	3.2 (15)	2.0 (-3.1; 7.2)	0.4278	4.4 (-1.5;10.3)	0.1459
T2	95	8.2 (13.0)	9.8 (16.1)	7.1 (16.6)	1.6 (-5.6; 8.8)	0.6551	2.7 (-5.4;11.0)	0.4993
T3	86	9.2 (16.9)	7.5 (19.3)	5.7 (15.8)	-1.7 (-11.3; 7.8)	0.7108	1.7 (-7.5;11.1)	0.7032
T4	84	10.6 (14.3)	11.2 (18.7)	7.5 (16.7)	0.5 (-8.8; 9.9)	0.9015	3.6 (-5.5;12.8)	0.4320
Δ in Caregiver Self-efficacy		**Mean (SD)**	**Mean (SD)**	**Mean (SD)**				
T1	318	4.5 (18.1)	4.8 (19.3)	2.5 (17.8)	0.3 (-4.8; 5.4)	0.9049	2.2 (-2.6;7.2)	0.3604
T2	277	3.5 (16.6)	5.4 (21.4)	2.9 (17.1)	1.9 (-3.5; 7.4)	0.4920	2.5 (-2.9;8.1)	0.3629
T3	245	7.4 (17.8)	10.4 (19.5)	2.1 (14.7)	2.9 (-2.7; 8.7)	0.3065	8.3 (3.1;13.5)	0.0010
T4	229	8.6 (17.5)	12.1 (25.1)	5.5 (19.7)	3.5 (-3.2; 10.3)	0.3052	6.5 (-0.4;13.6)	0.0641
Δ in Caregiver Preparedness Scale (CPS)		**Mean (SD)**	**Mean (SD)**	**Mean (SD)**				
T1	319	0.1 (0.5)	0.1 (0.5)	0.1 (0.6)	0 (-0.1; 0.1)	0.9277	-0.1 (-0.2;0.1)	0.4375
T2	278	0.1 (0.6)	0.2 (0.6)	0.2 (0.6)	0.1 (-0.1; 0.2)	0.3910	0 (-0.2;0.1)	0.6529
T3	248	0.2 (0.6)	0.3 (0.7)	0.2 (0.6)	0.1 (-0.1; 0.3)	0.2817	0.1 (-0.1;0.3)	0.2111
T4	232	0.2 (0.7)	0.3 (0.7)	0.3 (0.8)	0.1 (-0.1; 0.3)	0.2338	0.1 (-0.1;0.3)	0.6772
Δ in Hospital Anxiety Scale (HAS)		**Mean (SD)**	**Mean (SD)**	**Mean (SD)**				
T1	319	-0.6 (3.0)	-0.7 (3.1)	-0.4 (3.6)	-0.1 (-0.9;0.7)	0.8370	-0.3 (-1.1;0.5)	0.4877
T2	280	-1.9 (4.6)	-1.5 (4.8)	-2.2 (4.8)	0.4 (-0.9;1.7)	0.5533	0.6 (-0.6;2)	0.3352
T3	248	-2.3 (4.8)	-1.9 (5.3)	-2.3 (5.1)	0.3 (-1.2;1.8)	0.6597	0.3 (-1.2;1.9)	0.6683
T4	232	-1.7 (4.8)	-1.8 (5.3)	-2.2 (5.1)	-0.1 (-1.7;1.4)	0.8542	0.3 (-1.2;1.9)	0.6523
Δ in Hospital Depression Scale (HDS)		**Mean (SD)**	**Mean (SD)**	**Mean (SD)**				
T1	319	-0.5 (3.0)	-0.4 (3.0)	-0.1 (3.6)	0.1 (-0.6;0.9)	0.6672	-0.2 (-1.1;0.6)	0.6381
T2	280	-1.8 (4.9)	-1.1 (4.7)	-1.6 (5.0)	0.6 (-0.7;2)	0.3628	0.4 (-0.9;1.8)	0.5457
T3	248	-1.8 (5.0)	-1.2 (5.3)	-1.7 (5.6)	0.5 (-0.9;2.1)	0.4736	0.4 (-1.1;2.1)	0.5580
T4	232	-1.5 (5.1)	-1.4 (5.1)	-1.7 (5.4)	0.1 (-1.5;1.6)	0.9446	0.2 (-1.3;1.9)	0.7309
SF-12 Physical Component Summary (PCS)		**Mean (SD)**	**Mean (SD)**	**Mean (SD)**				
T1	319	2.5 (6.9)	1.4 (6.7)	0.6 (7.5)	-1.1 (-2.9;0.7)	0.2492	0.7 (-1.1;2.6)	0.4234
T2	281	2.4 (7.1)	2.3 (8.0)	1.6 (7.3)	0 (-2.2;2.1)	0.9748	0.7 (-1.4;2.9)	0.5022
T3	247	2.7 (6.4)	2.0 (7.5)	1.3 (7.4)	-0.6 (-2.8;1.4)	0.5444	0.7 (-1.5;3)	0.5195
T4	231	1.7 (7.7)	2.6 (8.5)	1.7 (6.3)	0.8 (-1.7;3.4)	0.5035	0.9 (-1.3;3.2)	0.4233
SF-12 Mental Component Summary (MCS)		**Mean (SD)**	**Mean (SD)**	**Mean (SD)**				
T1	319	1.5 (7.2)	0.6 (7.7)	0.2 (9.1)	-0.8 (-2.8;1.2)	0.4156	0.4 (-1.7;2.6)	0.6955
T2	281	2.6 (9.4)	0.7 (7.8)	2.2 (8.2)	-1.8 (-4.3;0.5)	0.1325	-1.5 (-3.7;0.7)	0.1897

T3	247	3.6 (7.9)	1.7 (10.4)	2.7 (7.5)	-1.8 (-4.6;0.9)	0.1834	-1 (-3.7;1.6)	0.4662
T4	231	2.8 (8.0)	2.4 (10.0)	2.0 (8.4)	-0.3 (-3.2;2.4)	0.7929	0.4 (-2.4;3.3)	0.7750
Δ in Global PSQI Score		**Mean (SD)**	**Mean (SD)**	**Mean (SD)**				
T1	312	-0.3 (2.8)	-0.5 (2.6)	0.4 (2.6)	-0.1 (-0.9;0.5)	0.6354	-0.9 (-1.6;-0.2)	0.0086
T2	271	-0.7 (2.7)	-0.4 (3.0)	-0.6 (2.5)	0.3 (-0.5;1.1)	0.4366	0.2 (-0.5;1)	0.5977
T3	245	-0.7 (2.8)	-0.4 (2.7)	-0.6 (2.7)	0.2 (-0.5;1.1)	0.5625	0.2 (-0.6;1)	0.6402
T4	232	-0.8 (2.3)	-0.1 (2.8)	-0.7 (2.1)	0.6 (-0.1;1.5)	0.0957	0.6 (-0.1;1.3)	0.1108
Δ in MS Love and affection		**Mean (SD)**	**Mean (SD)**	**Mean (SD)**				
T1	320	0.0 (0.7)	0.0 (0.6)	-0.1 (0.5)	0 (-0.2;0.2)	0.8084	0.1 (-0.1;0.2)	0.3113
T2	280	0.2 (0.7)	0.1 (0.6)	0.1 (0.6)	-0.1 (-0.2;0.2)	0.6438	0 (-0.2;0.2)	0.7611
T3	248	0.3 (0.7)	0.2 (0.6)	0.1 (0.6)	-0.1 (-0.3;0.1)	0.4800	0.1 (0;0.3)	0.1202
T4	232	0.3 (0.7)	0.1 (0.6)	0.1 (0.6)	-0.2 (-0.4;0.1)	0.1593	0 (-0.2;0.2)	0.7910
Δ in MS Shared pleasurable activities		**Mean (SD)**	**Mean (SD)**	**Mean (SD)**				
T1	320	0.1 (0.6)	0.1 (0.7)	0.1 (0.5)	0 (-0.2;0.2)	0.9553	0 (-0.2;0.2)	0.9690
T2	280	0.2 (0.7)	0.1 (0.7)	0.2 (0.7)	-0.1 (-0.3;0.1)	0.2386	-0.1 (-0.3;0.1)	0.3579
T3	248	0.3 (0.8)	0.3 (0.6)	0.1 (0.6)	0 (-0.3;0.2)	0.7275	0.2 (0;0.3)	0.1084
T4	232	0.3 (0.7)	0.3 (0.6)	0.2 (0.7)	0 (-0.2;0.2)	0.8417	0.1 (-0.1;0.4)	0.1951
Δ in MS Shared values		**Mean (SD)**	**Mean (SD)**	**Mean (SD)**				
T1	320	0.0 (0.7)	0.1 (0.7)	0.0 (0.7)	0.1 (0.0;0.3)	0.0989	0.1 (-0.1;0.3)	0.3340
T2	280	0.3 (0.8)	0.1 (0.7)	0.2 (0.8)	-0.2 (-0.4;0.1)	0.2316	-0.1 (-0.2;-0.2)	0.9217
T3	248	0.3 (1.0)	0.2 (0.7)	0.1 (0.6)	-0.1 (-0.3;0.2)	0.6809	0.1 (-0.1;0.4)	0.2278
T4	232	0.2 (0.8)	0.2 (0.7)	0.2 (0.6)	0 (-0.3;0.2)	0.8392	0 (-0.2;0.3)	0.6616
Δ in MS Reciprocity		**Mean (SD)**	**Mean (SD)**	**Mean (SD)**				
T1	320	0.1 (0.6)	0.1 (0.6)	0.1 (0.6)	0 (-0.1;0.2)	0.7588	0 (-0.2;0.2)	0.9912
T2	280	0.3 (0.8)	0.2 (0.7)	0.2 (0.7)	-0.1 (-0.3;0.1)	0.2340	0 (-0.3;0.1)	0.5587
T3	248	0.4 (0.7)	0.4 (0.7)	0.2 (0.6)	0 (-0.2;0.2)	0.8349	0.2 (0;0.4)	0.0666
T4	232	0.4 (0.8)	0.4 (0.7)	0.3 (0.7)	0 (-0.2;0.2)	0.9110	0.1 (-0.1;0.3)	0.2907
Δ in MS Total score		**Mean (SD)**	**Mean (SD)**	**Mean (SD)**				
T1	320	0.1 (0.5)	0.1 (0.5)	0.1 (0.4)	0 (-0.1;0.2)	0.6972	0 (-0.1;0.2)	0.6950
T2	280	0.3 (0.6)	0.1 (0.6)	0.2 (0.5)	-0.2 (-0.3;0.1)	0.2092	-0.1 (-0.2;0.1)	0.5852
T3	248	0.3 (0.6)	0.3 (0.5)	0.1 (0.5)	0 (-0.2;0.1)	0.7957	0.2 (0;0.3)	0.0314
T4	232	0.3 (0.7)	0.3 (0.5)	0.2 (0.5)	0 (-0.2;0.1)	0.6365	0.1 (-0.1;0.3)	0.2555

Note. T1, T2, T3 and T4 correspond to 3, 6, 9, and 12 months from enrolment.

CC, caregiver contribution; CI, confidence interval; MI, Motivational Interviewing; SD, standard deviation; CC: Caregiver Contribution; PSQI, Pittsburgh Sleep Quality Index; MS, Mutuality scale.

Δ scores. The columns for each arm report the delta (Δ) of the scores computed subtracting the score at baseline from the score at each follow-up time (T1, T2, T3, and T4).

5.4 Health-care service use among patients

The usage of health-care services among patients in each follow-up visit is detailed in Table 6. Throughout the observation period, 25 (16.1%) patients in Arm 1, 30 (17%) patients in Arm 2, and 20 (11.2%) patients in Arm 3 utilized medical services (hospitalizations and use of emergency services) at least once [67]. With no discernible trend over time and no statistically significant difference between the three arms, health-care service utilization ranged from 7.5% to 16.7%. Table 7 presents the results of the mixed model, in which the interaction between arm and visit number reported p = 0.836, while Figure 6 displays model-based estimations of patient use of health-care services.

Table 6. Health-care service use among patients with HF at each follow-up.

	Arm 1			**Arm 2**			**Arm 3**			
Follow-up	**N patients with available information**	**N health-care services**	**%**	**N patients with available information**	**N health-care services**	**%**	**N patients with available information**	**N health-care services**	**%**	**Fisher's exact test**
1	86	9	10.5	103	17	16.5	90	15	16.7	0.4097
2	80	6	7.5	91	11	12.1	77	7	9.1	0.6136
3	73	8	11.1	83	7	8.4	68	8	11.8	0.7614
4	62	10	16.1	73	11	15.1	62	10	16.1	1

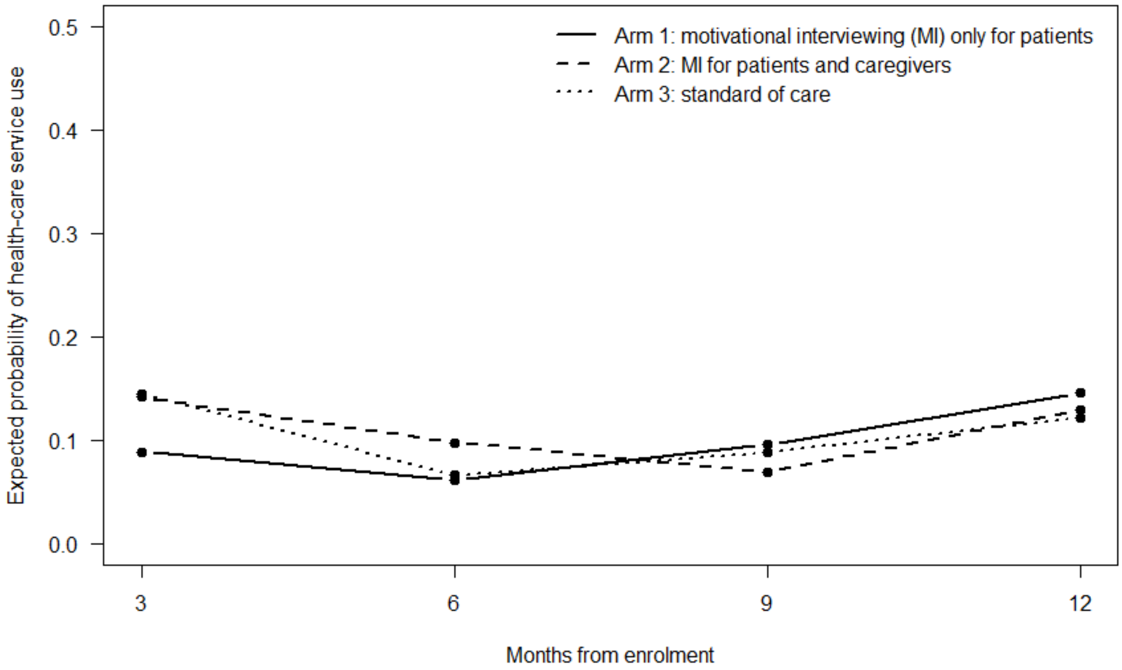
Note. Follow-up numbers 1, 2, 3 and 4 correspond to 3, 6, 9 and 12 months from enrolment, respectively. Arm 1: motivational interviewing (MI) only for patients; Arm 2: MI for patients and caregivers; and Arm 3: standard of care.

Table 7. Longitudinal generalized linear mixed model on health-care service use.

Effect	OR (95% CI)
Visit number: 2 vs 1	0.41 (0.14; 1.18)
Visit number: 3 vs 1	0.57 (0.20; 1.57)
Visit number: 4 vs 1	0.81 (0.30; 2.16)
Arm: 1 vs 3 at visit 1	0.58 (0.21; 1.58)
Arm: 2 vs 3 at visit 1	0.97 (0.40; 2.37)
Arm: 1 vs 3 at visit 2	0.92 (0.26; 3.29)
Arm: 2 vs 3 at visit 2	1.54 (0.49; 4.83)
Arm: 1 vs 3 at visit 3	1.09 (0.33; 3.59)
Arm: 2 vs 3 at visit 3	0.77 (0.23; 2.60)
Arm: 1 vs 3 at visit 4	1.22 (0.40; 3.73)
Arm: 2 vs 3 at visit 4	1.07 (0.36; 3.21)

Note. OR, odds ratio; CI, confidence interval. Arm 3 and visit number 1 considered as reference levels. Arm 1: motivational interviewing (MI) only for patients; Arm 2: MI for patients and caregivers; Arm 3: standard of care. Visits number 1, 2, 3 and 4 correspond to 3, 6, 9 and 12 months from enrolment.

Figure 6. Expected probabilities of health-care service use among patients with HF by treatment arm.



5.5 Mortality among patients

Throughout the course of twelve months of the study, 28 patients died [67]. Three (1.9%), one (0.6%), and nine (5.1%) patients had died in Arms 1, 2, and 3, respectively, at T1 (3 months after enrolment; Fisher test p = 0.026). Figure 7 displays the survival curve for the year of follow-up. Although the survival estimates in the control arm (Arm 3) were lower than those in the intervention arms (Arms 1 and 2), the log-rank test that took into account the entire follow-up showed no statistically significant differences among the three study arms (p = 0.2886). We split time at 3 months in a time-dependent Cox model since the hazard proportionality among the three arms was not respected (global Schoenfeld test p = 0.042). According to the Cox model, mortality was significantly lower in Arm 2 than Arm 3 during the first three months (HR = 0.11, 95% CI: 0.01; 0.88, p = 0.038), but there was no difference over the subsequent months (p = 0.690). In the first three months, there was also a potential for lower mortality in Arm 1 compared to Arm 3; however, this finding failed to reach statistical significance (HR = 0.38, 95% CI: 0.10; 1.41, p = 0.150, Table 8).

Figure 7. Life-table survival estimate of patients with HF in the three arms.

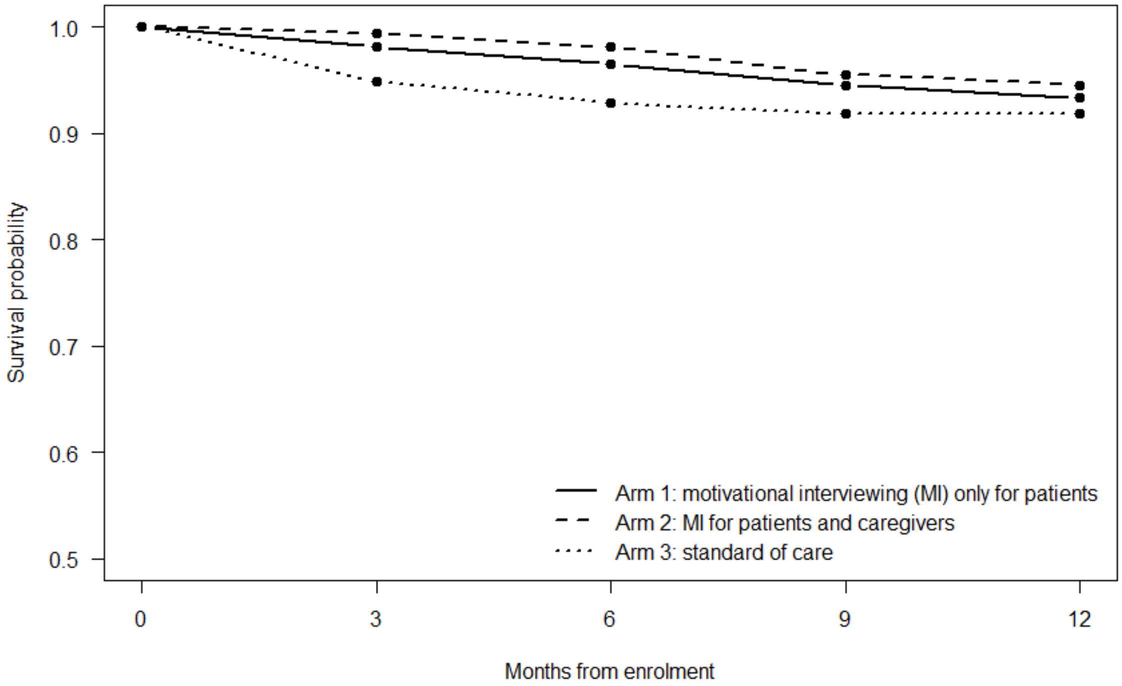


Table 8: Hazard ratios of all-cause mortality within (T0–T1) and over (T1–T4) 3 months after enrolment.

Time interval	Arm	HR (95% CI)	*P* value
0–3 months	Arm 1 vs Arm 3	0.38 (0.10–1.41)	0.1498
0–3 months	Arm 2 vs Arm 3	0.11 (0.01–0.88)	0.0376
3-12 months	Arm 1 vs Arm 3	1.26 (0.34–4.72)	0.7237
3-12 months	Arm 2 vs Arm 3	1.29 (0.36–4.58)	0.6896

Note. HR, hazard ratio; CI, confidence interval; Arm 1: motivational interviewing (MI) only for patients; Arm 2: MI for patients and caregivers; Arm 3: standard of care.

5.6 Change in scores over time in patients using linear mixed models

Figure 8 displays model-based trends in the primary and secondary end-point scores across time (from baseline to T4) measured on patients. Regarding self-care maintenance, management, and confidence scale scores, all of the arms saw improvements in their scores, with Arms 1 and 2 generally having a larger slope than Arm 3 (Table 9; Figure 8, panels a-c). In particular, Arm 1 improved in self-care maintenance over the course of the observation year significantly more than Arm 3 (β = 1.22, 95% CI: 0.02; 2.41, p = 0.046), while the difference between Arms 2 and 3 improved just barely enough to be statistically significant compared to Arm 3 (β = 1.12, 95% CI: -0.03; 2.26, p = 0.056). Arm 1 did not significantly differ from Arm 3 in terms of self-care management (p = 0.092), however Arm 2 considerably improved more than Arm 3 (β = 2.89, 95%CI: 0.9; 4.88, p = 0.005). In terms of self-care confidence, Arms 1 and 2 did not improved differently from Arm 3 (p = 0.172 and p = 0.289, respectively; Table 9). Even after accounting for additional factors (patient age, sex, income, cohabitation with caregiver, NYHA class, Charlson Comorbidity Index (CCI), MoCA score, time since diagnosis, number of medications, and baseline self-care confidence scale score), the results of these models remained consistent [47].

Figure 8, panels d-e, depicts model-based trends of HADS from baseline to T4. Neither anxiety nor depression improved more in either intervention Arms 1 and 2 than in Arm 3 during the course of the observational year (p = 0.354 and p = 0.321 in the HAS; p = 0.617 and p = 0.902 in the HDS, respectively). Additionally, these scores did not change over time (β = -0.15, 95% CI: -0.35; 0.05, p = 0.132 for HAS, and β = -0.11, 95% CI: -0.29; 0.07, p = 0.243 for HDS; Table 9). Same thing for sleep quality (PSQI), where Arms 1 and 2 did not improve significantly more than Arm 3 in model-based trends (p = 0.091 and p = 0.074, respectively), and there was no improvement over time (β = 0.06, 95% CI: -0.1; 0.23, p = 0.446) [64].

According to *P* values of 0.632 and 0.227 for the SF-12 Physical Component Summary (PCS) and 0.861 and 0.525 for the SF-12 Mental Component Summary (MCS), displayed in Figure 8, panels f-g, respectively, neither Arms 1 nor 2 significantly improved over the course of the observation year more than Arm 3 did. The longitudinal model, however, in terms of the disease-specific QOL (KCCQ overall summary score) over time showed that Arm 2 had significantly larger improvement than Arm 3 (β = 1.57, 95% CI: 0.26; 2.89, p =

0.019; Table 9, Figure 8h). Regarding the HFSPS total score, it significantly improved over time (β = -1.52, 95% CI: -2.36; -0.68, p = 0.0004) when examined in the longitudinal model (Figure 8i), with Arm 2 exhibiting a greater improvement than Arm 3 (β = -1.35, 95% CI: -2.50; -0.21, p = 0.020; Table 9) [64].

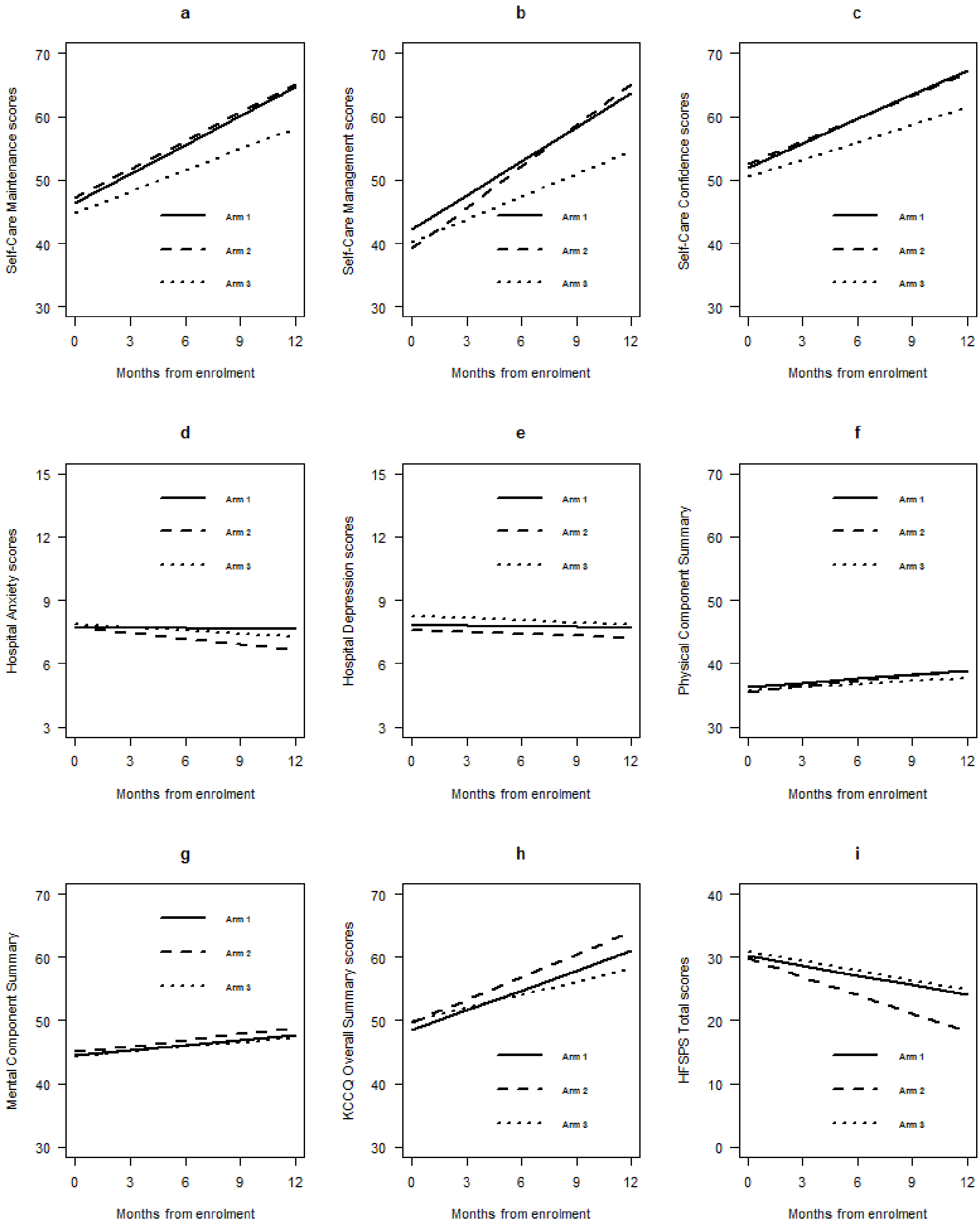
Table 9. Longitudinal linear mixed model results on primary and secondary outcome scores in patients

Effect	β	95% C.I.	*P* value
Self-care maintenance scores			
Time (for each visit)	3.38	(2.54; 4.21)	<.0001
MI only for patients vs Standard of care	1.60	(-1.98; 5.17)	0.3815
MI for patients and caregivers vs Standard of care	2.46	(-0.99; 5.92)	0.1621
Time * MI only for patients	1.22	(0.02; 2.41)	0.0460
Time * MI for patients and caregivers	1.12	(-0.03; 2.26)	0.0553
Self-care management scores			
Time (for each visit)	3.60	(2.15; 5.05)	<.0001
MI only for patients vs Standard of care	2.01	(-2.73; 6.76)	0.4053
MI for patients and caregivers vs Standard of care	-0.99	(-5.43; 3.45)	0.6604
Time * MI only for patients	1.78	(-0.28; 3.84)	0.0916
Time * MI for patients and caregivers	2.89	(0.90; 4.88)	0.0045
Self-care confidence scores			
Time (for each visit)	2.76	(1.70; 3.82)	<.0001
MI only for patients vs Standard of care	1.56	(-2.53; 5.65)	0.4554
MI for patients and caregivers vs Standard of care	2.17	(-1.78; 6.12)	0.2816
Time * MI only for patients	1.06	(-0.46; 2.57)	0.1720
Time * MI for patients and caregivers	0.79	(-0.67; 2.24)	0.2886
Hospital Anxiety Scale (HAS)			
Time (for each visit)	-0.15	(-0.35; 0.04)	0.1317
MI only for patients vs Standard of care	-0.14	(-0.93; 0.65)	0.7285
MI for patients and caregivers vs Standard of care	-0.13	(-0.90; 0.63)	0.7319
Time * MI only for patients	0.13	(-0.15; 0.42)	0.3543
Time * MI for patients and caregivers	-0.13	(-0.41; 0.13)	0.3210
Hospital Depression Scale (HDS)			
Time (for each visit)	-0.10	(-0.29; 0.07)	0.2433
MI only for patients vs Standard of care	-0.40	(-1.19; 0.39)	0.3181
MI for patients and caregivers vs Standard of care	-0.66	(-1.43; 0.10)	0.0885
Time * MI only for patients	0.06	(-0.19; 0.32)	0.6167
Time * MI for patients and caregivers	0.01	(-0.23; 0.26)	0.9017
Global PSQI Score			
Time (for each visit)	0.06	(-0.10; 0.23)	0.4462
MI only for patients vs Standard of care	0.31	(-0.43; 1.06)	0.4129
MI for patients and caregivers vs Standard of care	0.28	(-0.43; 1.01)	0.4401
Time * MI only for patients	-0.20	(-0.43; 0.03)	0.0911
Time * MI for patients and caregivers	-0.20	(-0.43; 0.01)	0.0737
SF-12 Physical Component Summary (PCS)			
Time (for each visit)	0.49	(0.03; 0.95)	0.0339
MI only for patients vs Standard of care	0.62	(-1.19; 2.45)	0.4998
MI for patients and caregivers vs Standard of care	-0.19	(-1.96; 1.56)	0.8278
Time * MI only for patients	0.15	(-0.49; 0.80)	0.6323
Time * MI for patients and caregivers	0.38	(-0.23; 1.00)	0.2265
SF-12 Mental Component Summary (MCS)			

Time (for each visit)	0.73	(0.19; 1.27)	0.0074
MI only for patients vs Standard of care	0.20	(-1.71; 2.13)	0.8311
MI for patients and caregivers vs Standard of care	0.62	(-1.23; 2.48)	0.5092
Time * MI only for patients	0.06	(-0.69; 0.83)	0.8607
Time * MI for patients and caregivers	0.23	(-0.49; 0.96)	0.5250
KCCQ Overall Summary Score			
Time (for each visit)	2.05	(1.084; 3.01)	<0.0001
MI only for patients vs Standard of care	-1.40	(-5.76; 2.95)	0.5262
MI for patients and caregivers vs Standard of care	-0.35	(-4.56; 3.85)	0.8692
Time * MI only for patients	1.05	(-0.31; 2.43)	0.1310
Time * MI for patients and caregivers	1.57	(0.25; 2.88)	0.0190
HFSPS Total Score			
Time (for each visit)	-1.52	(-2.36; -0.68)	0.0004
MI only for patients vs Standard of care	-0.74	(-4.24; 2.76)	0.6769
MI for patients and caregivers vs Standard of care	-1.15	(-4.54; 2.23)	0.5034
Time* MI only for patients	0.00	(-1.19; 1.19)	0.9965
Time* MI for patients and caregivers	-1.35	(-2.49; -0.21)	0.0201

Note. MI: Motivational Interviewing; CI: confidence interval; PSQI: Pittsburgh Sleep Quality Index; KCCQ: Kansas City Cardiomyopathy Questionnaire HFSPS: Heart Failure Somatic Perception Scale.

Figure 8. Model-based primary and secondary scores by follow-up time in patients.



Note. Self-care maintenance score (panel a), Self-care management score (panel b), Self-care confidence score (panel c), Hospital Anxiety Scale score (panel d), Hospital Depression Scale score (panel e), SF-12 Physical Component Summary score (panel f), SF-12 Mental Component Summary score (panel g), Kansas City Cardiomyopathy Questionnaire Overall Summary score (panel h), Heart Failure Somatic Perception Scale Total score (panel i). Arm 1: motivational interviewing (MI) only for patients; Arm 2: MI for patients and caregivers; and Arm 3: standard of care.

5.7 Change in scores over time in caregivers using linear mixed models

Figure 9 displays model-based trends of secondary outcome scores from baseline to T4 in caregivers. Imbalances in the variable "caregivers living with patients" at baseline were taken into account by the models [66]. In comparison to Arms 1 and 3, CC to self-care management, and maintenance scores in Arm 2 did not improve during the course of the observational year ($p = 0.515$ and $p = 0.694$ for CC to self-care management score, $p = 0.997$ and $p = 0.338$ for CC to self-care maintenance score, respectively; Table 10, Figure 9, panels a-b). However, Arm 2 significantly improved more than Arm 3 in terms of caregiver self-efficacy ($\beta = 1.39$, 95% CI: 0.02; 2.75, $p = 0.046$; Table 10, Figure 9c). Regarding Caregiver Preparedness Scale (CPS), Arm 2 once more did not significantly improve more than Arms 1 and 3 ($p = 0.112$ and $p = 0.227$, respectively; Table 10, Figure 9d).

Model-based trends of caregiver anxiety and depression levels from baseline to T4 are displayed in Figure 9 (panels e-f). These scores slightly decreased by about a half point for each visit during the course of the observation period (decrease of HAS in Arm 2: $\beta = -0.45$, 95% CI: -0.67; -0.24, $p < 0.0001$, and decrease of HDS in Arm 2: $\beta = -0.39$, 95% CI: -0.60; -0.18, $p = 0.0003$), although the decrement was not different among the three Arms (Table 10). In particular, the difference between caregivers participated in MI (Arm 2) and Arm 1 (MI exclusively for patient) was 0.04 (95% CI: -0.28; 0.36, $p = 0.816$) in HAS and 0.05 (95% CI: -0.27; 0.36, $p = 0.766$) in HDS. Caregivers living with the patient had higher HAS although the difference was not statistical significant ($\beta = 0.63$, 95% CI: -0.05; 1.30, $p = 0.069$), while they had a significant increase in HDS of nearly 1 point ($\beta = 0.81$, 95% CI: 0.17; 1.45, $p = 0.013$, Table 10) as compared with caregivers not living with the patient. In PSQI no differences over time were reported in Arm 2 compared to Arms 1 and 3 when the three Arms were analysed in the longitudinal model ($p = 0.252$ and $p = 0.389$, respectively; Figure 9, panel g). Living with patient was not significantly associated with changes in PSQI (Table 10).

Model-based trends of caregiver physical and mental QOL scores from baseline to T4 are also shown in Figure 9 (panels h and i). Over the year of observation, these scores increased of nearly half point for each visit (increase in Arm 2: β = 0.46, 95% CI: 0.12; 0.80, p = 0.008 in PCS and β = 0.64, 95% CI: 0.23; 1.04, p = 0.002 in MCS). The increment was not different in the three Arms. In fact, both scores did not change more in Arm 2 compared to Arms 1 and 3 (p = 0.992 and p = 0.368 in PCS, p = 0.552 and p = 0.543 in MCS, respectively). Living with patient condition decreased PCS (β = -2.75, 95% CI: -4.04; -1.46, p < 0.0001), but not MCS (β = -0.64, 95% CI: -2.06; 0.77, p = 0.373; Table 10).

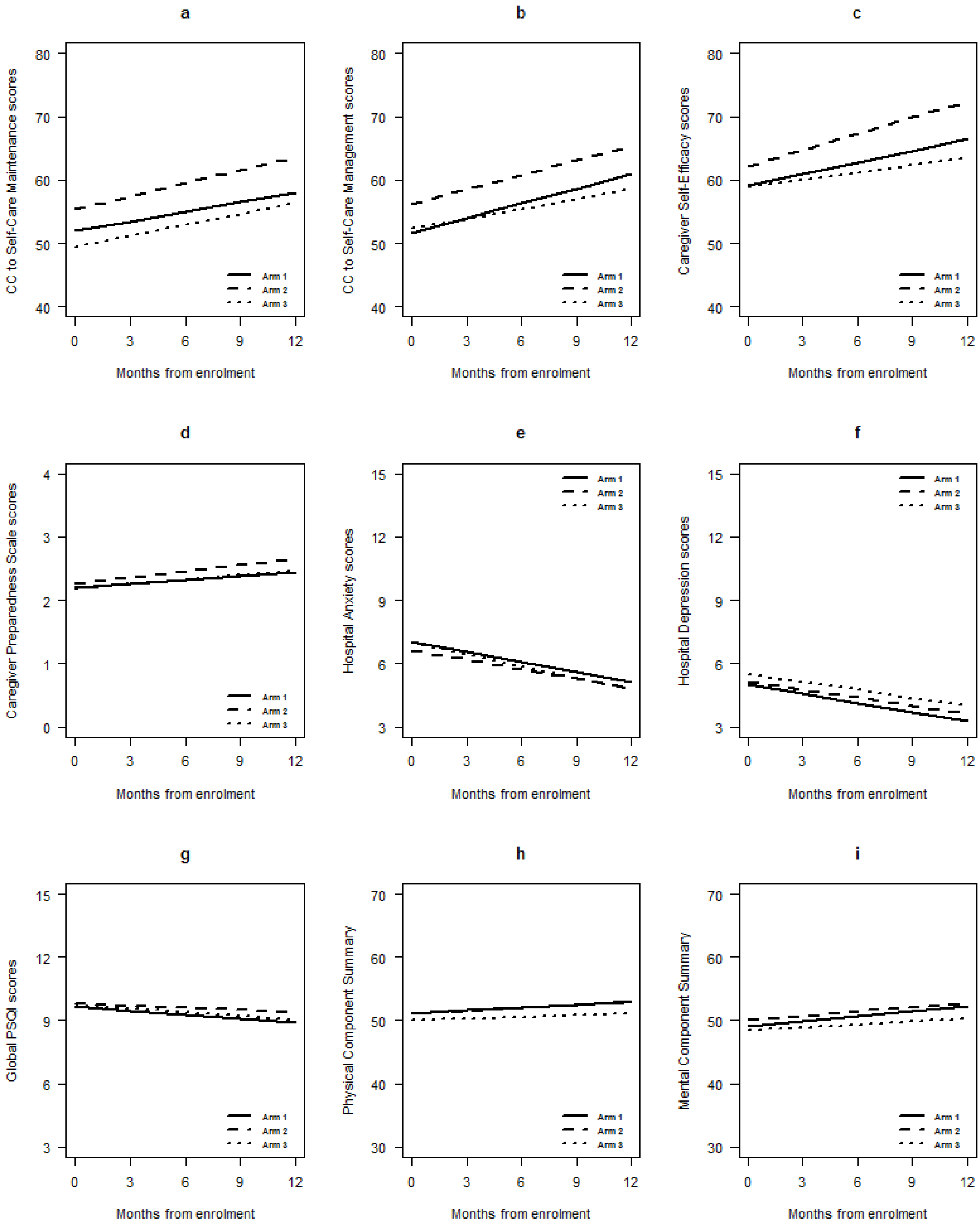
Table 10. Longitudinal linear mixed model results on secondary outcome scores in caregivers.

Effect	β	95% C.I.	*P* value
CC to self-care maintenance score			
Time for each visit (MI for patients and caregivers)	1.75	(0.60; 2.90)	0.0035
MI for patients and caregivers vs MI only for patients at baseline	3.49	(-0.54; 7.51)	0.0892
MI for patients and caregivers vs Standard of care at baseline	5.98	(2.15; 9.80)	0.0027
MI for patients and caregivers vs MI only for patients at follow-up	0.52	(-1.06; 2.10)	0.5156
MI for patients and caregivers vs Standard of care at follow-up	0.31	(-1.25; 1.87)	0.6942
Caregiver living with patient vs Not living with patient	-0.91	(-4.16; 2.34)	0.5828
CC to self-care management score			
Time for each visit (MI for patients and caregivers)	1.56	(0.40; 2.72)	0.0085
MI for patients and caregivers vs MI only for patients at baseline	4.53	(-0.46; 9.53)	0.0758
MI for patients and caregivers vs Standard of care at baseline	3.91	(-5.61; 4.35)	0.1025
MI for patients and caregivers vs MI only for patients at follow-up	0.00	(-1.59; 1.59)	0.9979
MI for patients and caregivers vs Standard of care at follow-up	0.76	(-0.81; 2.34)	0.3382
Caregiver living with patient vs Not living with patient	-2.06	(-6.09; 1.97)	0.3154
Caregiver self-efficacy score			
Time for each visit (MI for patients and caregivers)	1.15	(0.15; 2.16)	0.0248
MI for patients and caregivers vs MI only for patients at baseline	3.11	(-1.24; 7.46)	0.1616
MI for patients and caregivers vs Standard of care at baseline	3.34	(-0.79; 7.47)	0.1132
MI for patients and caregivers vs MI only for patients at follow-up	0.71	(-0.68; 2.09)	0.3164
MI for patients and caregivers vs Standard of care at follow-up	1.39	(0.02; 2.75)	0.0463
Caregiver living with patient vs Not living with patient	-2.09	(-5.59; 1.41)	0.2414
Caregiver Preparedness Scale (CPS)			
Time for each visit (MI for patients and caregivers)	0.07	(0.03; 0.11)	0.0005
MI for patients and caregivers vs MI only for patients at baseline	0.05	(-0.11; 0.21)	0.5361
MI for patients and caregivers vs Standard of care at baseline	0.07	(-0.08; 0.22)	0.3368
MI for patients and caregivers vs MI only for patients at follow-up	0.04	(-0.01; 0.09)	0.1123
MI for patients and caregivers vs Standard of care at follow-up	0.03	(-0.02; 0.08)	0.2275
Caregiver living with patient vs Not living with patient	-0.11	(-0.24; 0.01)	0.0776
Hospital Anxiety Scale (HAS)			
Time for each visit (MI for patients and caregivers)	-0.45	(-0.67; -0.24)	<0.0001
MI for patients and caregivers vs MI only for patients at baseline	-0.44	(-1.28; 0.41)	0.3100
MI for patients and caregivers vs Standard of care at baseline	-0.32	(-1.13; 0.48)	0.4269
MI for patients and caregivers vs MI only for patients at follow-up	0.04	(-0.28; 0.36)	0.8160
MI for patients and caregivers vs Standard of care at follow-up	0.09	(-0.22; 0.41)	0.5695
Caregiver living with patient vs Not living with patient	0.63	(-0.05; 1.30)	0.0692

Hospital Depression Scale (HDS)			
Time for each visit (MI for patients and caregivers)	-0.39	(-0.60; -0.18)	0.0003
MI for patients and caregivers vs MI only for patients at baseline	0.16	(-0.64; 0.97)	0.6894
MI for patients and caregivers vs Standard of care at baseline	-0.33	(-1.10; 0.43)	0.3942
MI for patients and caregivers vs MI only for patients at follow-up	0.05	(-0.27; 0.36)	0.7659
MI for patients and caregivers vs Standard of care at follow-up	-0.02	(-0.33; 0.29)	0.9133
Caregiver living with patient vs Not living with patient	0.81	(0.17; 1.45)	0.0133
Global PSQI Score			
Time for each visit (MI for patients and caregivers)	-0.10	(-0.22; 0.02)	0.1146
MI for patients and caregivers vs MI only for patients at baseline	0.16	(-0.53; 0.84)	0.6527
MI for patients and caregivers vs Standard of care at baseline	0.10	(-0.55; 0.76)	0.7569
MI for patients and caregivers vs MI only for patients at follow-up	0.10	(-0.07; 0.28)	0.2519
MI for patients and caregivers vs Standard of care at follow-up	0.08	(-0.10; 0.25)	0.3879
Caregiver living with patient vs Not living with patient	0.15	(-0.40; 0.69)	0.5940
SF-12 Physical Component Summary (PCS)			
Time for each visit (MI for patients and caregivers)	0.46	(0.12; 0.80)	0.0076
MI for patients and caregivers vs MI only for patients at baseline	-0.13	(-1.75; 1.48)	0.8743
MI for patients and caregivers vs Standard of care at baseline	0.94	(-0.60; 2.47)	0.2304
MI for patients and caregivers vs MI only for patients at follow-up	0.00	(-0.50; 0.51)	0.9915
MI for patients and caregivers vs Standard of care at follow-up	0.23	(-0.27; 0.72)	0.3683
Caregiver living with patient vs Not living with patient	-2.75	(-4.04; -1.46)	<0.0001
SF-12 Mental Component Summary (MCS)			
Time for each visit (MI for patients and caregivers)	0.64	(0.23; 1.04)	0.0021
MI for patients and caregivers vs MI only for patients at baseline	1.13	(-0.65; 2.92)	0.2142
MI for patients and caregivers vs Standard of care at baseline	1.68	(-0.02; 3.37)	0.0527
MI for patients and caregivers vs MI only for patients at follow-up	-0.18	(-0.79; 0.42)	0.5518
MI for patients and caregivers vs Standard of care at follow-up	0.19	(-0.41; 0.78)	0.5426
Caregiver living with patient vs Not living with patient	-0.64	(-2.06; 0.77)	0.3733

Note. MI: Motivational Interviewing; CI: confidence interval; CC: Caregiver Contribution; PSQI: Pittsburgh Sleep Quality Index.

Figure 9. Model-based secondary scores by follow-up time in caregivers.



Note. Caregiver Contribution to self-care maintenance score (panel a), Caregiver Contribution to self-care management score (panel b), Caregiver self-efficacy score (panel c), Caregiver Preparedness Scale score (panel d), Hospital Anxiety Scale score (panel e), Hospital Depression Scale score (panel f), Global Pittsburgh Sleep Quality Index score (panel g), SF-12 Physical Component Summary score (panel h), SF-12 Mental Component Summary score (panel i). Arm 1: motivational interviewing (MI) only for patients; Arm 2: MI for patients and caregivers; and Arm 3: standard of care.

5.8 Analysis of mutuality in patient-caregiver dyads

5.8.1 Level 1 model: within-dyad model

Table 11 provides an illustration of a single dyad data structure. Each dyad had 10 observations: 5 measurements for each subject of the dyad. The dyad ID is listed in the first column. The outcomes (mutuality score), with the responses of each subscale dimension (i.e., love and affection, shared pleasurable activities, shared values, and reciprocity) and for the total score, are shown from the second to the sixth column, with one row for each time point. Two dummy variables in the seventh and eighth columns stand in for the HF patient and the caregiver, respectively. The mutuality response was referring to the patient if the variable “HF Patient” was equal to 1 and to the caregiver if the variable “Caregiver” took value 1. The visit number for the HF patient and caregiver are shown in the ninth and tenth columns, respectively, and are referred to as “Visit time (HF Patient)” and “Visit time (Caregiver)”. The choice of the origin location has significant effects on how the growth curve trajectory parameters are to be interpreted, thus it should be carefully considered. In this application, as seen in the previous analyses, each time value has had a 1 subtracted from the visit number such that the intercept, which occurs when time equals 0, is at the initial or baseline assessment. The number 1 denotes the subsequent measurement, which took place three months after the study began, as well as the values 2, 3 and 4 denote the following evaluation points, which took place six, nine and twelve months after the study began. As a result, a time change of one unit corresponds to a period of three months.

Table 11. Level 1 record for one care dyad

Care Dyad ID	**Mutuality Scale (MS)**					**HF Patient**	**Caregiver**	**Visit time (HF Patient)**	**Visit time (Caregiver)**
	love and affection	Shared pleasurable activities	Shared values	Reciprocity	Total score				
1	4.0	3.5	4.0	3.8	3.8	1	0	0	0
1	3.3	2.7	2.5	3.0	2.9	1	0	1	0
1	.	.	.	.	.	1	0	2	0
1	3.6	1.7	1.5	2.5	2.4	1	0	3	0
1	3.3	3.2	3.0	3.0	3.1	1	0	4	0
1	4.0	3.7	2.5	3.8	3.6	0	1	0	0
1	3.0	4.0	3.0	3.0	3.2	0	1	0	1
1	.	.	.	.	.	0	1	0	2
1	3.0	2.2	2.5	2.3	2.4	0	1	0	3
1	3.3	3.2	3.0	2.6	3.0	0	1	0	4

Note. HF: heart failure.

5.8.2 Unconditional Level 2 model: between-dyad model

The maximum likelihood estimates of the fixed effects and the estimates of the variance components of the random effects, and their corresponding 95% confidence intervals for the unconditional model (i.e. no covariates included) on mutuality are shown in Table 12.

Regarding the fixed effects, when compared to caregivers, HF patients rated mutuality higher at baseline in the shared pleasurable activities ($\gamma_{10} = 2.76$ vs $\gamma_{30} = 2.71$), reciprocity ($\gamma_{10} = 2.83$ vs $\gamma_{30} = 2.68$) and total score ($\gamma_{10} = 2.87$ vs $\gamma_{30} = 2.82$) subscales; while they rated mutuality lower at baseline in love and affection ($\gamma_{10} = 3.21$ vs $\gamma_{30} = 3.34$) and shared values ($\gamma_{10} = 2.67$ vs $\gamma_{30} = 2.72$) subscales. Over the course of the 12-month period, there was evidence of significant positive linear trends for both HF patients and caregivers in all subscales of the Mutuality Scale (MS), as well as the total score ($\gamma_{20} = 0.06$, 95% CI: 0.04; 0.08, $p < 0.0001$ for HF patients, and $\gamma_{20} = 0.06$, 95% CI: 0.05; 0.08, $p < 0.0001$ for caregivers, respectively).

Regarding the random effects, we rejected the null hypothesis that the variance in the population is zero for each component. For both dyad members, there is a significant heterogeneity around the average score for each growth parameter (intercept and linear slope). The estimated correlations between the Level 2 random effects (the variance components), are shown in Table 13. The unconditional model, where the residual variance is interpreted as the degree of variability around the average growth parameter, can be used to determine the variances (and covariances) of the coefficients that make up the growth trajectories. The strength of shared variance in the outcomes for the dyad members is represented by notable correlations. In terms of the intercept, the correlation between HF patient and caregiver ranged from 0.53 to 0.70 ($p < 0.0001$) across the subscales of MS, indicating moderate-to-strong covariation within the dyads (Table 13). These estimates show a moderate amount of shared variance and justify our selection of a multilevel model.

Table 12. Multilevel model fixed and random effects results predicting patients' and caregivers' mutuality over time from the unconditional model

		Patient			**Caregiver**	
Model	**Estimate**	**95% C.I.**	***P* value**	**Estimate**	**95% C.I.**	***P* value**
Love and affection						
Fixed effects	Coefficient			Coefficient		
Intercept [§]	3.21	(3.16 ; 3.27)	<.0001	3.34	(3.28 ; 3.39)	<.0001
Linear slope [¥]	0.02	(0.00 ; 0.04)	0.0202	0.04	(0.02 ; 0.06)	0.0001
Random effects	Variance component			Variance component		
Intercept	0.29	(0.24 ; 0.35)	<.0001	0.25	(0.20 ; 0.30)	<.0001
Linear slope	0.01	(0.01 ; 0.02)	0.0002	0.01	(0.00 ; 0.01)	0.0035
Shared pleasurable activities						
Fixed effects	Coefficient			Coefficient		
Intercept [§]	2.76	(2.70 ; 2.81)	<.0001	2.71	(2.65 ; 2.78)	<.0001
Linear slope [¥]	0.07	(0.05 ; 0.09)	<.0001	0.07	(0.05 ; 0.09)	<.0001
Random effects	Variance component			Variance component		
Intercept	0.32	(0.27 ; 0.38)	<.0001	0.39	(0.33 ; 0.47)	<.0001
Linear slope	0.01	(0.01 ; 0.02)	<.0001	0.01	(0.01 ; 0.02)	<.0001
Shared values						
Fixed effects	Coefficient			Coefficient		
Intercept [§]	2.67	(2.61 ; 2.74)	<.0001	2.72	(2.65 ; 2.80)	<.0001
Linear slope [¥]	0.07	(0.04 ; 0.09)	<.0001	0.06	(0.03 ; 0.08)	<.0001
Random effects	Variance component			Variance component		
Intercept	0.47	(0.40 ; 0.56)	<.0001	0.51	(0.43 ; 0.60)	<.0001
Linear slope	0.03	(0.02 ; 0.04)	<.0001	0.02	(0.01 ; 0.04)	<.0001
Reciprocity						
Fixed effects	Coefficient			Coefficient		
Intercept [§]	2.83	(2.77 ; 2.89)	<.0001	2.68	(2.61 ; 2.74)	<.0001
Linear slope [¥]	0.07	(0.05 ; 0.09)	<.0001	0.08	(0.06 ; 0.10)	<.0001
Random effects	Variance component			Variance component		
Intercept	0.29	(0.24 ; 0.35)	<.0001	0.45	(0.39 ; 0.54)	<.0001
Linear slope	0.01	(0.01 ; 0.02)	<.0001	0.02	(0.01 ; 0.03)	<.0001
Total score						
Fixed effects	Coefficient			Coefficient		
Intercept [§]	2.87	(2.81 ; 2.92)	<.0001	2.82	(2.77 ; 2.88)	<.0001
Linear slope [¥]	0.06	(0.04 ; 0.08)	<.0001	0.06	(0.05 ; 0.08)	<.0001
Random effects	Variance component			Variance component		
Intercept	0.25	(0.21 ; 0.31)	<.0001	0.29	(0.25 ; 0.35)	<.0001
Linear slope	0.01	(0.01 ; 0.02)	<.0001	0.01	(0.01 ; 0.02)	<.0001

Note. CI: confidence interval.

[§] Mean MS at baseline.

[¥] Rate of change per 3 months in MS (assumed to be linear over the 12-month period).

Table 13. Multilevel model random-effect results: tau correlations between patient and caregiver mean MS scores at baseline

Mutuality Scale (MS)	**Tau correlation**	**95% C.I.**	***P* value**
Love and affection	0.62	(0.51; 0.73)	<.0001
Shared pleasurable activities	0.60	(0.51; 0.68)	<.0001
Shared values	0.70	(0.62; 0.77)	<.0001
Reciprocity	0.53	(0.44; 0.63)	<.0001
Total score	0.58	(0.49; 0.68)	<.0001

Note. CI: confidence interval.

5.8.3 Conditional Level 2 model: explaining variation in growth parameters

Predictors at Level 2 are added to a conditional model to account for the variation in change between dyads. The following is the conditional model that incorporates the predictors for treatment arm, the interaction with visit number and a binary indicator indicating the caregiver's cohabitation with the HF patient. The treatment arm and the cohabitation between patient and caregiver are two variables in this model that have common values for all members of the dyad. To investigate the relationship with the variation in HF patient's and caregiver's mutuality trajectories, the Level 2 model simultaneously included the variables treatment arm and caregiver's cohabitation with the HF patient. The coefficients γ_{13}, γ_{14}, γ_{23} and γ_{24} in this model reflect the effect of treatment arm on mutuality over time, controlling for the influence of cohabitation between HF patient and caregiver. Table 14 reports the results from the conditional dyadic model on the trends in Mutuality Scale (MS) over the follow-up time.

The motivational interview did not show any impact on changes in the HF patient's and caregiver's MS dimensions during the follow-up time. In fact, Arms 1 and 2 did not improve significantly more than Arm 3, neither in any subscale nor in the total score ($\gamma_{13} = 0$, 95% CI: -0.05; 0.04, $p = 0.872$ and $\gamma_{14} = -0.02$, 95%CI: -0.06; 0.03, $p = 0.407$ in patients; $\gamma_{23} = 0.03$, 95% CI: -0.02; 0.07, $p = 0.256$ and $\gamma_{24} = 0.02$, 95% CI: -0.02; 0.07, $p = 0.273$ in caregivers, respectively; Table 14, Figure 10). The difference that was closest to being significant was observed in the improvement in the love and affection subscale of MS for caregivers in Arm 1 compared to Arm 3 over time ($\gamma_{23} = 0.05$, 95% CI: 0; 0.09, $p = 0.063$).

The living together condition was significantly associated with HF patient's MS for love and affection ($\gamma_{15} = 0.22$, 95% CI: 0.11; 0.33, $p < 0.0001$), shared values ($\gamma_{15} = 0.22$, 95% CI: 0.09; 0.35, $p = 0.001$), and total score ($\gamma_{15} = 0.11$, 95% CI: 0.02; 0.21, $p = 0.023$), and in caregiver's MS for shared values ($\gamma_{25} = 0.22$, 95% CI: 0.09; 0.35, $p = 0.001$; Table 14). Taking into account the total score, on average, HF patients living with caregivers had significantly higher levels of mutuality at baseline than HF patients not living with caregivers, controlling for treatment arm. There was no significant effect of cohabitation between HF patient and caregiver on caregiver's baseline mutuality.

The random effects for the conditional model are shown in Table 14. The variance components showed that there was significant variability around the average trajectories for both the HF patient's and the caregiver's MS dimensions (i.e., total score: $u_{1p} = 0.25$, 95% CI: 0.21; 0.3, $p < 0.0001$ and $u_{2p} = 0.01$, 95% CI: 0.01; 0.02, $p < 0.0001$ in patient's intercept and linear slope, respectively; $u_{3p} = 0.29$, 95% CI: 0.24; 0.35, $p < 0.0001$ and $u_{4p} = 0.01$, 95% CI: 0.01; 0.02, $p < 0.0001$ in caregiver's intercept and linear slope, respectively; Table 14). By subtracting the residual variance component in the conditional model from the residual variance component in the unconditional model, and then converting the result to a percentage, it is possible to determine the percentage of variance in each growth parameter that is explained by the conditional model. For instance, the model for love and affection subscale explains 7% of the variance in the baseline mutuality of HF patient ($[0.29 - 0.27]/0.29 = 0.07$) and 4% of the variance in the baseline mutuality of caregiver ($[0.25 - 0.24]/0.25 = 0.04$). Overall, compared to the caregiver's trajectory, the two predictors are better at explaining variation in the trajectory of the HF patient, accounting for 7% of baseline mutuality variation. Both HF patients and caregivers had the same variance in the linear slope (Table 14).

Table 14. Multilevel model fixed and random effects results predicting patient and caregiver MS scores over time from treatment arm, its interaction with visit number and living together condition

Mutuality Scale (MS)	**Patient**			**Caregiver**		
Model	**Estimate**	**95% C.I.**	***P* value**	**Estimate**	**95% C.I.**	***P* value**
Love and affection						
Fixed effects	Coefficient			Coefficient		
Intercept §	3.04	(2.93; 3.16)	<.0001	3.33	(3.22; 3.44)	<.0001
Linear slope ¥	0.02	(-0.01; 0.06)	0.2194	0.02	(-0.02; 0.05)	0.3386
MI only for patients vs Standard of care at baseline	0.02	(-0.13; 0.16)	0.8184	-0.14	(-0.28; -0.01)	0.0413
MI for patients and caregivers vs Standard of care at baseline	0.09	(-0.05; 0.23)	0.1935	-0.01	(-0.14; 0.12)	0.8788
MI only for patients vs Standard of care at follow-up	0.03	(-0.03; 0.07)	0.3299	0.05	(0.00; 0.09)	0.0626
MI for patients and caregivers vs Standard of care at follow-up	-0.02	(-0.06; 0.03)	0.5282	0.02	(-0.03; 0.07)	0.3866
Living together (Yes vs No)	0.22	(0.11; 0.33)	<.0001	0.09	(-0.01; 0.19)	0.0797
Random effects	Variance component			Variance component		
Intercept	0.27	(0.22; 0.33)	<.0001	0.24	(0.19; 0.30)	<.0001
Linear slope	0.01	(0.01; 0.02)	0.0003	0.01	(0.00; 0.01)	0.0047
Shared pleasurable activities						
Fixed effects	Coefficient			Coefficient		
Intercept §	2.67	(2.55; 2.78)	<.0001	2.74	(2.61; 2.87)	<.0001
Linear slope ¥	0.09	(0.05; 0.12)	<.0001	0.04	(0.01; 0.08)	0.0202
MI only for patients vs Standard of care at baseline	0.07	(-0.08; 0.22)	0.3404	-0.10	(-0.26; 0.06)	0.2091
MI for patients and caregivers vs Standard of care at baseline	0.12	(-0.02; 0.26)	0.0848	-0.05	(-0.20; 0.10)	0.5113
MI only for patients vs Standard of care at follow-up	-0.03	(-0.08; 0.03)	0.3092	0.03	(-0.02; 0.09)	0.2274
MI for patients and caregivers vs Standard of care at follow-up	-0.01	(-0.06; 0.04)	0.6191	0.04	(-0.01; 0.09)	0.1057
Living together (Yes vs No)	0.04	(-0.07; 0.15)	0.4352	0.04	(-0.08; 0.16)	0.4997
Random effects	Variance component			Variance component		
Intercept	0.31	(0.26; 0.38)	<.0001	0.39	(0.33; 0.47)	<.0001
Linear slope	0.02	(0.01; 0.03)	<.0001	0.01	(0.01; 0.02)	<.0001
Shared values						
Fixed effects	Coefficient			Coefficient		
Intercept §	2.54	(2.40; 2.67)	<.0001	2.57	(2.42; 2.71)	<.0001
Linear slope ¥	0.07	(0.02; 0.11)	0.0026	0.05	(0.01; 0.10)	0.0137
MI only for patients vs Standard of care at baseline	-0.10	(-0.27; 0.07)	0.2430	-0.02	(-0.19; 0.16)	0.8679
MI for patients and caregivers vs Standard of care at baseline	0.08	(-0.09; 0.24)	0.3524	0.07	(-0.10; 0.24)	0.4317
MI only for patients vs Standard of care at follow-up	0.02	(-0.05; 0.08)	0.5550	0.02	(-0.04; 0.08)	0.5085
MI for patients and caregivers vs Standard of care at follow-up	-0.02	(-0.08; 0.04)	0.5042	0.00	(-0.06; 0.05)	0.9039
Living together (Yes vs No)	0.22	(0.09; 0.35)	0.0008	0.22	(0.09; 0.35)	0.0013
Random effects	Variance component			Variance component		
Intercept	0.44	(0.37; 0.53)	<.0001	0.49	(0.41; 0.58)	<.0001
Linear slope	0.03	(0.02; 0.04)	<.0001	0.02	(0.01; 0.04)	<.0001
Reciprocity						
Fixed effects	Coefficient			Coefficient		
Intercept §	2.74	(2.62; 2.85)	<.0001	2.60	(2.47; 2.74)	<.0001
Linear slope ¥	0.08	(0.05; 0.12)	<.0001	0.06	(0.02; 0.10)	0.0013
MI only for patients vs Standard of care at baseline	0.00	(-0.15; 0.14)	0.9580	0.00	(-0.17; 0.17)	0.9891
MI for patients and caregivers vs Standard of care at baseline	0.14	(0.01; 0.28)	0.0384	0.04	(-0.13; 0.20)	0.6629
MI only for patients vs Standard of care at follow-up	-0.01	(-0.06; 0.04)	0.7430	0.02	(-0.04; 0.07)	0.5018
MI for patients and caregivers vs Standard of care at follow-up	-0.03	(-0.07; 0.02)	0.3008	0.03	(-0.03; 0.08)	0.3323
Living together (Yes vs No)	0.07	(-0.03; 0.18)	0.1766	0.10	(-0.03; 0.23)	0.1272
Random effects	Variance component			Variance component		
Intercept	0.28	(0.23; 0.34)	<.0001	0.45	(0.38; 0.53)	<.0001

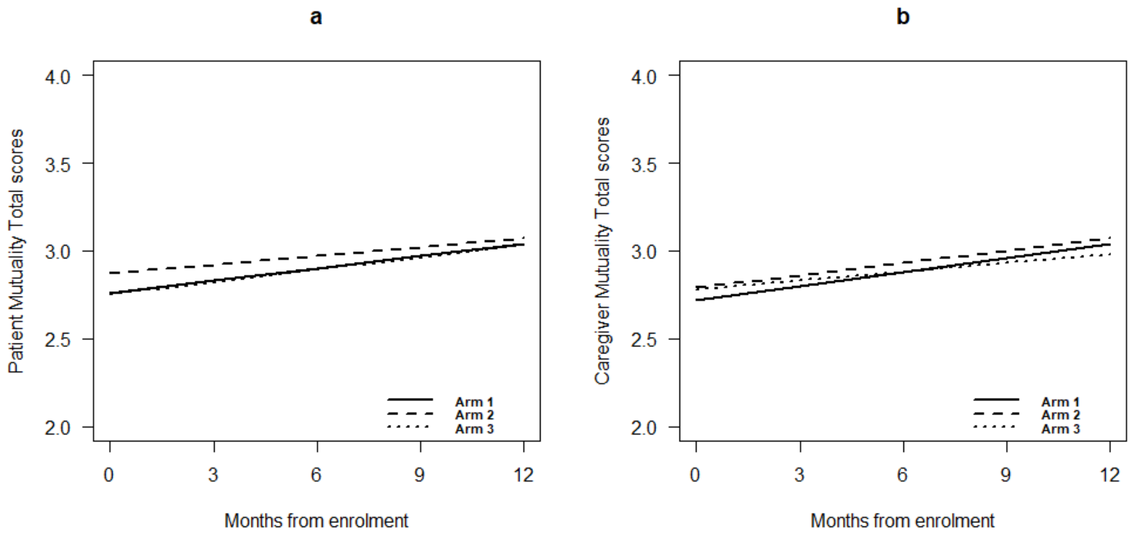
Linear slope	0.01	(0.01; 0.02)	<.0001	0.02	(0.01; 0.03)	<.0001
Total score						
Fixed effects	Coefficient			Coefficient		
Intercept §	2.75	(2.65; 2.86)	<.0001	2.78	(2.67; 2.89)	<.0001
Linear slope ¥	0.07	(0.04; 0.10)	<.0001	0.05	(0.01; 0.08)	0.0045
MI only for patients vs Standard of care at baseline	0.01	(-0.12; 0.14)	0.9086	-0.06	(-0.20; 0.08)	0.4088
MI for patients and caregivers vs Standard of care at baseline	0.12	(-0.01; 0.24)	0.0649	0.01	(-0.13; 0.14)	0.9175
MI only for patients vs Standard of care at follow-up	0.00	(-0.05; 0.04)	0.8715	0.03	(-0.02; 0.07)	0.2564
MI for patients and caregivers vs Standard of care at follow-up	-0.02	(-0.06; 0.03)	0.4071	0.02	(-0.02; 0.07)	0.2731
Living together (Yes vs No)	0.11	(0.02; 0.21)	0.0227	0.10	(-0.01; 0.20)	0.0632
Random effects	Variance component			Variance component		
Intercept	0.25	(0.21; 0.30)	<.0001	0.29	(0.24; 0.35)	<.0001
Linear slope	0.01	(0.01; 0.02)	<.0001	0.01	(0.01; 0.02)	<.0001

Note. MI: Motivational Interviewing; CI: confidence interval.

¥ Rate of change per 3 months in MS (assumed to be linear over the 12-month period).

§ Mean MS at baseline in Arm 3 for dyads not living together.

Figure 10. Model-based Mutuality total scores by follow-up time in patients (a) and caregivers (b)



Note. Arm 1: motivational interviewing (MI) only for patients; Arm 2: MI for patients and caregivers; and Arm 3: standard of care.

5.9 Analysis of dropouts

Like many other follow-up studies, the MOTIVATE-HF has a significant attrition rate. Table 15 lists the number and percentage of HF patients who dropped out at each study visit after the baseline. The analysis of the primary endpoint included a total of 406 patients who completed T1 (3 months after enrolment) resulting 20.4% in patient attrition. At the end of follow-up, 238 HF patients (46.7%) completed the full 12-month trial. The whole participant flow was shown in Figure 4. Furthermore, dropout might be closely related to self-care maintenance scores that are missing as a result of dropout, meaning that the unobserved self-care maintenance scores among those who drop out are consistently lower than those who continue follow-up, even after potentially adjusting for observed covariates. Pattern mixing models can be used to model the dependence between missing responses and dropout under a few specific but mostly untestable assumptions.

Table 15. Patterns of observing self-care maintenance score in patients by visit number

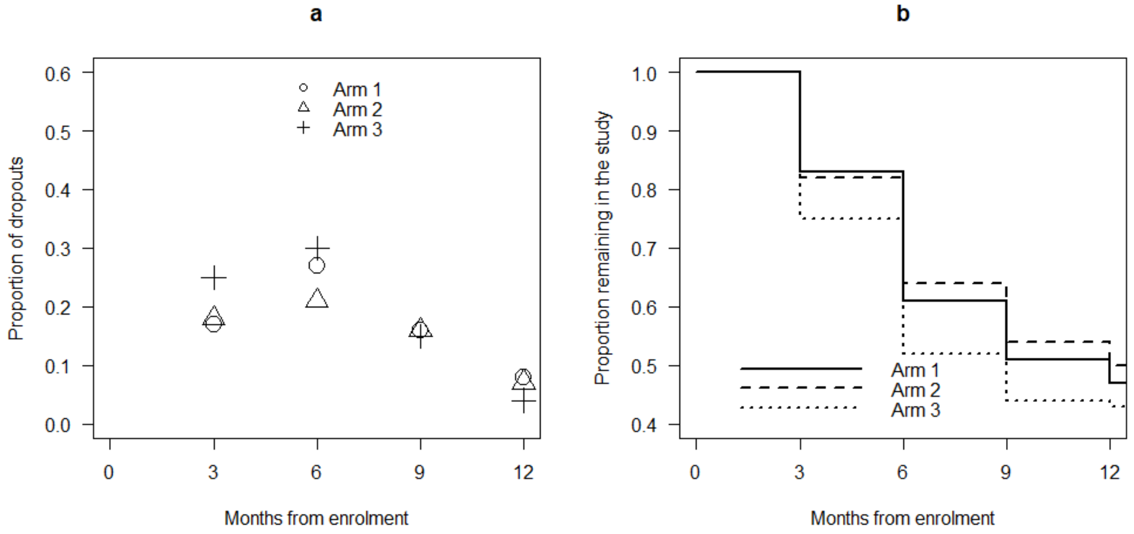
Visit	In follow-up, n (%)	No. of patients for which self-care maintenance was observed, n (%)	Dropped out
Baseline	510 (100)	510 (100)	0
T1	406 (79.6)	363 (89.4)	104 (13 died)
T2	301 (59.0)	293 (97.3)	105 (7 died)
T3	254 (49.8)	252 (99.2)	47 (6 died)
T4	238 (46.7)	238 (100)	16 (2 died)

Note. Missing data at each visit is a combination of intermittent missingness and dropout. T1, T2, T3 and T4 correspond to 3, 6, 9, and 12 months from enrolment.

Figure 11 illustrates two crucial steps in the cessation study. Panel “a” indicates the ratio of participants who have dropped out by visit, stratified by treatment group. The ratio was calculated utilizing the number of dropouts at the current follow-up as the numerator and the observed patients in the previous follow-up as the denominator. With the pattern of missing data depicted in this panel, we can obtain a valid estimate under the MCAR assumption (observed data is a random draw from the whole data). On the other hand, under the MAR assumption, the probability of dropout at each occasion can vary depending on the outcomes that have already been recorded up to, but excluding the current occasion; while under the MNAR assumption, this probability can depend on current and future unobserved outcomes.

Panel “b” displays the proportion of remaining in follow-up as a function of visit. Even while the proportion of completers is quite similar at the end of the trial (47%, 50% and 43% in Arms 1, 2 and 3, respectively), their pattern is different: most dropouts on the control arm (Arm 3) did so immediately in the first 6 months. Within this first half period of the trial, 64% (114/177) of those in Arm 2 and similarly to 61% (94/155) in Arm 1 remained in the study compared to just 52% (93/178) of those in the control group.

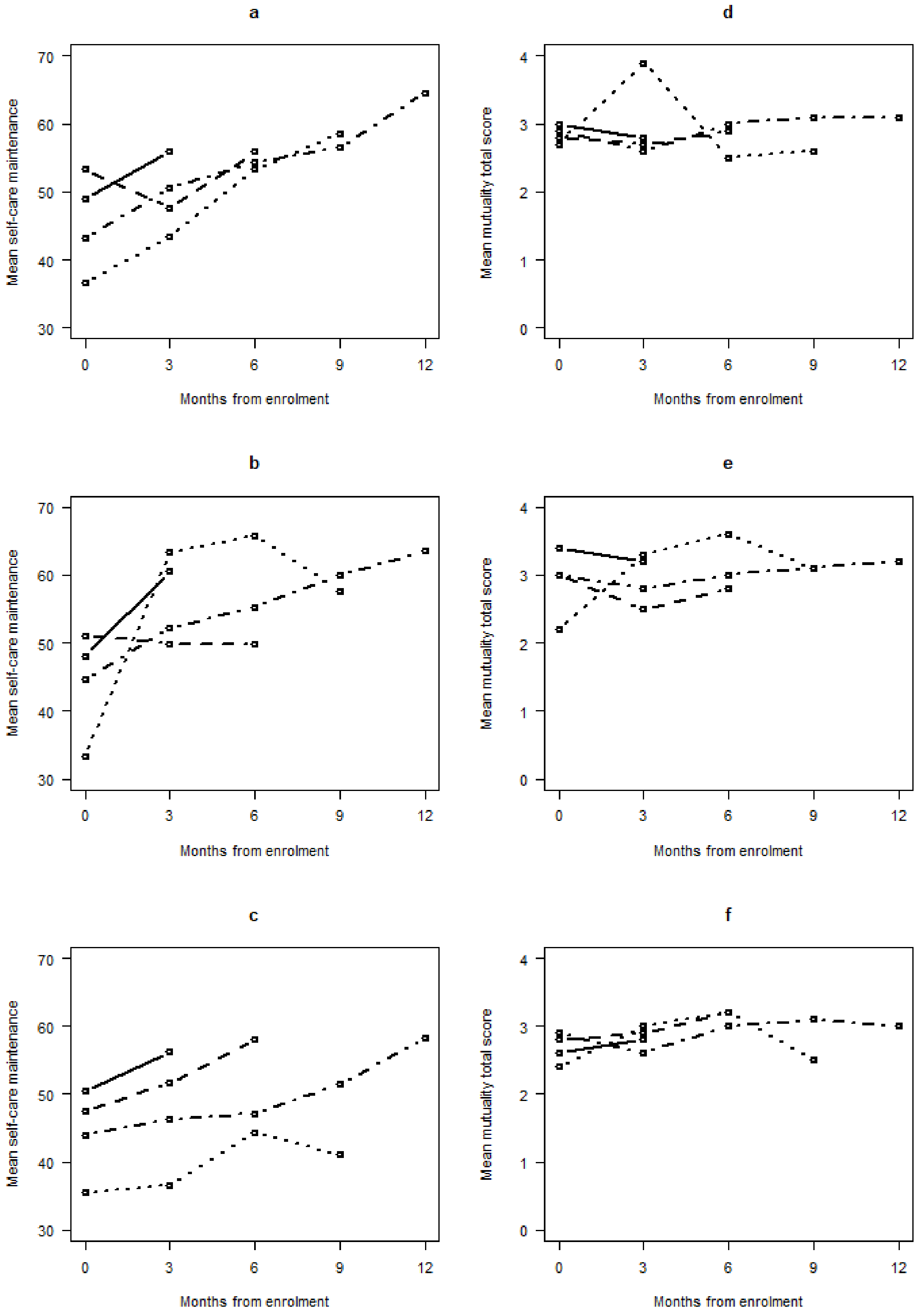
Figure 11. Proportion of dropouts (a) and proportion remaining in the study (b) in each visit, by treatment group



Note. Arm 1: motivational interviewing (MI) only for patients; Arm 2: MI for patients and caregivers; and Arm 3: standard of care.

Next, we looked at the visit-specific means of self-care maintenance and mutuality total scores in patients with different dropout, at each visit. Figure 12 compares these observed means among the three treatment arms in terms of self-care maintenance score (panels a-c) and mutuality total score (panels d-f). Completers (i.e., patients in the fourth dropout category) did not show a different pattern as compared with patients who dropped out in terms of visit-specific means of both self-care maintenance and mutuality total scores.

Figure 12. Observed means of self-care maintenance (a-c) and mutuality total score (d-f) in patients, stratified by dropout categories at each visit, and treatment arm.



Note. Panels a-c refer to mean self-care maintenance scores stratified by dropout categories at each visit in Arms 1, 2 and 3, respectively. Panels d-f refer to mean mutuality total scores stratified by dropout categories at each visit in Arms 1, 2 and 3, respectively.

Finally, a multivariable logistic regression model was applied to evaluate the association between dropouts during follow-up and several characteristics measured at baseline. Starting from the participant flow of the MOTIVATE-HF trial shown in Figure 4, a binary variable indicating patient lost at follow-up with the reason of refusal was created as response of the model. Dead patients were not considered as lost at follow-up. Specifically, 266 (238+28 deaths, 52.2%) patients completed the trial, while 244 (47.8%) dropped out before the end of the study. The baseline characteristics of patients present in Table 3 were included in the adjusted model of Table 16 as covariates, including the treatment arm. Self-care management score was not included since, as already specified, it can be computed only if patients have had HF symptoms in the last month (354, 69.4%).

Particularly, lower age (OR: 0.97, 95% CI: 0.95; 0.99, p = 0.032) and MoCA scores (OR: 0.91, 95% CI: 0.87; 0.94, p < 0.0001), higher self-care confidence score (OR: 1.02, 95% CI: 1.01; 1.03, p = 0.004), male gender (OR: 1.71, 95% CI: 1.08; 2.72, p = 0.022), lower income (higher vs lower income: OR: 0.28, 95% CI: 0.08; 0.88, p = 0.009) and not cohabitation between patient and caregiver (OR: 1.99, 95% C: 1.27; 3.10, p = 0.002) were associated with dropouts during follow-up among patients at multivariable analysis.

Table 16. Logistic regression model on dropouts during follow-up and baseline characteristics

Effect (n = 479, dropout = 226)	**OR**	**95% C.I.**	***P* value**
Arm 1: MI only for patients vs Arm 3: standard care	0.82	(0.49; 1.33)	0.4276
Arm 2: MI for patients and caregivers vs Arm 3: standard care	0.95	(0.59; 1.52)	0.8069
Age (years)	0.97	(0.95; 0.99)	0.0320
Time with HF (months)	0.99	(0.99; 1.00)	0.2889
N. of medications	1.00	(0.93; 1.08)	0.9419
CCI scores	1.03	(0.93; 1.15)	0.5820
MoCA scores	0.91	(0.87; 0.94)	<.0001
Self-care maintenance score	1.01	(0.99; 1.02)	0.4494
Self-care confidence score	1.02	(1.01; 1.03)	0.0040
Sex (male vs female)	1.71	(1.08; 2.72)	0.0223
Marital status (married vs not married)	1.06	(0.64; 1.74)	0.8265
Education (high schools or higher vs lower)	1.59	(0.95; 2.68)	0.0753
Employment (retired vs employed)	1.42	(0.74; 2.70)	0.2903
Income (the necessary to live vs not the necessary to live)	0.62	(0.22; 1.71)	0.6157
Income (more than the necessary to live vs not the necessary to live)	0.28	(0.08; 0.88)	0.0088
NYHA Class (III vs II)	1.17	(0.74; 1.86)	0.5763
NYHA Class (IV vs II)	1.88	(0.79; 4.49)	0.1959
Caregiver living with patient (no vs yes)	1.99	(1.27; 3.10)	0.0024

Note. OR: odds ratio; CI: confidence interval; HF: heart failure; CCI: Charlson Comorbidity Index; MoCA: Montreal Cognitive Assessment; NYHA: New York Heart Association.

As a sensitivity analysis on endpoints that resulted significantly affected by MI, the variables significantly associated with dropouts during follow-up (model of Table 16) were used as covariates in adjusted linear mixed models. Results are shown in Table 17 on self-care maintenance, self-care management, KCCQ and HFSPS scores. In addition, we also performed a sensitivity analysis on the dyadic endpoint mutuality total score that is the main focus of this thesis in Table 18. Even after accounting for these additional factors (patient age, sex, income, cohabitation with caregiver, MoCA score, and baseline self-care confidence scale score), the results of the models remained consistent. Reminding that, in the primary analysis in Table 9, Arm 1 improved in self-care maintenance over the year just barely enough to be statistically significant compared to Arm 3 ($p = 0.046$), in the adjusted model Arm 1 did not significantly differ from Arm 3 ($p = 0.066$, Table 17). In terms of self-care management, both Arms 1 and 2 considerably improved more than Arm 3 ($\beta = 2.47$, 95% CI: 0.40; 4.54, $p = 0.019$ and $\beta = 3.19$, 95% CI: 1.21; 5.17, $p = 0.002$ in Arms 1 and 2, respectively, Table 17), while in the primary analysis only Arm 2 showed a significant improvement compared to Arm 3 (Table 9).

Table 17. Longitudinal linear mixed model results on primary and secondary outcome scores in patients adjusted for variables significantly associated with dropout during follow-up

Effect (n = 493)	**β**	**95% C.I.**	***P* value**
Self-care maintenance scores			
Time (for each visit)	3.59	(2.77; 4.41)	<.0001
MI only for patients vs Standard of care	1.84	(-1.52; 5.19)	0.2846
MI for patients and caregivers vs Standard of care	2.01	(-1.19; 5.20)	0.2188
Time * MI only for patients	1.11	(-0.07; 2.28)	0.0657
Time * MI for patients and caregivers	1.00	(-0.11; 2.12)	0.0795
Age (years)	0.24	(0.13; 0.35)	<.0001
MoCA Scores	0.09	(-0.11; 0.31)	0.3656
Self-care confidence score	0.25	(0.20; 0.31)	<.0001
Sex (male vs female)	0.46	(-2.07; 3.00)	0.7209
Income (the necessary to live vs not the necessary to live)	-0.03	(-5.95; 5.88)	0.9917
Income (more than the necessary vs not the necessary to live)	-7.01	(-13.53; -0.49)	0.0352
Caregiver living with patient (yes vs no)	2.64	(0.07; 5.20)	0.0437
Self-care management scores			
Time (for each visit)	3.37	(1.90; 4.84)	<.0001
MI only for patients vs Standard of care	1.15	(-3.47; 5.77)	0.6255
MI for patients and caregivers vs Standard of care	-1.54	(-5.82; 2.74)	0.4803
Time * MI only for patients	2.47	(0.40; 4.54)	0.0192
Time * MI for patients and caregivers	3.19	(1.21; 5.17)	0.0016
Age (years)	-0.03	(-0.17; 0.11)	0.6881
MoCA Scores	0.15	(-0.12; 0.42)	0.2905
Self-care confidence score	0.28	(0.20; 0.35)	<.0001
Sex (male vs female)	-2.28	(-5.52; 0.94)	0.1649
Income (the necessary to live vs not the necessary to live)	4.12	(-3.34; 11.58)	0.2785
Income (more than the necessary vs not the necessary to live)	4.32	(-3.88; 12.53)	0.3013
Caregiver living with patient (yes vs no)	0.04	(-3.24; 3.32)	0.9814
KCCQ Overall Summary Score			
Time (for each visit)	2.27	(1.26; 3.27)	<0.0001
MI only for patients vs Standard of care	-1.25	(-5.30; 2.80)	0.5446
MI for patients and caregivers vs Standard of care	-0.34	(-4.19; 3.51)	0.8621
Time * MI only for patients	1.02	(-0.41; 2.45)	0.1611
Time * MI for patients and caregivers	1.48	(0.12; 2.84)	0.0330
Age (years)	-0.35	(-0.49; -0.21)	<.0001
MoCA Scores	0.09	(-0.17; 0.37)	0.4687
Self-care confidence score	0.23	(0.15; 0.30)	<.0001
Sex (male vs female)	7.81	(4.56; 11.05)	<.0001
Income (the necessary to live vs not the necessary to live)	1.53	(-6.01; 9.07)	0.6904
Income (more than the necessary vs not the necessary to live)	-1.68	(-10.03; 6.66)	0.6922
Caregiver living with patient (yes vs no)	-0.39	(-3.66; 2.88)	0.8142
HFSPS Total Score			
Time (for each visit)	-1.65	(-2.52; -0.78)	0.0002
MI only for patients vs Standard of care	-1.50	(-4.86; 1.85)	0.3798
MI for patients and caregivers vs Standard of care	-0.76	(-3.96; 2.43)	0.6396
Time* MI only for patients	0.07	(-1.16; 1.31)	0.9061
Time* MI for patients and caregivers	-1.36	(-2.54; -0.18)	0.0231
Age (years)	0.05	(-0.05; 0.17)	0.3304
MoCA Scores	-0.24	(-0.46; -0.01)	0.0339
Self-care confidence score	-0.15	(-0.21; -0.09)	<.0001
Sex (male vs female)	-6.68	(-9.37; -4.00)	<.0001
Income (the necessary to live vs not the necessary to live)	-2.14	(-8.38; 4.09)	0.5002
Income (more than the necessary vs not the necessary to live)	-2.46	(-9.36; 4.43)	0.4834

Caregiver living with patient (yes vs no)	-0.47	(-3.17; 2.23)	0.7320

Note. MI: Motivational Interviewing; CI: confidence interval; MoCA: Montreal Cognitive Assessment; KCCQ: Kansas City Cardiomyopathy Questionnaire; HFSPS: Heart Failure Somatic Perception Scale.

Table 18. Multilevel model fixed and random effects results predicting patient and caregiver mutuality total score over time, adjusted for variables significantly associated with dropout during follow-up

Mutuality Scale (MS) Total score (n = 493)	**Patient**			**Caregiver**		
	Estimate	**95% C.I.**	***P* value**	**Estimate**	**95% C.I.**	***P* value**
Fixed effects	Coefficient			Coefficient		
Intercept	2.20	(1.71; 2.69)	<.0001	2.96	(2.42; 3.48)	<.0001
Linear slope ¥	0.07	(0.04; 0.11)	<.0001	0.05	(0.02; 0.08)	0.0030
MI only for patients vs Standard of care at baseline	0.00	(-0.12; 0.13)	0.9546	-0.05	(-0.19; 0.09)	0.4461
MI for patients and caregivers vs Standard of care at baseline	0.11	(-0.01; 0.23)	0.0801	0.01	(-0.12; 0.14)	0.8827
MI only for patients vs Standard of care at follow-up	0.00	(-0.05; 0.05)	0.9704	0.03	(-0.02; 0.07)	0.2500
MI for patients and caregivers vs Standard of care at follow-up	-0.02	(-0.06; 0.02)	0.4683	0.02	(-0.02; 0.06)	0.3632
Living together (yes vs no)	0.13	(0.02; 0.22)	0.0157	0.08	(-0.02; 0.19)	0.1282
Patient age (years)	0.00	(0.00; 0.01)	0.1775	0.00	(-0.01; 0.00)	0.1098
Patient MoCA Score	0.00	(-0.01; 0.01)	0.6672	0.00	(-0.01; 0.01)	0.4565
Patient self-care confidence score	0.00	(0.00; 0.01)	<.0001	0.00	(-0.01; 0.00)	0.0979
Patient sex (male vs female)	-0.03	(-0.13; 0.07)	0.5913	0.01	(-0.10; 0.11)	0.9255
Patient income (the necessary to live vs not the necessary to live)	0.07	(-0.17; 0.31)	0.5539	-0.11	(-0.36; 0.15)	0.4098
Patient income (more than the necessary to live vs not the necessary to live)	0.04	(-0.22; 0.29)	0.7796	-0.02	(-0.29; 0.26)	0.9060
Random effects	Variance component			Variance component		
Intercept	0.23	(0.19; 0.28)	<.0001	0.28	(0.24; 0.34)	<.0001
Linear slope	0.01	(0.01; 0.02)	<.0001	0.01	(0.01; 0.02)	<.0001

Note. MI: Motivational Interviewing; CI: confidence interval; MoCA: Montreal Cognitive Assessment.

¥ Rate of change per 3 months in MS (assumed to be linear over the 12-month period).

6. Discussion

6.1 Findings of the MOTIVATE-HF trial

The objective of this randomized controlled trial was to assess the effect of a motivational interview (MI) in improving self-care maintenance in patients with heart failure. This investigation also enabled us to determine whether MI affected other secondary variables. In this study, we showed that MI was beneficial in improving both self-care maintenance (the primary endpoint) and self-care management three months following enrolment. The two active intervention groups also showed improvements in self-care maintenance, management, and confidence over time. These results are crucial since altering self-care behaviours in HF is challenging, and a prior study found that self-care maintenance improved after three months only after adjusting for confounding factors [15].

In this project, we examined the impact of MI over a one-year period. Unexpectedly, we saw that MI had a mixed and protracted effect on self-care. For instance, MI had a significant effect on self-care maintenance at all follow-up intervals, on self-care management 6 months after enrolment, on self-care confidence 6 and 9 months after enrolment. When we used the mixed model analysis from baseline to 12 months, controlling for missing data and confounding variables, these results were partially confirmed. This study demonstrated that, over a year, even if the caregiver in the same dyad did not get the intervention (Arm 1), this intervention barely enough significantly increased patient self-care maintenance compared to the control arm. However, MI was effective for self-care management only when it was also performed with caregivers (Arm 2) but, after adjusting for potential confounders, MI was effective even when performed in Arm 1. Previous studies that employed MI to improve HF patients' self-care reported only improvements at 12 and 8 weeks following enrolment. We were able to demonstrate that MI can have an impact on self-care up to a year after enrolment [47].

In the secondary analyses of the MOTIVATE-HF trial, the effect of MI on levels of anxiety, depression, sleep quality, quality of life and burden of physical symptoms on patients with HF was assessed. In the 12 months after the MI intervention, we discovered that MI had no impact on anxiety, depression, or sleep quality. When compared to the control arm, only disease-specific QOL and burden of physical symptoms improved with time in the intervention arms, although these improvements occurred from 9 months following the intervention [64].

According to earlier research, the smallest clinically significant change in the KCCQ score is 5 points [68,69]. Even with a power of 71% in the sample size, we saw a statistically significant improvement in KCCQ overall summary score, ranging from 6.73 to 8.41 points. This indicates that the disease-specific QOL was significantly impacted by this intervention. Physical limits, symptoms, QOL, and social constraints are all reflected in the KCCQ overall summary score, which tended to get better later on in the follow-up period in the group where both the patient and caregiver received the intervention. This development seems to be the result of assistance from the shared MI experience.

On the other hand, over a year, HFSPS total score showed a substantial improvement only when MI was given also to the caregiver (Arm 2). This result is not unexpected because, in the primary endpoint, we saw that the intervention had a stronger effect on self-care when it was administrated on both patients and caregivers. So, in order to reduce the burden of HF symptoms, clinicians are advised to incorporate informal caregivers in the care process and encourage them to provide patients with acceptable levels of self-care. From a scientific point of view, this result highlighted how critical it is to comprehend the relationship between self-care and the severity of HF symptoms [65]. Knowing how much self-care activities or other factors affect symptom burden is crucial because it is a subjective experience that is influenced by a variety of circumstances. With this knowledge, therapies for HF patients could be better tailored to their symptoms.

Additionally, physical symptoms are prognostic of survival, and reducing the severity of symptoms may decrease HF patient mortality. In fact, we demonstrated a decrease in patient mortality in Arms 1 and 2 compared to Arm 3. It is interesting to note that while improvement in symptom burden was seen at 9 and 12 months, mortality decrease was only noticed at 3 months following the intervention [67]. This suggests that while an increase in self-care may have an immediate impact on mortality, its impact on symptoms may take longer. Changes in self-care behaviours may be responsible for the substantial decline in mortality in HF patients at three months, but this result should be interpreted with care given the low number of events. The experimental arms (Arm 1 and 2) experienced lower mortality rates, but only in the group that received the intervention for both patients and caregivers this reduction reached statistical significance. This is important because it suggests, once again, that performing this intervention in dyads rather than patients alone may increase its effectiveness.

Regarding the analyses on caregivers, we discovered that a MI intervention was effective at increasing caregiver self-efficacy but not effective at increasing CC to self-care or caregiver preparedness [66]. The discovery that caregiver self-efficacy has increased as a result of MI is significant because it sheds light on how to increase this factor, which is a potent mediator between predictors that influence CC to self-care and CC to self-care itself [22]. It is interesting to note that the intervention improved caregiver self-efficacy 9 months following the intervention, but not directly after 3 months. This result agrees with a study showing an increase in caregiver self-efficacy 12 months following the intervention [70]. This consistency in the results is significant because it shows that interventions intended to increase caregiver self-efficacy should be assessed after a set amount of time has passed since the intervention, rather than soon after.

Furthermore, we found that the intervention, in which caregivers were guided to improve their support towards patient self-care, did not increase caregivers' levels of anxiety and depression and did not decrease their QOL and sleep. This can lead to different considerations regarding the results. First, it may mean that caregivers' levels of anxiety, depression, QOL and sleep did not significantly change because MI did not significantly improve CC to self-care. Second, it could be that the significant improvement of caregiver self-efficacy prevented these variables to worsen. Third, these results may mean that delivering interventions to caregivers aiming to improve their contribution to patient self-care does not cause a worsening in caregivers' levels of anxiety, depression, QOL and sleep.

The HF patient–caregiver dyadic analysis showed that, although the difference that was closest to being significant was in the improvement in the mutuality for love and affection for caregivers in the intervention arm only for patient respect to the standard care, MI did not show an impact on changes in the patients' and caregivers' general mutuality domains during the follow-up time. However, the living together condition was significantly associated with changes in the patients' mutuality dimensions (i.e., love and affection, shared values) and total score, while in caregivers it was associated only with the MS for shared values. Although no significant differences were observed in the three arms, this represents an important result. From a clinical point of view, clinicians and nurses should be encouraged not to use MI to improve the mutuality between patients and caregivers in order to manage the disease. Furthermore, knowing that the greatest effects have been observed in HF patients who live with their caregivers, clinicians should reflect on adopting specific interventions for patients who may be most at risk, such as those who live alone.

In this trial, a number of potential biases were found. Patient attrition was 20.4% 3 months following the enrolment and increased to 46.7% 12 months later. This attrition rate at 3 months was lower than that of comparable studies (33%), despite being high [15]. By the end of the study, when attrition was at its highest, we used a mixed model analysis to account for confounders associated with dropouts, which supported the effectiveness of MI on self-care maintenance and management, disease-specific QOL and burden of HF symptoms, in the intervention arms. MI quality scores lower than desired for the technical and relational components may also be another source of bias. Nevertheless, the MI intervention was effective. The improvement in self-care maintenance, management, and confidence would have been much better with a better intervention quality and higher dosage. Furthermore, a potential bias is that MI is a highly customized method that is challenging to standardize [47]. This because the interventionists let the patients choose which self-care behaviour to address during the intervention, in accordance with the method (e.g. physical activity and diet). Despite these limitations, the results of this study may have strong external validity because the patients were included from a variety of healthcare settings and the inclusion and exclusion criteria were quite lax.

6.2 Multilevel modelling in dyad context

The analysis of data with multiple levels can be done using a variety of techniques. However, there is still a lot of misunderstanding regarding the similarities and differences between these strategies. The use of linear mixed models, in contrast to the paired t-test applied at each time point, allowed to control for covariates, variation within and between group levels, as well as longitudinal measures. They also accounted for dropouts in the data: the paired t-test worked on complete data at each follow-up time, while the mixed model allowed to include all randomized subjects considering their trajectory in time. By the use of the model we also got an overall estimate of the effect of MI over the entire follow-up.

The use of mixed model was also very advantageous when studying mutuality, in fact its multilevel structure nicely fitted this dyadic outcome. Although mutuality was not affected by MI, the application showed various benefits of multilevel modelling over traditional dyadic analysis:

1. while controlling for the interdependency of the outcome scores, the models compared the average trajectory for both members to assess the relational character

of change within dyads. Caregivers displayed a slightly higher level of mutuality in almost all domains of the scale at baseline, and HF patients had a higher significant linear trend;

2. the models supported differential patterns of prediction across dyad members. For instance, for love and affection and total score domains, the living together condition (living with caregiver vs. living alone) was found to be a significant predictor for HF patient mutuality, but not for caregiver mutuality. Insight into important dyadic processes can be gained by exploring differential patterns of prediction using more complex models;
3. the model estimated cross-care member effects. For example, in the dropout analysis, the self-care confidence score of the HF patient at baseline was significantly associated with how the patient perceived the care relationship, but it did not influence the caregiver’s perception of the care relationship. Such cross-member effects help us comprehend how the dyad relationship evolves over time and how it affects the wellbeing of both members.

These methods have the limitation that they do not specifically address differences in care relationship roles. Since every member of the dyad is treated equally, the methods do not specifically model discrepancies between patient and caregiver [71]. These kinds of analyses are theoretically possible, but they are unlikely to occur in reality given the typical low number of members within a dyad and the variety of patterns presenting for assessment.

Multilevel modelling has significantly improved family researchers' capacity to look at the contextual effects of communities on individual development over the past decades. Both family science and clinical work have advanced the ability to model such nested data. In this direction, family research can be developed in several ways if multilevel modelling will become more widely used to investigate the context of both dyads and families. First, multilevel modelling will allow to model within-dyad (and within-family) variation and make new questions concerning within-family relationships. The dynamics of the family context over time and the various effects of individual roles within the same family can both be studied by researchers. Second, multilevel modelling can change how we assess interventions given to both individuals and families. To gain a deeper knowledge of the costs and benefits of individual- and family-targeted interventions, multilevel modelling enables to assess the effects on both levels. We may start to determine whether an intervention is beneficial for both the target individual and other family members by looking at how it affects a dyad or family. Individually targeted interventions could be advantageous to all members, disadvantageous to all or none, or have little to no effect on all or none. Additionally, we can look also whether the intervention has any effect on the family as a whole, such as reducing discrepancy among members [46].

Thinking about possible extensions of the model, under the suppositions that (i) family members have distinguishable roles and (ii) outcome data collected from each member are relational in nature (that is all individuals are answering to the same scale or survey items), the multivariate outcomes dyad models presented in this project can be extended to the situation of the triad or family. When the model is extended to families where the number of members varies among families, there could be issues with data sparsity because there are not many large families. In multilevel modelling, this reflects to the robustness of the model to large amounts of missing data. A strategy would be to sample chosen family members in order to guarantee that the same number is included for each family. Another extension could be the use of a three-level model to examine effects that are truly contextual, such as care dyads that are nested inside various healthcare organizations, communities, or clinics. Finally, these models can also incorporate time-varying covariates that reflect the association between changes in several features of health status and changes in mutuality over time [43,72].

Concluding, the MOTIVATE-HF trial on the efficacy of a motivational interview in improving self-care and other outcomes in patients with HF involves dyadic and familial relationships. When dealing with dyadic outcomes, such as mutuality, the multilevel approach has been shown to be useful to take into account the hierarchical composition of families and reveal variation both within and between families.

References

1. Ponikowski P, Voors AA, Anker SD, et al. 2016 ESC Guidelines for the diagnosis and treatment of acute and chronic heart failure: The Task Force for the diagnosis and treatment of acute and chronic heart failure of the European Society of Cardiology (ESC)Developed with the special contribution of the Heart Failure Association (HFA) of the ESC. *Eur Heart J*. 2016;37(27):2129-2200. doi:10.1093/eurheartj/ehw128

2. Mozaffarian D, Benjamin EJ, Go AS, et al. Heart Disease and Stroke Statistics—2015 Update: A Report From the American Heart Association. *Circulation*. 2015;131(4). doi:10.1161/CIR.0000000000000152

3. Bekelman DB, Rumsfeld JS, Havranek EP, et al. Symptom Burden, Depression, and Spiritual Well-Being: A Comparison of Heart Failure and Advanced Cancer Patients. *J Gen Intern Med*. 2009;24(5):592-598. doi:10.1007/s11606-009-0931-y

4. Davidson PM, Abernethy AP, Newton PJ, Clark K, Currow DC. The caregiving perspective in heart failure: a population based study. *BMC Health Serv Res*. 2013;13(1):342. doi:10.1186/1472-6963-13-342

5. Smith CE, Piamjariyakul U, Wick JA, et al. Multidisciplinary Group Clinic Appointments: The Self-Management and Care of Heart Failure (SMAC-HF) Trial. *Circ Heart Fail*. 2014;7(6):888-894. doi:10.1161/CIRCHEARTFAILURE.113.001246

6. Riegel B, Dickson VV, Faulkner KM. The Situation-Specific Theory of Heart Failure Self-Care: Revised and Updated. *J Cardiovasc Nurs*. 2016;31(3):226-235. doi:10.1097/JCN.0000000000000244

7. Freedland KE, Carney RM, Rich MW, Steinmeyer BC, Rubin EH. Cognitive Behavior Therapy for Depression and Self-Care in Heart Failure Patients: A Randomized Clinical Trial. *JAMA Intern Med*. 2015;175(11):1773. doi:10.1001/jamainternmed.2015.5220

8. Dracup K, Moser DK, Pelter MM, et al. Randomized, Controlled Trial to Improve Self-Care in Patients With Heart Failure Living in Rural Areas. *Circulation*. 2014;130(3):256-264. doi:10.1161/CIRCULATIONAHA.113.003542

9. Jonkman NH, Westland H, Groenwold RHH, et al. What Are Effective Program Characteristics of Self-Management Interventions in Patients With Heart Failure? An Individual Patient Data Meta-analysis. *J Card Fail*. 2016;22(11):861-871. doi:10.1016/j.cardfail.2016.06.422

10. Clark AM, Wiens KS, Banner D, et al. A systematic review of the main mechanisms of heart failure disease management interventions. *Heart*. 2016;102(9):707-711. doi:10.1136/heartjnl-2015-308551

11. Miller WR, Rollnick S. Ten Things that Motivational Interviewing Is Not. *Behav Cogn Psychother*. 2009;37(2):129-140. doi:10.1017/S1352465809005128

12. Palacio A, Garay D, Langer B, Taylor J, Wood BA, Tamariz L. Motivational Interviewing Improves Medication Adherence: a Systematic Review and Meta-analysis. *J Gen Intern Med*. 2016;31(8):929-940. doi:10.1007/s11606-016-3685-3

13. Ekong G, Kavookjian J. Motivational interviewing and outcomes in adults with type 2 diabetes: A systematic review. *Patient Educ Couns*. 2016;99(6):944-952. doi:10.1016/j.pec.2015.11.022

14. Lee WWM, Choi KC, Yum RWY, Yu DSF, Chair SY. Effectiveness of motivational interviewing on lifestyle modification and health outcomes of clients at risk or diagnosed with cardiovascular diseases: A systematic review. *Int J Nurs Stud*. 2016;53:331-341. doi:10.1016/j.ijnurstu.2015.09.010

15. Masterson Creber R, Patey M, Lee CS, Kuan A, Jurgens C, Riegel B. Motivational interviewing to improve self-care for patients with chronic heart failure: MITI-HF randomized controlled trial. *Patient Educ Couns*. 2016;99(2):256-264. doi:10.1016/j.pec.2015.08.031

16. Riegel B, Masterson Creber R, Hill J, Chittams J, Hoke L. Effectiveness of Motivational Interviewing in Decreasing Hospital Readmission in Adults With Heart Failure and Multimorbidity. *Clin Nurs Res*. 2016;25(4):362-377. doi:10.1177/1054773815623252

17. Paradis V, Cossette S, Frasure-Smith N, Heppell S, Guertin MC. The Efficacy of a Motivational Nursing Intervention Based on the Stages of Change on Self-care in Heart Failure Patients. *J Cardiovasc Nurs*. 2010;25(2):130-141. doi:10.1097/JCN.0b013e3181c52497

18. Brodie DA, Inoue A, Shaw DG. Motivational interviewing to change quality of life for people with chronic heart failure: A randomised controlled trial. *Int J Nurs Stud*. 2008;45(4):489-500. doi:10.1016/j.ijnurstu.2006.11.009

19. Riegel B, Dickson VV, Garcia LE, Masterson Creber R, Streur M. Mechanisms of change in self-care in adults with heart failure receiving a tailored, motivational interviewing intervention. *Patient Educ Couns*. 2017;100(2):283-288. doi:10.1016/j.pec.2016.08.030

20. Cameron J, Worrall-Carter L, Page K, Riegel B, Lo SK, Stewart S. Does cognitive impairment predict poor self-care in patients with heart failure? *Eur J Heart Fail*. 2010;12(5):508-515. doi:10.1093/eurjhf/hfq042

21. Vellone E, Riegel B, Cocchieri A, et al. Validity and Reliability of the Caregiver Contribution to Self-care of Heart Failure Index. *J Cardiovasc Nurs*. 2013;28(3):245-255. doi:10.1097/JCN.0b013e318256385e

22. Vellone E, Riegel B, Alvaro R. A Situation-Specific Theory of Caregiver Contributions to Heart Failure Self-care. *J Cardiovasc Nurs*. 2019;34(2):166-173. doi:10.1097/JCN.0000000000000549

23. Cairo Notari S, Favez N, Notari L, Charvoz L, Delaloye JF. The caregiver burden in male romantic partners of women with non-metastatic breast cancer: The protective role of couple satisfaction. *J Health Psychol*. 2017;22(13):1668-1677. doi:10.1177/1359105316633285

24. Stamp KD, Dunbar SB, Clark PC, et al. Family partner intervention influences self-care confidence and treatment self-regulation in patients with heart failure. *Eur J Cardiovasc Nurs*. 2016;15(5):317-327. doi:10.1177/1474515115572047

25. Bidwell JT, Vellone E, Lyons KS, et al. Caregiver determinants of patient clinical event risk in heart failure. *Eur J Cardiovasc Nurs*. 2017;16(8):707-714. doi:10.1177/1474515117711305

26. Vellone E, Barbaranelli C, Pucciarelli G, Zeffiro V, Alvaro R, Riegel B. Validity and Reliability of the Caregiver Contribution to Self-Care of Heart Failure Index Version 2. *J Cardiovasc Nurs*. 2020;35(3):280-290. doi:10.1097/JCN.0000000000000655

27. Carter RE, Carter CA. Marital adjustment and effects of illness in married pairs with one or both spouses chronically ill. *Am J Fam Ther*. 1994;22(4):315-326. doi:10.1080/01926189408251324

28. Druley JA, Stephens MAP, Coyne JC. Emotional and physical intimacy in coping with lupus: Women’s dilemmas of disclosure and approach. *Health Psychol*. 1997;16(6):506-514. doi:10.1037/0278-6133.16.6.506

29. Schulz R, Beach SR. Caregiving as a Risk Factor for Mortality: The Caregiver Health Effects Study. *JAMA*. 1999;282(23):2215. doi:10.1001/jama.282.23.2215

30. Lawrence RH, Tennstedt SL, Assmann SF. Quality of the caregiver–care recipient relationship: Does it offset negative consequences of caregiving for family caregivers? *Psychol Aging*. 1998;13(1):150-158. doi:10.1037/0882-7974.13.1.150

31. Spruytte N, Van Audenhove C, Lammertyn F. Predictors of institutionalization of cognitively-impaired elderly cared for by their relatives. *Int J Geriatr Psychiatry*. 2001;16(12):1119-1128. doi:10.1002/gps.484

32. Sterling MR, Barbaranelli C, Riegel B, et al. The Influence of Preparedness, Mutuality, and Self-efficacy on Home Care Workers’ Contribution to Self-care in Heart Failure: A Structural Equation Modeling Analysis. *J Cardiovasc Nurs*. 2022;37(2):146-157. doi:10.1097/JCN.0000000000000768

33. Vellone E, Chung ML, Alvaro R, Paturzo M, Dellafiore F. The Influence of Mutuality on Self-Care in Heart Failure Patients and Caregivers: A Dyadic Analysis. *J Fam Nurs*. 2018;24(4):563-584. doi:10.1177/1074840718809484

34. Hooker SA, Schmiege SJ, Trivedi RB, Amoyal NR, Bekelman DB. Mutuality and heart failure self-care in patients and their informal caregivers. *Eur J Cardiovasc Nurs*. 2018;17(2):102-113. doi:10.1177/1474515117730184

35. Rohrbaugh MJ, Shoham V, Coyne JC. Effect of Marital Quality on Eight-Year Survival of Patients With Heart Failure. *Am J Cardiol*. 2006;98(8):1069-1072. doi:10.1016/j.amjcard.2006.05.034

36. Dellafiore F, Chung ML, Alvaro R, et al. The Association Between Mutuality, Anxiety, and Depression in Heart Failure Patient-Caregiver Dyads: An Actor-Partner Interdependence Model Analysis. *J Cardiovasc Nurs*. 2019;34(6):465-473. doi:10.1097/JCN.0000000000000599

37. Cooney TM, Proulx CM, Bekelman DB. Changes in Social Support and Relational Mutuality as Moderators in the Association Between Heart Failure Patient Functioning and Caregiver Burden. *J Cardiovasc Nurs*. 2021;36(3):212-220. doi:10.1097/JCN.0000000000000726

38. Lyons KS, Lee CS. The Theory of Dyadic Illness Management. *J Fam Nurs*. 2018;24(1):8-28. doi:10.1177/1074840717745669

39. Bookwala J, Schulz R. Spousal similarity in subjective well-being: The cardiovascular health study. *Psychol Aging*. 1996;11(4):582-590. doi:10.1037/0882-7974.11.4.582

40. Lyons KS, Zarit SH, Sayer AG, Whitlatch CJ. Caregiving as a Dyadic Process: Perspectives From Caregiver and Receiver. *J Gerontol B Psychol Sci Soc Sci*. 2002;57(3):P195-P204. doi:10.1093/geronb/57.3.P195

41. Kenny DA, Kashy DA. The Design and Analysis of Data from Dyads and Groups. In: Reis HT, Judd CM, eds. *Handbook of Research Methods in Social and Personality Psychology*. 2nd ed. Cambridge University Press; 2014:589-607. doi:10.1017/CBO9780511996481.027

42. Cook WL, Kenny DA. The Actor–Partner Interdependence Model: A model of bidirectional effects in developmental studies. *Int J Behav Dev*. 2005;29(2):101-109. doi:10.1080/01650250444000405

43. Raudenbush SW, Brennan RT, Barnett RC. A multivariate hierarchical model for studying psychological change within married couples. *J Fam Psychol*. 1995;9(2):161-174. doi:10.1037/0893-3200.9.2.161

44. Barnett RC, Marshall NL, Raudenbush SW, Brennan RT. Gender and the relationship between job experiences and psychological distress: A study of dual-earner couples. *J Pers Soc Psychol*. 1993;64(5):794-806. doi:10.1037/0022-3514.64.5.794

45. Kenny DA, Kashy DA, Cook WL. *Dyadic Data Analysis*. Paperback edition. The Guilford Press; 2020.

46. Lyons KS, Sayer AG. Longitudinal Dyad Models in Family Research. *J Marriage Fam*. 2005;67(4):1048-1060. doi:10.1111/j.1741-3737.2005.00193.x

47. Vellone E, Paturzo M, D’Agostino F, et al. MOTIVATional intErviewing to improve self-care in Heart Failure patients (MOTIVATE-HF): Study protocol of a three-arm multicenter randomized controlled trial. *Contemp Clin Trials*. 2017;55:34-38. doi:10.1016/j.cct.2017.02.003

48. Callahan CM, Unverzagt FW, Hui SL, Perkins AJ, Hendrie HC. Six-Item Screener to Identify Cognitive Impairment Among Potential Subjects for Clinical Research: *Med Care*. 2002;40(9):771-781. doi:10.1097/00005650-200209000-00007

49. Vellone E, Riegel B, Cocchieri A, et al. Psychometric testing of the self-care of heart failure index version 6.2. *Res Nurs Health*. 2013;36(5):500-511. doi:10.1002/nur.21554

50. Jurgens CY, Lee CS, Riegel B. Psychometric Analysis of the Heart Failure Somatic Perception Scale as a Measure of Patient Symptom Perception. *J Cardiovasc Nurs*. 2017;32(2):140-147. doi:10.1097/JCN.0000000000000320

51. Ware J, Kosinski M, Keller SD. A 12-Item Short-Form Health Survey: construction of scales and preliminary tests of reliability and validity. *Med Care*. 1996;34(3):220-233. doi:10.1097/00005650-199603000-00003

52. Green CP, Porter CB, Bresnahan DR, Spertus JA. Development and evaluation of the Kansas City Cardiomyopathy Questionnaire: a new health status measure for heart failure. *J Am Coll Cardiol*. 2000;35(5):1245-1255. doi:10.1016/S0735-1097(00)00531-3

53. Zigmond AS, Snaith RP. The Hospital Anxiety and Depression Scale. *Acta Psychiatr Scand*. 1983;67(6):361-370. doi:10.1111/j.1600-0447.1983.tb09716.x

54. Buysse DJ, Reynolds CF, Monk TH, Berman SR, Kupfer DJ. The Pittsburgh sleep quality index: A new instrument for psychiatric practice and research. *Psychiatry Res*. 1989;28(2):193-213. doi:10.1016/0165-1781(89)90047-4

55. Dellafiore F, Buck HG, Pucciarelli G, et al. Psychometric characteristics of the mutuality scale in heart failure patients and caregivers. *Heart Lung*. 2018;47(6):553-561. doi:10.1016/j.hrtlng.2018.05.018

56. Petruzzo A, Paturzo M, Buck HG, et al. Psychometric evaluation of the Caregiver Preparedness Scale in caregivers of adults with heart failure. *Res Nurs Health*. 2017;40(5):470-478. doi:10.1002/nur.21811

57. Vellone E, Rebora P, Ausili D, et al. Motivational interviewing to improve self-care in heart failure patients (MOTIVATE-HF): a randomized controlled trial. *ESC Heart Fail*. 2020;7(3):1309-1318. doi:10.1002/ehf2.12733

58. Fitzmaurice GM, Laird NM, Ware JH. *Applied Longitudinal Analysis*. John Wiley & Sons, Inc.; 2011. doi:10.1002/9781119513469

59. Hogan JW, Roy J, Korkontzelou C. Handling drop-out in longitudinal studies. *Stat Med*. 2004;23(9):1455-1497. doi:10.1002/sim.1728

60. *EuroQol Research Foundation. EQ-5D-3L User Guide, 2018. Available from: Https://Euroqol.Org/Publications/User-Guides.*

61. Rubin DB, Frangakis CE. Estimation of the Causal Effect of a Time-Varying Exposure on the Marginal Mean of a Repeated Binary Outcome: Comment. *J Am Stat Assoc*. 1999;94(447):702. doi:10.2307/2669980

62. Robins JM, Greenland S, Hu FC. Estimation of the Causal Effect of a Time-Varying Exposure on the Marginal Mean of a Repeated Binary Outcome: Rejoinder. *J Am Stat Assoc*. 1999;94(447):708. doi:10.2307/2669983

63. Raudenbush SW, Bryk AS. *Hierarchical Linear Models: Applications and Data Analysis Methods*. 2. ed., [Nachdr.]. Sage Publ; 2010.

64. Rebora P, Spedale V, Occhino G, et al. Effectiveness of motivational interviewing on anxiety, depression, sleep quality and quality of life in heart failure patients: secondary analysis of the MOTIVATE-HF randomized controlled trial. *Qual Life Res*. 2021;30(7):1939-1949. doi:10.1007/s11136-021-02788-3

65. Caggianelli G, Iovino P, Rebora P, et al. A Motivational Interviewing Intervention Improves Physical Symptoms in Patients with Heart Failure: A Secondary Outcome Analysis of the Motivate-HF Randomized Controlled Trial. *J Pain Symptom Manage*. 2022;63(2):221-229.e1. doi:10.1016/j.jpainsymman.2021.09.006

66. Locatelli G, Zeffiro V, Occhino G, et al. Effectiveness of Motivational Interviewing on contribution to self-care, self-efficacy, and preparedness in caregivers of patients with

heart failure: a secondary outcome analysis of the MOTIVATE-HF randomized controlled trial. *Eur J Cardiovasc Nurs*. Published online March 15, 2022:zvac013. doi:10.1093/eurjcn/zvac013

67. Iovino P, Rebora P, Occhino G, et al. Effectiveness of motivational interviewing on health-service use and mortality: a secondary outcome analysis of the MOTIVATE-HF trial. *ESC Heart Fail*. 2021;8(4):2920-2927. doi:10.1002/ehf2.13373

68. Butler J, Khan MS, Mori C, et al. Minimal clinically important difference in quality of life scores for patients with heart failure and reduced ejection fraction. *Eur J Heart Fail*. 2020;22(6):999-1005. doi:10.1002/ejhf.1810

69. Spertus J, Peterson E, Conard MW, et al. Monitoring clinical changes in patients with heart failure: A comparison of methods. *Am Heart J*. 2005;150(4):707-715. doi:10.1016/j.ahj.2004.12.010

70. Wingham J, Frost J, Britten N, et al. Caregiver outcomes of the REACH-HF multicentre randomized controlled trial of home-based rehabilitation for heart failure with reduced ejection fraction. *Eur J Cardiovasc Nurs*. 2019;18(7):611-620. doi:10.1177/1474515119850011

71. Feaster D, Brincks A, Robbins M, Szapocznik J. Multilevel Models to Identify Contextual Effects on Individual Group Member Outcomes: A Family Example. *Fam Process*. 2011;50(2):167-183. doi:10.1111/j.1545-5300.2011.01353.x

72. Barnett RC, Raudenbush SW, Brennan RT, Pleck JH, Marshall NL. Change in job and marital experiences and change in psychological distress: A longitudinal study of dual-earner couples. *J Pers Soc Psychol*. 1995;69(5):839-850. doi:10.1037/0022-3514.69.5.839