Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Beyond Symptom Reduction: A Research Protocol for a Family Psychoeducational Intervention Focused on Personal Recovery and Empowerment in Major Depressive Disorder

Salvatore Iuso

salvatore.iuso@unifg.it

University of Foggia Department of Humanities Arts Cultural Heritage Education Sciences: Universita degli Studi di Foggia Dipartimento di Studi Umanistici Lettere Beni Culturali Scienze della Formazione

https://orcid.org/0000-0002-9372-1899

Trotta Nicoletta

Fondazione IRCCS Ca' Granda Ospedale Maggiore Policlinico

Antonio Ventriglio

University of Foggia

Nadia Genzano

University of Fogg

Adriana Leccese

University of Foggia

Annamaria Petito

University of Foggia

Research Article

Keywords: Family Psychoeducation, Major Depressive Disorder, pilot randomized controlled trial, personal recovery, Expressed Emotion

Posted Date: November 3rd, 2026

DOI: https://doi.org/10.21203/rs.3.rs-11595515/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: The authors declare no competing interests.

Abstract

Background

Family psychoeducation shows promise in major depressive disorder (MDD), but evidence is limited and heterogeneous, and person-centred outcomes such as personal recovery and empowerment have rarely been examined.

Objective

To explore, in a pilot randomized controlled trial, the effects of a Falloon-based Family Psychoeducation Intervention (FPI), compared with an active family healthy-lifestyle program of equal intensity, on personal recovery, empowerment, expressed emotion, symptoms, functioning and quality of life.

Methods

Seventeen families of clinically stable outpatients with MDD (Foggia, Italy) were randomized to FPI (n = 8) or control (n = 9); both programs (12–14 sessions over about six months) were adjunct to usual care. Assessments took place pre-treatment (T0), post-intervention (T1) and at six-month follow-up (T2). Data were analyzed with nonparametric ANOVA-type statistics (nparLD) and Holm-adjusted post hoc tests.

Results

Significant group × time interactions emerged for the Recovery Assessment Scale total score (ATS = 9.28, $p < .001$; higher in FPI at T1 and T2, adjusted $p = .010$) and goal-orientation subscale ($p = .039$), and for family-reported expressed emotion (ATS = 3.22, $p = .041$), which decreased progressively in the FPI group. No significant interactions were found for depressive and anxiety symptoms, functioning, quality of life, medication adherence, internalized stigma or empowerment (whose scale showed poor internal consistency). Several outcomes improved over time in both arms.

Conclusions

In this small pilot, FPI appeared to influence personal recovery and family emotional climate rather than symptoms, whose improvement was shared by both conditions, consistent with non-specific factors. Findings are exploratory (no a priori power calculation or correction across outcomes) and need confirmation in adequately powered multicenter trials with longer follow-up.

INTRODUCTION

Major Depressive Disorder (MDD) is a clinical condition associated with significant distress, impaired personal and social functioning, and a non-negligible risk of persistence or recurrence. MDD can alter family dynamics, the distribution of responsibilities, communication and mutual expectations, making the family unit both a potential source of support and an environment exposed to prolonged stress (Katsuki et al., 2022). More recent meta-analyses have confirmed the findings of a now-outdated longitudinal study conducted by Hayhurst et al. (1997), which highlighted an association between persistent criticism from partners and the presence of residual depressive symptoms, while also emphasizing the difficulty of establishing a clear causal relationship. In this context, providing family members with information about the disorder and tools to understand its manifestations can be a significant component of a therapeutic approach that takes into account not only the patient's symptoms, but also the relational context in which care takes place. Data collected on specific diagnostic populations have increased interest in manual-based family interventions. However, most studies focus on schizophrenia and related disorders, and findings from one population are not automatically transferable to another (Sarkhel et al., 2020). The most recent evidence suggests possible benefits of family psychoeducation on depressive symptoms, although the available literature remains limited and the interventions studied are heterogeneous, making it impossible to draw definitive conclusions about their effectiveness or their effects on family outcomes (Obieche et al., 2025).

The systematic review by Brady, Kangas and McGill (2017) – the first to focus specifically on family psychoeducation in MDD – identified promising results regarding patients' functioning and carers' well-being, whilst noting the heterogeneity of the studies and the need to strengthen the empirical evidence base. The subsequent meta-analysis by Katsuki and colleagues (2022), based on five RCTs involving a total of 301 patients and family members, found a statistically significant effect on the severity of depressive symptoms, both at 16 weeks and at the final follow-up, whilst the effects on family functioning did not reach statistical significance. A randomized trial of brief multi-family psychoeducation for family members of patients with chronic DDM also revealed a mixed picture: no significant effect on the primary outcome relating to family members' distress, whilst some secondary analyses suggested possible improvements in family members' depressive symptoms and in specific dimensions of family functioning (Katsuki et al., 2018). These findings suggest that the intervention should not be described as uniformly effective across all outcomes, but rather that consideration should be given to which dimensions are most responsive to change and under what conditions.

A construct frequently discussed in the literature on family interventions is that of 'expressed emotionality', which encompasses relational dimensions such as criticism, hostility and excessive emotional involvement. Expressed emotionality should not be interpreted as an immutable characteristic of the family nor as an indicator of causal responsibility: rather, it is a dimension of the interpersonal climate that can be influenced by distress, available resources and communication styles (Fahrer et al., 2022). In the presence of a MDD, criticism may also stem from frustration, concern or difficulty in understanding the symptoms; similarly, very intense emotional involvement may reflect an attempt to protect a loved one. A psychoeducational intervention can help individuals recognize these processes and put into practice less critical and more effective communication strategies (Girdhar et al., 2024). In

practical terms, a psychoeducational intervention may include information on symptoms, risk and protective factors, and treatment options; recognizing early warning signs of a deterioration in condition; support for continuing treatment; communication and problem-management strategies; and attention to the needs of family members (Sarkhel et al., 2020; Di Lorenzo et al., 2024). The informative component becomes particularly valuable when linked to the specific situations described by participants, thereby avoiding a purely didactic approach. From a recovery-oriented perspective, psychoeducation should not be viewed as a process through which professionals transfer knowledge unidirectionally to patients and family members. It is preferable to see it as a space for the shared construction of knowledge and skills, in which participants' lived experiences are acknowledged and brought into dialogue with clinical information (Guarino et al., 2024; Neckar & Boxall, 2026). This approach supports the patient's self-determination and, as regards family members, helps them to recognize their own scope for intervention without taking over the person's decisions, whilst respecting their autonomy, preferences and boundaries (Stawnychy et al., 2021).

In recent decades, mental health services have complemented the paradigm centred on symptom reduction with an approach centred on the individual, on participation and on the possibility of building a meaningful life even in the face of persistent difficulties. Within this framework, personal recovery is neither synonymous with clinical cure nor with the absence of symptoms: it is a subjective, non-linear and individual process through which a person reconstructs meanings, goals, relationships and an identity that is not entirely defined by their illness (Leamy et al., 2011; Boardman, 2010; van Weeghel et al., 2019). It is useful to distinguish – whilst acknowledging their interdependence – between clinical recovery, which refers to the course of symptoms; functional recovery, which relates to the performance of daily activities and the maintenance of social roles; and personal recovery, which concerns the ability to regain agency, hope and direction in one's life (Rossi et al., 2022).). In MDD, a loss of hope, low self-esteem, social withdrawal and difficulty in envisaging future goals can hinder active participation in treatment and the recovery of meaningful roles (Yavari et al., 2023). A recovery-oriented approach acknowledges difficulties without equating the person with them, values existing resources and supports goals consistent with individual preferences: the role of professionals is to create conditions that facilitate informed choices and participation, rather than imposing a single model of a fulfilling life from the outside.

Empowerment, which is closely linked to this perspective, involves strengthening the ability to understand one's own situation, participate in decision-making, recognize one's own resources and act consciously in pursuit of one's goals. It does not equate to absolute self-sufficiency nor to shifting the entire burden of managing the illness onto the patient: it implies a collaborative care relationship, in which accessible information, social support and opportunities for choice foster active participation (Hormazábal-Salgado et al., 2024; Saavedra et al., 2023). Personal recovery and empowerment are therefore particularly relevant outcomes for evaluating psychosocial interventions that aim to address not only symptoms, but also the way in which a person experiences and copes with their condition.

The experience of caregiving can also involve practical and emotional burdens, concern for the future, a reduction in personal time and difficulties in maintaining one's social roles. The well-being of family members is therefore a significant outcome in its own right, as well as a potential factor facilitating the continuity of support and the overall effectiveness of treatment pathways. The evaluation of family-based interventions should therefore include not only the patient's clinical indicators, but also the relational and subjective aspects affecting those involved in the care pathway.

Despite the clear relevance of these theoretical and clinical considerations, the current literature has not yet systematically explored the impact of manualized family psychoeducational interventions for MDD using person-centred constructs, such as personal recovery and empowerment, as primary outcomes. To address this methodological gap, this protocol describes the design and methodology of a two-arm, parallel-group randomized controlled trial (RCT) aimed at evaluating the efficacy of a structured family psychoeducational intervention.

The primary objective of the study is to determine whether the intervention is capable of significantly promoting patients' personal recovery and empowerment, whilst simultaneously improving the well-being and quality of life of their carers. Secondary objectives include assessing changes in the severity of clinical symptoms of MDD, in overall functioning and in the modulation of levels of Expressed Emotionality within the family unit, whilst verifying their stability at a six-month follow-up after the end-of-intervention assessment. It is expected that a psychoeducational approach based on co-construction and oriented towards recovery will increase the agency of the person receiving care and reduce stress and criticism within the care-giving relationship.

METHODS

Procedure

This study was designed as a pilot randomized controlled trial. A total of 17 families of patients diagnosed with major depressive disorder (MDD) in a clinically stable phase were voluntarily recruited at the University Psychiatric Outpatient Unit of the Ospedale Riuniti in Foggia, Italy, between April 2023 and September 2024. The diagnosis of MDD was confirmed through psychiatric clinical assessment and evaluation of depressive symptom severity using the Hamilton Depression Rating Scale (HAM-D), with scores above 7 considered indicative of the presence of depressive symptoms. Clinical stability at baseline was established on the basis of the psychiatrist's assessment, considering the patients' clinical condition. All participants provided written informed consent, and the study protocol was approved by the Ethics Committee on March 10, 2023.

Randomization was performed using an online platform, with an allocation sequence generated independently and concealed from the researchers responsible for participant recruitment. Group allocation was therefore unknown and unpredictable to the recruiters during enrollment and could not be subsequently modified. Through this procedure, eight families were assigned to the Family

Psychoeducation Intervention (FPI; experimental group), while nine families were allocated to the control group, which received an intervention focused on promoting a healthy lifestyle. Both groups continued to receive their usual pharmacological treatment and routine outpatient psychiatric monitoring throughout the study. The interventions specified in the study protocol were therefore delivered as an adjunct to usual care and did not replace it. All patients participated in their assigned intervention together with at least one family member, who also underwent pre- and post-treatment assessments.

Exclusion criteria included: moderate or severe cognitive impairment; physical conditions, in patients or family members, that would prevent participation in the sessions; the presence, in family members, of a severe mental disorder (schizophrenia or other psychotic disorders, bipolar disorder, major depressive disorder); hospitalization or significant clinical worsening in the two months prior to enrollment, such that it required a substantial change in medication or intensive crisis management interventions.

Given the small sample size, the study should be considered a preliminary feasibility study aimed at estimating the effect size and identifying protocol challenges, rather than generalizing the results.

The experimental group received the Family Psychoeducation Intervention based on Falloon's model, administered individually to each family and structured into 12–14 sessions spread over approximately six months, initially on a weekly basis and subsequently every two weeks. The program included the phases of family intake, individual and family assessment, informational sessions, and sessions dedicated to communication skills and problem-solving aimed at achieving shared goals. Each module was based on structured informational materials and practical exercises, with the goal of making patients and their family members active participants in managing the disorder.

The control group received a family-based program promoting a healthy lifestyle, of equal intensity and duration, and with the same meeting format. The program took a psychoeducational approach to the main behavioral determinants of physical and mental well-being, as well as adherence to medication, with a particular focus on recognizing and managing side effects. The sessions were conducted interactively and included self-monitoring tools, with the goal of promoting sustainable habits and an active role for the patient and family in the treatment process.

Measures

The study protocol adopted a longitudinal design comprising three assessment time points: pre-treatment (T0), post-intervention (T1), and a six-month follow-up (T2). The assessment battery included measures of clinical symptoms, psychosocial functioning, and subjective quality of life, alongside instruments specifically addressing recovery, empowerment, and self-stigma. The LEE and MSAQL scales were administered to both patients and family members.

Hamilton Anxiety Rating Scale (HAM-A). The severity of anxiety symptoms was quantified using the Hamilton Anxiety Rating Scale (HAM-A; Hamilton, 1959), a 14-item scale administered by qualified clinicians via a semistructured interview regarding the previous week (Maier et al., 1988). The instrument

assesses both the psychological anxiety dimension (items 1–6 and 14) and the somatic anxiety dimension (items 7–13) (Manzar et al., 2025), with individual item scores ranging from 0 to 4 and a total score ranging from 0 to 56. Standard cut-off scores classify the severity of anxiety as mild (< 17), mild to moderate (18–24), and moderate to severe (25–30), based on robust psychometric properties of validity and reliability [Yang, 2020]. In the present sample, Cronbach's alpha at T0 was 0.74.

Hamilton Depression Rating Scale (HAM-D). The severity of depressive symptoms was assessed using the 17-item version of the HAM-D, administered by experienced clinicians during semistructured interviews (Hamilton, 1967). The instrument assesses areas such as depressed mood, suicidal ideation, insight, and somatic symptoms. The total score (range 0–52) classifies severity as mild (8–16), moderate (17–23), and severe (≥ 24). The Italian version has demonstrated excellent psychometric properties (Carrozzino et al., 2020); in the present sample, Cronbach's alpha at T0 was 0.69.

Morisky Medication Adherence Scale (MMAS-4). The Morisky Medication Adherence Scale (MMAS-4) is a self-report questionnaire consisting of four dichotomous (yes/no) questions designed to identify behavioral barriers to medication adherence (Morisky et al., 1986). The instrument is widely used as a screening test in primary care due to its simplicity and its ability to highlight specific mechanisms of nonadherence (Tan et al., 2014). In the present sample, Cronbach's alpha at T0 was 0.63.

Recovery Assessment Scale (RAS). The recovery process and levels of empowerment and self-determination were assessed using the Italian version of the Recovery Assessment Scale (RAS; Boggian et al., 2020). This instrument, designed specifically for individuals with severe mental disorders, is a self-report questionnaire consisting of 41 items on a five-point Likert scale (from 1 = "Strongly Disagree" to 5 = "Strongly Agree," with a "not applicable" option). The total score ranges from 41 to 205, with the option to calculate scores for the five factor dimensions: self-confidence and hope, willingness to seek help, goal orientation and success, trust in others, and not feeling overwhelmed by symptoms. Participants completed the questionnaire with the support of a staff member to clarify any interpretive doubts, ensuring the reliability of the measurement. In the present sample, Cronbach's alpha at T0 was 0.88.

Level of Expressed Emotion Scale (LEE). The instrument is a 60-item self-report questionnaire developed according to Vaughn and Leff's theory of expressed emotion (Di Paola et al., 2008). It assesses four dimensions, each comprising 15 items: intrusiveness, emotional response to the illness, emotional attitude, and tolerance/expectations. Items are rated on a true/false scale, with 1 assigned to clinically significant behaviors or attitudes and 0 otherwise. Higher scores indicate greater impairment in the perceived emotional climate of the relationship. The questionnaire requires no specific training for administration and does not rely on the identification of a "key" family member. In Italy, it is used in epidemiological research, psychoeducational interventions, and screening as a practical and reliable measure of the perceived emotional climate within significant interpersonal relationships (Di Paola et al., 2008). In the present sample, Cronbach's alpha at T0 was 0.73.

Netherlands Empowerment List (NEL). Levels of empowerment and personal and relational resources were measured using the Netherlands Empowerment List (NEL; Boevink et al., 2017). The self-report

instrument consists of 40 items rated on a five-point Likert scale (from 1 = “Strongly disagree” to 5 = “Strongly agree”). Four items related to care services include an additional “not applicable” option. The scale examines six multidimensional factors: trust and purpose in life (13 items), social support (10 items), caring community (4 items), feeling connected (4 items), self-management (5 items), and professional help (4 items) (Boevink et al., 2017). Higher scores reflect a greater degree of empowerment and social integration among participants. In the present sample, Cronbach’s alpha at T0 was 0.14.

Personal and Social Performance Scale (PSP): Personal and social functioning over the four weeks preceding each assessment was evaluated using the Personal and Social Performance Scale (PSP; White et al., 2016). The scale, administered by specifically trained clinicians, assesses patients’ autonomy and adaptive functioning across four main domains: socially useful activities, personal and social relationships, self-care, and disturbing or aggressive behaviors (Vauth et al., 2021). The instrument provides a global score and domain-specific ratings and has demonstrated high sensitivity to changes in clinical status and everyday functioning (White et al., 2016). In the present sample, Cronbach’s alpha for the PSP was 0.75 at T0, indicating acceptable internal consistency at baseline but lower internal consistency at subsequent assessments.

Manchester Short Assessment of Quality of Life (MSAQL). Subjective quality of life was assessed using the Manchester Short Assessment of Quality of Life (MANSA; Petkari et al., 2020), a brief instrument derived from the Lancashire Quality of Life Profile (LQoLP) and administered to both patients and family members. The questionnaire includes 16 items, 12 of which measure satisfaction in specific domains of daily life (including physical and mental health, work, finances, social and family relationships, and safety) on a seven-point Likert scale (from 1 = “It couldn’t be worse than this” to 7 = “It couldn’t be better than this”). Data analysis is based on a two-factor structure that distinguishes “Life and Health Aspects” (8 items) from “Quality of the Living Environment” (4 items), ensuring a rapid (approximately 15 minutes) and psychometrically sound assessment of the disorder’s impact on perceived well-being (Petkari et al., 2020). In the present sample, Cronbach’s alpha at T0 was 0.71.

Internalized Stigma of Mental Illness Inventory (ISMI). A 29-item self-assessment tool for self-stigma in people with psychiatric disorders, developed using a participatory approach that directly involved people with lived experience of mental illness and based on a multidimensional model of the subjective experience of stigma (Boyd et al., 2014). It comprises five subscales: alienation (the feeling of being devalued or excluded as members of society), endorsement of stereotypes (agreement with negative stereotypes about mental illness), perceived discrimination (perception of discriminatory treatment due to others’ prejudices), social withdrawal (avoidance of interactions related to one’s mental illness), and resistance to stigma (perceived ability to cope with and counteract it). The items are rated on a four-point Likert scale, and the overall mean score quantifies the level of self-stigma, with higher values indicating greater internalization. The original validation study demonstrated high internal consistency and satisfactory test-retest reliability; construct validity is supported by positive correlations with stigmatizing beliefs and depressive symptoms, and negative correlations with self-esteem,

empowerment, and recovery orientation. It is widely used internationally, with translations into over 40 languages; in Italy, it is used in randomized controlled clinical trials of interventions aimed at reducing self-stigma, with inclusion criteria based on a score greater than 1 on a scale from 0 to 3 (Boyd et al., 2014). In the present sample, Cronbach's alpha at T0 was 0.88.

Data Analysis

Prior to the main analysis, comparisons were conducted between the two groups for all variables measured at baseline using independent samples t-tests (SPSS v.29) to verify the homogeneity and comparability of the initial conditions.

The main analyses were conducted in R (RStudio 2025.09.1 Build 401) using the nparLD package, specifically developed for the analysis of nonparametric longitudinal data. Since the assumptions of normality and homoscedasticity were not met, and given the small sample sizes of the groups, a nonparametric ANOVA for mixed designs was used, with a between-subjects factor (group: control vs. experimental) and a within-subjects factor with three time levels (T0, T1, T2). For each outcome, ANOVA-Type Statistics (ATS) were calculated and used to test the main effect of group, the main effect of time, and the group × time interaction. Significance was assessed by applying Box's correction for degrees of freedom, with a significance level set at $p < 0.05$.

To further investigate significant interactions or those approaching significance, post-hoc analyses were performed: the Wilcoxon signed-rank test for paired data for within-group comparisons (T0–T1, T0–T2, T1–T2), with Holm's correction for multiple comparisons; Mann–Whitney tests for between-group comparisons at each time point, with effect size estimated using Cliff's delta. Finally, the medians and interquartile ranges for each group × time condition are reported, along with the Relative Treatment Effects (RTE) provided by nparLD, which can be interpreted as nonparametric indicators of effect size.

RESULTS

Participant Characteristics and Descriptive Statistics

Participant sociodemographic characteristics are summarized in Table 1.

Table 1
Sociodemographic characteristics of the sample

Variable	Category	Total (N = 17)	Control (n = 9)	Experimental (n = 8)
N		17	9	8
Age (years)	Mean ± SD	43.47 ± 16.48	41.89 ± 15.95	45.25 ± 17.98
	Range	18–69	22–66	18–69
Sex, n (%)	Female	8 (47.1%)	5 (55.6%)	3 (37.5%)
	Male	9 (52.9%)	4 (44.4%)	5 (62.5%)
Note. SD = standard deviation. Percentages are calculated within each group.				

Comparisons between the two groups for all variables measured at baseline, using t-tests on independent samples (SPSS v. 29), confirmed the homogeneity and comparability of the initial conditions (Table 2).

Table 2
Between-group comparisons at baseline (T0)

Variable	Control, mean (SD)	Experimental, mean (SD)	t	p
HAM-A	20.22 (10.10)	22.00 (9.77)	-0.37	0.718
HAM-D	14.89 (7.15)	18.75 (8.48)	-1.01	0.331
LEE_F	14.33 (8.22)	19.38 (5.04)	-1.54	0.146
LEE	15.56 (9.32)	21.00 (16.34)	-0.83	0.424
ISMI TOT	52.67 (10.11)	57.75 (15.25)	-0.80	0.440
ISMI %	1.82 (0.35)	1.99 (0.53)	-0.80	0.440
MSAQL_F	74.11 (14.46)	70.88 (8.03)	0.58	0.573
MSAQL	56.89 (14.13)	53.62 (11.50)	0.52	0.608
MMAS	1.22 (0.97)	1.12 (1.36)	0.17	0.869
PSP_1	1.33 (1.66)	1.62 (1.19)	-0.42	0.681
PSP_2	1.33 (1.32)	1.12 (1.55)	0.30	0.772
PSP_3	0.78 (1.56)	0.75 (1.16)	0.04	0.967
PSP_4	0.56 (0.88)	1.00 (1.41)	-0.77	0.459
NEL_C	42.11 (4.54)	41.62 (2.72)	0.27	0.791
NEL_SS	25.22 (4.60)	24.88 (3.31)	0.18	0.860
NEL_CC	22.56 (3.13)	21.00 (1.77)	1.28	0.223
NEL_CN	22.33 (2.92)	22.25 (3.20)	0.06	0.956
NEL_SM	17.11 (2.47)	19.25 (1.83)	-2.04	0.060
NEL_PH	15.22 (3.11)	13.12 (2.23)	1.61	0.129
NEL total	144.56 (9.08)	142.12 (5.67)	0.67	0.514
RAS_F	28.44 (5.46)	25.62 (5.85)	1.02	0.323
RAS_D	9.00 (2.74)	8.38 (2.26)	0.51	0.614
RAS_OR	15.22 (4.32)	14.00 (3.02)	0.68	0.507
RAS_FA	12.00 (3.20)	11.38 (2.56)	0.45	0.662
RAS_S	9.78 (2.28)	8.25 (2.43)	1.33	0.204

Note. Independent-samples t tests (SPSS, version 29). Values are means (SDs). P values are not adjusted for multiple comparisons. SD = standard deviation. Outcome and subscale codes are retained from the source document.

Variable	Control, mean (SD)	Experimental, mean (SD)	t	p
RAS total	127.78 (26.96)	115.50 (25.41)	0.97	0.349

Note. Independent-samples t tests (SPSS, version 29). Values are means (SDs). P values are not adjusted for multiple comparisons. SD = standard deviation. Outcome and subscale codes are retained from the source document.

The descriptive statistics for the measurements taken are shown in Table 3.

Table 3
Outcome scores at the three assessment time points

Outcome	Group	T0	T1	T2
RAS	Control	127.78 ± 26.96	138.56 ± 12.10	139.00 ± 10.61
	Experimental	115.50 ± 25.41	160.88 ± 11.73	155.38 ± 6.44
	Total	122.00 ± 26.19	149.06 ± 16.29	146.71 ± 12.06
NEL	Control	144.56 ± 9.08	149.67 ± 7.89	148.78 ± 5.21
	Experimental	142.13 ± 5.67	151.00 ± 7.01	151.13 ± 7.49
	Total	143.41 ± 7.54	150.29 ± 7.29	149.88 ± 6.29
HAM-D	Control	14.89 ± 7.15	9.56 ± 9.22	14.11 ± 9.08
	Experimental	18.75 ± 8.48	10.63 ± 3.85	11.25 ± 4.86
	Total	16.71 ± 7.81	10.06 ± 7.02	12.76 ± 7.33
HAM-A	Control	20.22 ± 10.10	11.78 ± 11.79	20.56 ± 11.31
	Experimental	22.00 ± 9.77	13.63 ± 7.89	13.88 ± 5.25
	Total	21.06 ± 9.67	12.65 ± 9.88	17.41 ± 9.37
LEE	Control	15.56 ± 9.32	9.11 ± 5.04	12.78 ± 4.76
	Experimental	21.00 ± 16.34	14.75 ± 16.56	16.13 ± 6.94
	Total	18.12 ± 12.97	11.76 ± 11.88	14.35 ± 5.95
ISMI	Control	52.67 ± 10.11	46.67 ± 17.71	59.56 ± 13.57
	Experimental	57.75 ± 15.25	50.88 ± 9.95	57.00 ± 3.78
	Total	55.06 ± 12.64	48.65 ± 14.31	58.35 ± 10.01
MSAQL	Control	56.89 ± 14.13	48.78 ± 29.15	61.00 ± 14.10
	Experimental	53.63 ± 11.50	59.13 ± 12.05	63.88 ± 10.84
	Total	55.35 ± 12.67	53.65 ± 22.73	62.35 ± 12.37
MMAS-4	Control	1.22 ± 0.97	0.67 ± 1.00	1.33 ± 1.12

Note. Values are means ± SDs. T0, T1, and T2 denote the three assessment time points. RAS = Recovery Assessment Scale, total score (range 41–205); NEL = Netherlands Empowerment List, total score (range 40–200); HAM-D = Hamilton Depression Rating Scale; HAM-A = Hamilton Anxiety Rating Scale; LEE = Level of Expressed Emotion Scale (family members, as described in the source); ISMI = Internalized Stigma of Mental Illness Inventory; MANSA = Manchester Short Assessment of Quality of Life (labelled MSAQL in the source tables); MMAS-4 = Morisky Medication Adherence Scale; PSP = sum of four item scores, each rated from 0 to 4 (sum range 0–16). SD = standard deviation. Instrument labels and score ranges are retained as reported and require confirmation before submission.

Outcome	Group	T0	T1	T2
	Experimental	1.13 ± 1.36	0.75 ± 0.71	1.13 ± 1.13
	Total	1.18 ± 1.13	0.71 ± 0.85	1.24 ± 1.09
PSP	Control	4.00 ± 3.94	2.56 ± 2.46	2.11 ± 1.90
	Experimental	4.50 ± 4.38	2.25 ± 2.19	1.88 ± 2.10
	Total	4.24 ± 4.02	2.41 ± 2.27	2.00 ± 1.94

Note. Values are means ± SDs. T0, T1, and T2 denote the three assessment time points. RAS = Recovery Assessment Scale, total score (range 41–205); NEL = Netherlands Empowerment List, total score (range 40–200); HAM-D = Hamilton Depression Rating Scale; HAM-A = Hamilton Anxiety Rating Scale; LEE = Level of Expressed Emotion Scale (family members, as described in the source); ISMI = Internalized Stigma of Mental Illness Inventory; MANSA = Manchester Short Assessment of Quality of Life (labelled MSAQL in the source tables); MMAS-4 = Morisky Medication Adherence Scale; PSP = sum of four item scores, each rated from 0 to 4 (sum range 0–16). SD = standard deviation. Instrument labels and score ranges are retained as reported and require confirmation before submission.

For each outcome, the main effect of time, the main effect of group, and the group × time interaction are reported (ANOVA-type statistic, ATS, with approximate degrees of freedom and p-values; Table 4). Post hoc within-group comparisons were performed using paired Wilcoxon signed-rank tests and are reported in Table A1, whereas between-group comparisons at each assessment time point were performed using Mann-Whitney tests and are reported in Table A2. P values were adjusted using the Holm method across the three temporal comparisons within each outcome and group, and across the three assessment time points within each outcome for between-group comparisons. All p values reported below are Holm-adjusted.

Recovery Assessment Scale, total score (RAS). The group × time interaction was significant (ATS = 9.28, df = 1.39, $p < 0.001$). A significant main effect of time (ATS = 13.96, df = 1.39, $p < 0.001$) and a significant main effect of group (ATS = 6.76, df = 1, $p = 0.009$) emerged. In summary, the experimental intervention produced a faster and more sustained overall improvement in the recovery process compared to the control group, with effects that consolidate over time. For the Recovery Assessment Scale total score, within-group post hoc comparisons did not identify statistically significant changes in either the control group (all $p \geq 0.5764$) or the experimental group (all $p \geq 0.06259$). Between-group comparisons showed no statistically significant difference at baseline ($p = 0.4131$). However, the experimental group had higher scores than the control group at both T1 ($W = 5.5$, $p = 0.009895$) and T2 ($W = 5$, $p = 0.009895$).

Self-confidence and hope (RAS_F). A significant main effect of time was found (ATS = 12.833, df = 1.80, $p < 0.001$), with no significant main effect of group (ATS = 3.194, df = 1, $p = 0.074$). The group × time interaction was not significant (ATS = 2.904, df = 1.80, $p = 0.060$). In summary, the results indicate that both interventions promoted a significant increase in self-confidence and hope over the course of six months. Although the experimental group showed a nominally greater descriptive increase than the control group, the absence of a statistically significant group × time interaction ($p = 0.060$) does not

allow us to confirm that the Family Psychoeducation Intervention is truly superior or more rapid than the active control on this specific dimension. For the RAS_F subscale, no statistically significant within-group changes were observed in the control group (all $p \geq 0.3231$). In the experimental group, a statistically significant difference was found between T0 and T2 ($W = 0$, $p = 0.04244$), whereas the T0-to-T1 and T1-to-T2 comparisons were not statistically significant ($p = 0.05955$ and $p = 0.3240$, respectively). Between-group comparisons identified a statistically significant difference at T1 ($W = 9$, $p = 0.02678$), but not at T0 or T2 (both $p = 0.5549$).

Goal-Orientation (RAS_OR). The group × time interaction was significant (ATS = 3.674, df = 1.46, $p = 0.039$). A significant main effect of time emerged (ATS = 10.338, df = 1.46, $p < 0.001$), in the absence of a significant main effect of group (ATS = 1.543, df = 1, $p = 0.214$). In summary, the significant interaction indicates that the two interventions follow distinct trajectories of improvement over time. The Family Psychoeducation Intervention appears to promote a more rapid optimization of goal orientation, with a marked increase already observed at T1 and subsequent maintenance at T2. In contrast, the control group shows a more gradual and linear progression, yet by the end of the six months (T2) it reaches an outcome level comparable to that of the experimental group. For the RAS_OR subscale, no statistically significant within-group changes were observed in the control group (all $p = 0.4632$). In the experimental group, statistically significant differences were found between T0 and T1 and between T0 and T2 (both $W = 0$, $p = 0.04103$), whereas the difference between T1 and T2 was not statistically significant ($p = 0.5176$). Between-group comparisons showed a statistically significant difference at T1 ($W = 10$, $p = 0.03979$), but not at T0 or T2 (both $p = 1.000$).

Willingness to Ask for Help (RAS_D). A significant main effect of time was found (ATS = 7.669, df = 1.54, $p = 0.0015$), with no significant main effect of group (ATS = 1.450, df = 1, $p = 0.229$). The group × time interaction was not significant (ATS = 2.175, df = 1.54, $p = 0.127$). In summary, the results show a significant main effect of time, indicating that both groups exhibited a significant increase in willingness to ask for help over the course of the six-month follow-up. Although the experimental group recorded a clinically higher score in purely descriptive terms at T2 compared to the control group, the absence of a statistically significant group × time interaction ($p = 0.127$) indicates that the trajectories of improvement in the two arms do not differ substantially. Both structured interventions were therefore shown to be equally effective in promoting more proactive attitudes toward seeking support in the target population. For the RAS_D subscale, within-group comparisons did not reveal statistically significant changes in either the control group (all $p \geq 0.5345$) or the experimental group (all $p \geq 0.06391$). Between-group differences were not statistically significant at T0, T1, or T2 ($p = 0.7339$, $p = 0.2901$, and $p = 0.1700$, respectively).

Trust in Others (RAS_FA). A significant main effect of time was found (ATS = 7.301, df = 1.76, $p = 0.0012$), with no significant main effect of group (ATS = 0.363, df = 1, $p = 0.547$). The group × time interaction was not significant (ATS = 1.066, df = 1.76, $p = 0.338$). Both groups showed similar temporal patterns, but the RTE scores suggest that the Experimental group exhibited a more pronounced increase from T0 (RTE = 0.270 vs. 0.362 for the Control group) toward T1 and T2. In summary, the results indicate

a significant main effect of time, with no group effect or group × time interaction. Both groups showed identical temporal patterns in their medians. RTE scores confirmed stable and parallel probabilistic estimation trajectories between the two conditions throughout the follow-up period. This subscale documents a common and homogeneous improvement over time, demonstrating that the ability to trust others develops favorably in response to both structured interventions, regardless of the specificity of the psychoeducational model received. For the RAS_FA subscale, no statistically significant within-group changes were identified in the control group (all $p \geq 0.6297$) or the experimental group (all $p \geq 0.06165$). Between-group comparisons were not statistically significant at any assessment time point (all $p = 0.8902$).

Not feeling overwhelmed by symptoms (RAS_S). A significant main effect of time was found (ATS = 6.477, df = 1.73, $p = 0.0026$), with no significant main effect of group (ATS = 0.609, df = 1, $p = 0.435$). The group × time interaction was not significant (ATS = 2.665, df = 1.73, $p = 0.078$). In summary, the results show a significant main effect of time, with no significant main effect of group. The group × time interaction showed only a marginal trend that was not statistically significant ($p = 0.078$). At the descriptive level, the Control group maintained a largely stable trend, while the Experimental group exhibited a temporal fluctuation with a nominal increase at the first follow-up, followed by a partial return to baseline levels. Although the profile suggests a possible initial, transient benefit of the Family Psychoeducation Intervention on the perception of not being dominated by symptoms, the absence of a robust statistical interaction calls for caution and does not allow for the confirmation of a clear superiority over active control in this pilot sample. For the RAS_S subscale, no statistically significant within-group changes were observed in the control group (all $p \geq 0.7531$). In the experimental group, a statistically significant difference was found between T0 and T1 ($W = 0$, $p = 0.04209$). The T0-to-T2 and T1-to-T2 comparisons were not statistically significant (both $p = 0.07272$). Between-group differences were not statistically significant at any assessment time point (all $p \geq 0.2334$).

Netherlands Empowerment List, total score (NEL). A significant main effect of time was found (ATS = 9.675, df = 1,992, $p < 0.001$), with no significant main effect of group (ATS = 0.030, df = 1, $p = 0.863$). The group × time interaction was not significant (ATS = 1.382, df = 1,992, $p = 0.251$). In summary, both groups showed a significant improvement in overall empowerment over time, peaking at the first follow-up with substantial maintenance of the effects at the second follow-up. No statistically significant differences emerged in the trajectory of change between the experimental and control groups. For the Netherlands Empowerment List total score, within-group post hoc comparisons did not identify statistically significant changes in the control group (all $p \geq 0.4149$) or the experimental group (all $p \geq 0.1269$). Between-group differences were not statistically significant at T0, T1, or T2 ($p = 0.5744$, $p = 0.6640$, and $p = 0.5744$, respectively).

Confidence (NEL_C). A significant main effect of time was found (ATS = 5.042, df = 1.83, $p = 0.0081$), with no significant main effect of group (ATS = 0.398, df = 1, $p = 0.528$). The group × time interaction was not significant (ATS = 1.632, df = 1.83, $p = 0.198$). The subscale showed significant improvements over time common to both arms, suggesting that self-confidence and confidence in one's own intentions

develop favorably within the context of a structured course of treatment and family support, regardless of the specific nature of the intervention received. For the NEL_C subscale, no statistically significant within-group changes were observed in the control group (all $p = 1.000$) or the experimental group (all $p \geq 0.1020$). Between-group comparisons were not statistically significant at T0 or T1 (both $p = 1.000$), or at T2 ($p = 0.3242$).

Professional Help (NEL_PH). A significant main effect of time was found (ATS = 10.146, df = 1.69, $p < 0.001$), with no significant main effect of group (ATS = 0.920, df = 1, $p = 0.338$). The group × time interaction was not significant (ATS = 2.131, df = 1.69, $p = 0.127$). In summary, both groups showed a significant improvement in their attitude toward professional help over time. The absence of a statistically significant group × time interaction ($p = 0.127$) indicates that the overall experience of participating in a structured, family-based therapeutic program fosters greater openness toward professional services, regardless of the specific psychoeducational model received. For the NEL_PH subscale, no statistically significant within-group changes were observed in the control group (all $p = 0.5153$). In the experimental group, the T0-to-T1 and T0-to-T2 comparisons were not statistically significant (both $p = 0.06676$), nor was the T1-to-T2 comparison ($p = 0.2021$). Between-group differences were not statistically significant at T0 ($p = 0.4340$), T1, or T2 (both $p = 1.000$).

Self Management (NEL_SM). No significant main effects were found for either time (ATS = 0.524, df = 1.68, $p = 0.561$) or group (ATS = 3.768, df = 1, $p = 0.052$), although a trend toward significance was observed. The group × time interaction was not significant (ATS = 0.112, df = 1.68, $p = 0.860$). Descriptive patterns showed a slight increase from T0 to T1 followed by a return to baseline at T2, which was more pronounced in the Experimental group than in the Control group. The Experimental group showed a trend toward greater self-management skills, suggesting a possible benefit of the intervention on this dimension of empowerment, although not statistically significant. In summary, no significant main effects of time or group × time interaction emerged, indicating the absence of a stable or differential change in self-management skills. A trend toward significance was observed for the main effect of group ($p = 0.052$), due to consistently higher scores in the experimental group at all time points, regardless of treatment. At the descriptive level, both conditions showed minimal transient fluctuation, characterized by a slight increase at T1 followed by a complete return to baseline at T2. The results indicate that neither intervention modality produced a lasting impact on the Self-Management dimension in this pilot sample. For the NEL_SM subscale, within-group comparisons did not identify statistically significant changes in the control group (all $p \geq 0.5207$) or the experimental group (all $p = 1.000$). Between-group comparisons were not statistically significant at T0, T1, or T2 ($p = 0.1895$, $p = 0.2630$, and $p = 0.4969$, respectively).

Social Support (NEL_SS). No significant main effects were found for either time (ATS = 2.967, df = 1.58, $p = 0.064$) or group (ATS = 0.107, df = 1, $p = 0.743$). The group × time interaction was not significant (ATS = 0.688, df = 1.58, $p = 0.470$). The temporal patterns differed slightly: the Control group showed an increase from T0 to T1 followed by stability, while the Experimental group remained largely stable with minimal variations. The perception of social support remained essentially stable over time for both

groups, suggesting that this dimension of empowerment may require more targeted or prolonged interventions to yield significant changes. In summary, no significant main effects of group or group × time interaction emerged. There was a trend toward significance for the main effect of time ($p = 0.064$). From a descriptive standpoint, the two groups show partially overlapping trajectories: the Control group shows a slight nominal increase from T0 to T1, followed by stabilization, while the Experimental group remains stable. For the NEL_SS subscale, no statistically significant within-group changes were observed in the control group (all $p \geq 0.1050$) or the experimental group (all $p \geq 0.8079$). Between-group differences were not statistically significant at any assessment time point (all $p = 1.000$).

Caring Community (NEL_CC). No significant main effects were found for either time (ATS = 0.877, df = 1.57, $p = 0.394$) or group (ATS = 0.493, df = 1, $p = 0.482$). The group × time interaction was not significant (ATS = 1.981, df = 1.57, $p = 0.148$). The temporal patterns indicated a slight increase at T1 followed by a return close to baseline at T2, which was more pronounced in the Experimental group than in the Control group, but without significant differences. The perception of living in a caring community showed minimal variations over the course of the observation, indicating that this dimension of empowerment remains relatively stable in the short- to medium-term. For the NEL_CC subscale, within-group comparisons showed no statistically significant changes in the control group (all $p = 1.000$) or the experimental group (all $p = 0.8356$). Between-group differences were not statistically significant at baseline ($p = 0.4377$), T1, or T2 (both $p = 1.000$).

Connectedness (NEL_CN). No significant effects were found: the main effect of group was not significant (ATS = 0.094, df = 1, $p = 0.759$), nor were the effects of time (ATS = 0.8412, df ≈ 1.80, $p = 0.421$), nor were time (ATS = 0.131, df ≈ 1.84, $p = 0.86$) or the group × time interaction. The sense of social connectedness remained essentially unchanged over time for both groups, suggesting that this dimension requires specific approaches to be effectively influenced. For the NEL_CN subscale, no statistically significant within-group changes were observed in either group across the three temporal comparisons (all $p = 1.000$). Between-group comparisons were also not statistically significant at any assessment time point (all $p = 1.000$).

Hamilton Anxiety Rating Scale (HAM-A). A significant main effect of time was found (ATS = 6.70, df = 1.73, $p = 0.0021$), with no significant main effect of group (ATS = 0.12, df = 1, $p = 0.729$). The group × time interaction showed a trend but did not reach significance (ATS = 2.84, df = 1.73, $p = 0.066$). Anxiety symptoms showed a significant reduction over time in both groups, with the most marked improvement at the first follow-up, followed by a partial increase, suggesting that the experience of participating in treatment facilitated a reduction in anxiety regardless of the specific intervention received. The results show a significant reduction in anxiety symptoms over time in both groups ($p = 0.0021$), with no statistical differences between treatments ($p = 0.729$). Although the Group × Time interaction showed only a marginal trend toward significance ($p = 0.066$), the descriptive trend at T2 suggests the need to further investigate a possible maintenance effect of the experimental treatment in samples with greater statistical power. At present, the data support the non-specific efficacy of both approaches in reducing participants' anxiety. For the Hamilton Anxiety Rating Scale, within-group post hoc comparisons did not

identify statistically significant changes in either the control group (all p ≥ 0.1293) or the experimental group (all p ≥ 0.2402). Between-group differences were not statistically significant at baseline or T1 (both p = 1.000), or at T2 (p = 0.2209).

Hamilton Rating Scale for Depression (HAM-D). No significant group effect was found (ATS = 0.071; df ≈ 1.00; p = 0.790), while a significant time effect was found (ATS = 7.610; df ≈ 1.99; p = 0.0005). The group × time interaction did not reach significance (ATS = 0.943; df ≈ 1.83; p = 0.382). Depressive symptoms showed a reduction from T0 to T1 followed by a subsequent increase at T2, with a pattern that was essentially similar in both groups. The levels observed at baseline are consistent with mild-to-moderate depressive symptoms, which partially subsided during the intervention period. The observed trend indicates a temporal variation in depression scores shared by both groups, with no evidence of a different course of the condition related to the treatment received. The initial reduction did not remain stable at the six-month follow-up, suggesting that the benefits observed at the end of the intervention tend to diminish over time. For the Hamilton Depression Rating Scale, no statistically significant within-group changes were observed in the control group (all p ≥ 0.2872) or the experimental group (all p ≥ 0.1057). Between-group comparisons were not statistically significant at any assessment time point (all p = 1.000).

Morisky Medication Adherence Scale (MMAS). The analysis of medication adherence did not reveal any statistically significant differences. No significant main effects were found for time (ATS = 5.1290, df ≈ 2, p = 0.0769), nor for group (ATS = 0.0679, df ≈ 1, p = 0.794), nor for the group × time interaction (ATS = 1.1015, df ≈ 2, p = 0.5765). Therapeutic compliance scores remained stable and consistently high in both groups throughout the study. For the Morisky Medication Adherence Scale score, within-group comparisons did not reveal statistically significant changes in the control group (all p ≥ 0.05897) or the experimental group (all p = 1.000). Between-group differences were not statistically significant at any assessment time point (all p = 1.000).

Internalized Stigma of Mental Illness Scale, punteggio totale (ISMI). No significant main effects were found for either time (ATS = 2.311, df ≈ 1.80, p = 0.102) or group (ATS = 0.249, df ≈ 1, p = 0.618). The group × time interaction was not significant (ATS = 0.104, df ≈ 1.88, p = 0.893). Internalized stigma remained essentially stable over time for both groups, indicating that this psychosocial dimension requires more specific and prolonged interventions to be effectively modified. Analysis of the mean score per item on the ISMI scale confirmed the absence of statistically significant variations. No significant main effects were found for time (ATS = 2.940, df = 1.84, p = 0.057) or for group (ATS = 0.100, df = 1. p = 0.752), nor was there a significant group × time interaction (ATS = 0.356, df = 1.79, p = 0.683). At the descriptive level, both arms showed a non-significant fluctuating trend, characterized by a partial reduction in scores at the end of the intervention (T1) followed by an increase at follow-up (T2). This pattern was marginally more pronounced in the control group than in the experimental group. Overall, the convergence of analytical results between the total score and the mean score per item highlights that internalized stigma represents a psychosocial dimension that is particularly resistant to change in the short and medium term. For the Internalized Stigma of Mental Illness total score, no statistically

significant within-group changes were observed in the control group (all p ≥ 0.3220) or the experimental group (all p ≥ 0.3207). Between-group comparisons were not statistically significant at T0, T1, or T2 (all p = 1.000).

Socially Useful Activities (PSP_1). The analysis of the PSP_1 subscale, which measures impairment in socially useful activities, did not reveal any statistically significant differences. No significant main effects were found for time (ATS = 0.790, df = 1.88, p = 0.447) or for group (ATS = 0.004, df = 1, p = 0.952), and the group × time interaction did not reach statistical significance (ATS = 0.894, df = 1.88, p = 0.403). At the descriptive level, the two arms showed partially divergent trajectories, although these were not supported by statistical significance: the control group maintained absolute stability across the three time points, while the experimental group showed a progressive descriptive shift toward lower scores. Overall, the data indicate substantial long-term invariance in levels of specific social functioning for both groups. For PSP item 1, no statistically significant within-group changes were observed in the control group (all p = 1.000) or the experimental group (all p ≥ 0.5087). Between-group comparisons were not statistically significant at any assessment time point (all p = 1.000).

Personal and Social Relationships (PSP_2). Analysis of the second domain of the PSP scale, which pertains to the quality and integrity of personal and social relationships, confirmed a pattern of longitudinal stability. No significant main effects emerged for the time factor (ATS = 1.819, df = 1.59, p = 0.170) or for the group factor (ATS = 0.016, df = 1, p = 0.899), and the group × time interaction did not reach statistical significance (ATS = 0.441, df = 1.59, p = 0.598). The descriptive scores on this subscale remained comparable between the two study arms throughout the duration of the protocol. For PSP item 2, within-group comparisons did not identify statistically significant changes in the control group (all p ≥ 0.1383) or the experimental group (all p = 1.000). Between-group differences were not statistically significant at T0, T1, or T2 (all p = 1.000).

Personal Appearance and Hygiene (PSP_3). Analysis of the third domain of the PSP scale, focused on personal appearance and hygiene, revealed a significant main effect of time (ATS = 4.251, df = 1.47, p = 0.0245), indicating a longitudinal change in scores. In contrast, no significant effects were found for the group factor (ATS = 0.059, df = 1, p = 0.808) or for the group × time interaction (ATS = 1.242, df = 1.47, p = 0.281). At the descriptive level, analysis of the distributions and mean ranks indicates a slight reduction in impairment over time: the control group showed a gradual and steady decline in scores, while the experimental group exhibited a more pronounced reduction between T0 and T1, followed by stabilization at follow-up. Overall, the data indicate a progressive and generalized clinical improvement in personal appearance and hygiene. This positive trend, which occurred independently of the type of treatment received, suggests that ongoing participation in the clinical protocol and periodic monitoring acted as effective nonspecific factors in supporting and stimulating patients' daily self-management skills. For PSP item 3, no statistically significant within-group changes were observed in the control group (all p = 1.000). In the experimental group, the T0-to-T1 and T0-to-T2 comparisons were not statistically significant (both p = 0.3629), whereas the T1-to-T2 comparison was reported as not estimable. Between-

group differences were not statistically significant at T0 or T1 (p = 0.7651 and p = 0.3934, respectively). The between-group comparison at T2 was reported as not estimable.

Disruptive and Aggressive Behaviors (PSP_4). Analysis of the fourth domain of the PSP scale, concerning the presence and severity of disruptive and aggressive behaviors, revealed a significant main effect of time (ATS = 4.212, df = 1.87, p = 0.0169). In contrast, no significant main effects were found for the group factor (ATS = 0.260, df = 1, p = 0.610) or for the group × time interaction (ATS = 0.330, df = 1.87, p = 0.704). The longitudinal rank analysis shows a partial descriptive reduction in the frequency or intensity of these behaviors over time: the control group showed a gradual and steady decrease in scores, while the experimental group exhibited a more pronounced reduction between T0 and T1, which subsequently stabilized at the six-month follow-up. Overall, the results indicate a significant longitudinal reduction in disruptive and aggressive behaviors across the entire sample. For PSP item 4, no statistically significant within-group changes were identified in the control group (all p ≥ 0.2669) or the experimental group (all p ≥ 0.5087). Between-group comparisons were not statistically significant at any assessment time point (all p = 1.000).

Patient-Perceived Expressed Emotionality (LEE). Analysis of patient-perceived expressed emotionality using the LEE scale revealed a significant main effect of time (ATS = 4.762; df ≈ 1.87; p = 0.010). In contrast, no significant main effects were found for the group factor (ATS = 0.394; df = 1; p = 0.530) or for the group × time interaction (ATS = 0.035; df ≈ 1.89; p = 0.959). The Relative Treatment Effects (RTE) analysis confirmed a marked longitudinal fluctuation across the entire sample, characterized by a decrease in scores at the end of the intervention followed by an increase at follow-up. This temporal pattern indicates that the reduction in perceived expressed emotionality is a transient phenomenon. The initial reduction may reflect the immediate clinical impact of the therapeutic alliance and initial commitment to the treatment protocol—nonspecific factors capable of promoting a temporarily more favorable perception of the family's emotional climate (Girdhar et al., 2024). However, the substantial return of scores to baseline levels six months after the end of treatment (T2) suggests that the benefits achieved in terms of the perception of relational dynamics tend to diminish once structured clinical support has ended. For the Level of Expressed Emotion Scale score labelled LEE, within-group comparisons did not reveal statistically significant changes in the control group (all p ≥ 0.1239) or the experimental group (all p ≥ 0.4524). Between-group differences were not statistically significant at T0, T1, or T2 (all p = 1.000).

Expressed Emotionality of Family Members (LEE_F). The analysis of expressed emotionality, assessed directly in family members using the LEE_F scale, revealed results of significant clinical interest. A statistically significant group × time interaction emerged (ATS = 3.215, df = 1.98, p = 0.0407), indicating markedly divergent developmental trajectories between the two study arms. A significant main effect of time was also found (ATS = 5.870, df = 1.98, p = 0.0029), in the absence of a significant main effect of the group factor (ATS = 0.003, df = 1, p = 0.958). From a descriptive perspective, the two groups exhibited opposite temporal patterns. The experimental group showed a progressive, consistent, and clinically relevant reduction in expressed emotionality at all time points considered, nearly halving the initial score

by the end of the follow-up. The significance of the interaction confirms the differential efficacy of the experimental intervention in permanently altering the family's emotional climate.

For the LEE_F score, no statistically significant within-group changes were observed in the control group (all $p \geq 0.3686$). In the experimental group, a statistically significant difference was found between T0 and T2 ($W = 36$, $p = 0.04244$), whereas the T0-to-T1 and T1-to-T2 comparisons were not statistically significant ($p = 0.1168$ and $p = 0.2359$, respectively). Between-group comparisons were not statistically significant at any assessment time point (all $p = 0.6829$).

Patients' Quality of Life (MSAQL). The analysis of perceived quality of life, as assessed using the MSAQL scale, did not reveal any statistically significant differences between groups or over time. No significant main effects were found for the group factor (ATS = 0.014, df = 1, $p = 0.905$) or for the group × time interaction (ATS = 1.441, df = 1.94, $p = 0.237$). The main effect of time showed a marginal trend toward longitudinal change, but did not reach the conventional threshold for statistical significance (ATS = 2.817, df = 1.94, $p = 0.061$). From a descriptive perspective, both study arms showed essentially stable satisfaction scores from baseline to the six-month follow-up. Although the temporal trend ($p = 0.061$) suggests a slight and generalized numerical increase in quality-of-life scores concurrent with participation in the protocol, the lack of statistical significance precludes the possibility of attributing stable structural changes to this functional dimension. Overall, the data indicate that perceived quality of life remained unchanged regardless of the specific type of intervention received. For the quality-of-life score labelled MSAQL, within-group post hoc comparisons did not identify statistically significant changes in the control group (all $p \geq 0.6382$) or the experimental group (all $p \geq 0.2380$). Between-group differences were not statistically significant at T0 or T1 (both $p = 1.000$), or at T2 ($p = 0.9952$).

Quality of life as assessed by family members (MSAQL_F). The analysis of quality of life reported by family members using the MSAQL_F scale did not reveal any statistically significant differences. No significant main effects were found for time (ATS = 0.950; df = 1.625; $p = 0.371$) or for group (ATS = 0.519; df = 1; $p = 0.471$), and the group × time interaction was largely nonsignificant (ATS = 0.138; df = 1.806; $p = 0.851$). From a descriptive perspective, the two arms showed slight numerical fluctuations that were not supported by statistical significance. For the family-rated quality-of-life score labelled MSAQL_F, no statistically significant within-group changes were observed in either group (all $p = 1.000$). Between-group comparisons were also not statistically significant at any assessment time point (all $p = 1.000$).

Table 4
Longitudinal outcomes: group, time, and group × time effects

Outcome	Group ATS (df), p	Time ATS (df), p	Group × time ATS (df), p
Recovery: RAS total and subscales			
RAS TOT	6.76 (1), **p = 0.009**	13.96 (1.39), **p < 0.001**	9.28 (1.39), **p < 0.001**
RAS_F	3.194 (1), p = 0.074	12.833 (1.80), **p < 0.001**	2.904 (1.80), p = 0.060
RAS_OR	1.543 (1), p = 0.214	10.338 (1.46), **p < 0.001**	3.674 (1.46), **p = 0.039**
RAS_D	1.450 (1), p = 0.229	7.669 (1.54), **p = 0.0015**	2.175 (1.54), p = 0.127
RAS_FA	0.363 (1), p = 0.547	7.301 (1.76), **p = 0.0012**	1.066 (1.76), p = 0.338
RAS_S	0.609 (1), p = 0.435	6.477 (1.73), **p = 0.0026**	2.665 (1.73), p = 0.078
Empowerment			
NEL	0.030 (1), p = 0.863	9.675 (1.992), **p < 0.001**	1.382 (1.992), p = 0.251
NEL_C	0.398 (1), p = 0.528	5.042 (1.83), **p = 0.0081**	1.632 (1.83), p = 0.198
NEL_PH	0.920 (1), p = 0.338	10.146 (1.69), **p < 0.001**	2.131 (1.69), p = 0.127
NEL_SM	3.768 (1), p = 0.052	0.524 (1.68), p = 0.561	0.112 (1.68), p = 0.860
NEL_SS	0.107 (1), p = 0.743	2.967 (1.58), p = 0.064	0.688 (1.58), p = 0.470
NEL_CC	0.493 (1), p = 0.482	0.877 (1.57), p = 0.394	1.981 (1.57), p = 0.148
NEL_CN	0.093 (1), p = 0.761	0.8412 (1.80), p = 0.421	0.131 (1.84), p = 0.860
Anxiety and depressive symptoms			

Note. ATS = ANOVA-type statistic, calculated using the nparLD package; df = approximate degrees of freedom, shown in parentheses. Each cell reports ATS (df), followed by the p value. Bold p values indicate p < 0.05. P values are not adjusted for multiplicity across outcomes. Numerical precision is retained from the source, with trailing zeros added where needed for consistent minimum precision. Outcome and subscale codes are retained from the source document.

Outcome	Group ATS (df), p	Time ATS (df), p	Group × time ATS (df), p
Recovery: RAS total and subscales			
HAM-A	0.12 (1), p = 0.729	6.70 (1.73), **p = 0.0021**	2.84 (1.73), p = 0.066
HAM-D	0.071 (1), p = 0.789	7.61 (1.99), **p = 0.0005**	0.9433 (1.82), p = 0.382
Quality of life			
Patient quality of life (MSAQL)	0.014 (1), p = 0.905	2.817 (1.94), p = 0.061	1.441 (1.94), p = 0.237
Family-rated quality of life (MSAQL_F)	0.519 (1), p = 0.471	0.950 (1.625), p = 0.371	0.138 (1.806), p = 0.851
Secondary outcomes: personal and social functioning (PSP)			
PSP_1	0.004 (1), p = 0.952	0.790 (1.88), p = 0.447	0.894 (1.88), p = 0.403
PSP_2	0.016 (1), p = 0.899	1.819 (1.59), p = 0.170	0.441 (1.59), p = 0.598
PSP_3	0.059 (1), p = 0.808	4.251 (1.47), **p = 0.0245**	1.242 (1.47), p = 0.281
PSP_4	0.260 (1), p = 0.610	4.212 (1.87), **p = 0.0169**	0.330 (1.87), p = 0.704
Medication adherence and internalized stigma			
MMAS	0.0679 (1), p = 0.076	5.1290 (2), p = 0.0769	1.1015 (2), p = 0.576
ISMI TOT	0.249 (1), p = 0.618	2.311 (1.80), p = 0.102	0.104 (1.88), p = 0.893
ISMI %	0.100 (1), p = 0.752	2.940 (1.84), p = 0.057	0.356 (1.79), p = 0.683
Perceived expressed emotion			
LEE	0.394 (1), p = 0.530	4.762 (1.87), **p = 0.010**	0.035 (1.89), p = 0.959
LEE_F	0.003 (1), p = 0.958	5.870 (1.98), **p = 0.0029**	3.215 (1.98), **p = 0.0407**

Note. ATS = ANOVA-type statistic, calculated using the nparLD package; df = approximate degrees of freedom, shown in parentheses. Each cell reports ATS (df), followed by the p value. Bold p values indicate $p < 0.05$. P values are not adjusted for multiplicity across outcomes. Numerical precision is retained from the source, with trailing zeros added where needed for consistent minimum precision. Outcome and subscale codes are retained from the source document.

DISCUSSION

This study evaluated the effect of a family psychoeducation intervention based on Falloon's model, compared with a family program promoting a healthy lifestyle on a set of outcomes focused on the individual and the relational context in 17 families of patients with DDM. The comparison was therefore made using an active control group: both conditions shared the same program structure, frequency of contact, professional attention, and family involvement. This study design is essential for interpreting the results, as it makes the comparison more rigorous and attributes a portion of the effect to common therapeutic factors, in the absence of an active control, would have been attributed to the specific intervention.

The statistically most robust result concerns the personal recovery process. The total score on the Recovery Assessment Scale (RAS) showed a statistically significant group × time interaction, with differences between the groups favoring the experimental arm at both T1 and T2, despite comparable baseline conditions. The second significant finding from a relational perspective concerns the family dimension: EE reported by family members showed a significant interaction, characterized by a progressive reduction in the experimental group and a substantially stable trend in the control group. In contrast, symptomatic outcomes, overall psychosocial functioning, medication adherence, and internalized stigma did not show any statistically significant group × time interactions. The preliminary picture that emerges suggests that the intervention exerts a differential effect primarily on dimensions that are not strictly symptomatic, such as the patient's regaining of hope, sense of purpose, and agency, as well as the quality of the family's emotional climate. This is a distinction that the literature on recovery has long placed at the center of the evaluation of psychosocial interventions, and which finds in this study an initial exploratory empirical confirmation. The higher total RAS score observed in the experimental group appears theoretically consistent with the model underlying Falloon's intervention. Indeed, this intervention was not limited to a psychoeducational component focused solely on conveying information about the disorder, but included structured modules dedicated to communication and problem-solving, aimed at defining and pursuing shared goals through practical exercises and standardized materials. These specific components may have accelerated processes inherent to personal recovery rather than producing direct clinical benefits superior to those of the active control group. The international literature confirms that the collaborative setting of goals and systematic support for their pursuit are closely associated with higher levels of personal recovery (Stewart et al., 2022). The observed pattern fits within the theoretical framework of interventions focused on personal recovery, in which dimensions related to agency and the ability to set personal goals are central elements. Approaches that support the identification of individual values and the acquisition of related coping skills are associated with significant improvements in these areas (Tse et al., 2016; Sánchez-Guarnido et al., 2024). From this perspective, goal planning can serve as one of the key mechanisms through which the intervention promotes motivation and engagement in change-oriented behaviors (Stewart et al., 2022), while simultaneously fostering self-determination and a sense of control over one's health journey (Francis et al., 2024). Overall, the stronger evidence regarding goal orientation, accompanied by less conclusive and purely descriptive findings in the subscales of hope and symptom

provides an element of theoretical consistency but not definitive proof of causal mechanisms, as these findings are compromised by the statistical instability inherent in small samples. One aspect worthy of note is the temporal pattern of the change. In the experimental group, the increase in RAS scores is concentrated between T0 and T1, after which it stabilizes, whereas in the control group, the trend is more gradual. This profile is consistent with the hypothesis that the intervention produces an initial acceleration of recovery processes, followed by a consolidation phase; the lack of any further increase at T2 should be interpreted as the maintenance of the gains already achieved. In the context of DDM, compromised hope and difficulty in envisioning future goals pose significant obstacles to active participation in treatment (Yavari et al., 2023). An intervention that addresses these dimensions at an early stage can foster more stable engagement in the care process. A notion that aligns with the established role of family psychoeducation in bolstering confidence through the collaborative planning of goals (Miklowitz, 2023).

The reduction in EE reported by family members constitutes a specific indicator of a change in the relational context that the intervention was intended to address. This outcome is of crucial clinical importance in DDM: the classic meta-analytic literature, in fact, highlights that high EE strongly predicts relapse in mood disorders, showing a moderate association ($r = .39$) that is even higher than that documented in schizophrenia ($r = .30$; Butzlaff & Hooley, 1998). Within this construct, family members' criticism and hostility act as primary psychosocial risk factors for the depressed patient, while relational warmth plays a protective role. Therefore, the significance of the findings lies in the potential modification of a factor closely associated with long-term vulnerability to depression. In this study, the LEE trajectory among family members in the experimental group decreased progressively across the three assessment time points, whereas the control group remained stable. Although the interaction is statistically significant, the small sample size requires that this effect be considered a preliminary estimate. Nevertheless, the progressive nature of the decrease supports the hypothesis of a process of learning and consolidation of family skills. Falloon's program did not merely expand theoretical knowledge of DDM but also addressed the regulation of emotional responses and shared problem-solving. A greater understanding of the disorder can reduce blame-based attributions; communication training can curb criticism; and stress management strategies can mitigate the family's emotional reactivity (Fahrer et al., 2022; Girdhar et al., 2024). This interpretation is supported by experimental evidence from structured interventions based on psychoeducation and communication, which have documented significant reductions in EE among caregivers (Öksüz et al., 2017; Shetty et al., 2023). However, the exclusively self-report nature of the LEE scale calls for careful inferential caution. The correspondence between self-report measures and observational of EE is significant but not perfect (Kazarian, 1992; Van Humbeeck et al., 2002); the finding must therefore be described, strictly speaking, as a reduction in self-reported EE. A change in the way one perceives one's own attitudes may nonetheless constitute a fundamental proximal outcome and a preliminary step toward behaviors that are concretely less critical or intrusive in daily life. The profile of the EE as perceived by patients is more complex; here, a main effect of time emerged in the absence of a significant interaction: both groups showed a reduction at T1 followed by a partial increase at T2. From an inferential standpoint, this does

not demonstrate a specific effect of the Falloon model, but indicates that both pathways influenced the perception of the family climate. Since the control group also included a structured component involving family members, the shared change likely reflects common nonspecific factors, such as professional support, regular contact, and family involvement. The divergence between family members and patients also suggests that behavioral changes require different incubation periods and levels of sensitivity to be fully recognized by the patient in everyday life.

Psychosocial functioning showed an improvement over time limited to the areas of self-care and disruptive behaviors, representing an effect shared by both conditions. It is plausible that both programs supported self-care through different but converging mechanisms: psychoeducation by fostering active management of the disorder, and the lifestyle program by promoting healthy daily routines. This finding is consistent with evidence that self-management interventions produce small- to moderate-sized benefits on subjective and functional outcomes (Lean et al., 2019). The other areas of the PSP remained unchanged, confirming that functional recovery in social roles typically requires specific and context-based rehabilitative interventions, which are not fully addressed by interventions focused on the nuclear family (Vita et al., 2023; Barlati et al., 2024).

Internalized stigma did not show any significant changes, consistent with the program's lack of modules explicitly aimed at the cognitive restructuring of self-stigmatizing beliefs or at structured peer-to-peer dialogue—strategies that have historically been necessary to influence this construct (Tsang et al., 2016; Moreno Dias et al., 2024). Similarly, the analysis of empowerment is dramatically affected by the severe psychometric instability of the instrument in our small sample (NEL Cronbach's alpha ranging from 0.05 to 0.25). Although a favorable descriptive trend was observed in the self-management domain, the psychometric fragility of the measure requires that the data on empowerment be considered purely speculative and lacking the methodological robustness found in the RAS.

The literature on family psychoeducation in DDM is still in the process of consolidation. Although the meta-analysis by Katsuki et al. (2022) of 5 RCTs found a significant effect on reducing the severity of depressive symptoms but no significant effects on family functioning, the present study shows the opposite pattern. This divergence does not constitute a methodological contradiction but rather an advancement in knowledge related to the experimental design. The studies included in Katsuki's meta-analysis predominantly used passive controls, whereas the present trial employs an active control of equal intensity. When the comparison is structured equivalently, the proportion of the effect attributable to nonspecific therapeutic factors is controlled for. The interpretive implication is of great significance: a large part of the benefit on depressive symptoms that the previous literature attributes to specific psychoeducation is likely attributable to common components of family interventions, as demonstrated by the improvement in symptoms observed in both of our treatment arms.

In contrast, the net advantage maintained by the experimental group regarding RAS and the caregiver's expressed emotionality demonstrates the incremental efficacy of the specific factors of the Falloon model when common factors are controlled for. From a theoretical perspective, this pattern of results

supports the need to carefully distinguish between clinical recovery and personal recovery, confirming that the latter addresses distinct therapeutic targets and cannot be reduced to the mere attenuation of the clinical severity of the disorder (Salzer & Brusilovskiy, 2014; Boggian et al., 2020; Winsper et al., 2020).

Limitations of the Study

First, the sample size is small (N = 17 households; 8 in the experimental group, 9 in the control group), and no a priori statistical power calculation was performed. The study is therefore significantly underpowered, increasing the risk of Type II errors. Consequently, sensitivity in detecting small or moderate effects is reduced, and parameter estimates are characterized by high sampling variability. The observed effect sizes should be regarded as preliminary estimates subject to uncertainty and not as definitive indicators of efficacy. This limitation particularly affects the interpretation of outcomes near the significance threshold, which are of purely exploratory value.

Second, conducting multiple statistical tests on an extensive set of outcomes, without an overall correction for multiple comparisons, significantly increases the risk of Type I error inflation. Although the Holm procedure was routinely applied for post-hoc within-group comparisons, the absence of an adjustment at the level of the entire family of hypotheses requires that the overall framework be considered as generating hypotheses rather than as confirmatory. Furthermore, the exploration of post-hoc comparisons on variables that had not shown a statistically significant omnibus interaction constitutes a violation of the logic of hypothesis testing, the value of which remains strictly ancillary and descriptive.

Third, the internal consistency indices for the NEL scale were found to be unsatisfactory in the present sample, with Cronbach's alpha coefficients ranging from 0.05 to 0.25. This finding does not necessarily reflect a structural property of the instrument, but rather the mathematical instability of the coefficient when calculated on a highly limited sample, coupled with the potential error variance introduced by the dimensionality of the scale. Given the low psychometric reliability observed in our sample, the results related to empowerment do not possess the same methodological robustness as those related to the RAS and must be interpreted with extreme caution.

Fourth, the longitudinal design does not allow for the assessment of the trajectory and long-term stability of changes. Complex constructs such as social functioning and expressed emotionality typically exhibit nonlinear trends over longer time periods. The partial increase in symptom scores and patient-reported emotional distress observed at T2 suggests a possible decline in therapeutic benefit following the cessation of active treatment. It will therefore be necessary to implement longer follow-up periods to determine whether gains in personal recovery require subsequent maintenance or booster interventions, as suggested by the anti-stigma literature, which identifies longer incubation periods for the restructuring of internalized stigma.

In terms of measurement tools, the assessment of expressed emotionality was conducted exclusively through self-report questionnaires completed by family members. This approach captures the caregiver's subjective perception of their own attitudes and emotional responses and does not constitute an objective or third-party assessment. It is therefore not possible to infer a perfect correlation between a reduction in the LEE score and an actual change in everyday communication patterns.

Finally, the external validity of the study is limited by its single-center recruitment, the inclusion of patients in a phase of clinical stability, and the voluntary enrollment of families. This design exposes the results to a potential self-selection bias, limiting the generalizability of the data to families with lower motivation, to different care settings, or to acute phases of major depressive disorder. The complete absence of missing data across the three assessment time points, while representing a methodological strength, confirms the high and atypical compliance of the selected sample, preventing a realistic assessment of attrition and dropout rates associated with the intervention in routine clinical practice.

CONCLUSIONS

In conclusion, this pilot study suggests that a family psychoeducation intervention based on Falloon's model, integrated into the routine treatment of patients with MDD who are in a state of clinical stability, may primarily influence the dimensions of personal recovery and the family's emotional climate, rather than depressive and anxiety symptoms in the strict sense. The total score on the RAS and the self-reported emotionality expressed by family members were the outcomes that showed a favorable difference compared to an active control group of equal intensity and design. In contrast, symptomatic outcomes, quality of life, and overall psychosocial functioning showed improvements common to both conditions, consistent with the nonspecific therapeutic effects exerted by common factors such as structured monitoring, the therapeutic alliance, and active engagement in a treatment program. These preliminary results must be interpreted strictly in light of the methodological limitations described. The small sample size and the absence of a priori power calculation make the estimates unstable and the effect sizes non-generalizable as point estimates. The large number of comparisons analyzed without an overall adjustment for the multiplicity of secondary outcomes means that the entire picture should be considered solely as a basis for generating hypotheses. Furthermore, the unsatisfactory internal consistency of the NEL in our sample precludes any robust inferences regarding empowerment constructs. Finally, the observation window limited to six months and the partial increase in perceived EE and distress scores at T2 do not allow us to conclude that the benefits are stable in the long term after the cessation of active treatment. The study therefore does not provide definitive evidence of efficacy, and the findings regarding personal recovery and expressed emotionality require formal validation.

The main contribution of this study lies in its methodological and planning value. The documented effect size estimates provide the empirical basis for the proper sample size determination of a future confirmatory multicenter trial. Based on these indices, a subsequent study should aim to enroll a sample of at least 60–80 households, implementing a priori power calculation and a rigorous analysis protocol

with correction for multiple comparisons. From a psychometric perspective, the use of empowerment scales with greater mathematical stability in the face of reduced sample variances will optimize the robustness of the statistical models. From a temporal perspective, extending the follow-up to 12 and 24 months will be essential for mapping nonlinear longitudinal trajectories, distinguishing structural changes from temporary fluctuations, and assessing the potential need for subsequent follow-up sessions. Finally, the inclusion of cost-effectiveness analyses will provide essential data to support the scalability and sustainability of these interventions in local mental health departments.

From a clinical perspective, the data support the hypothesis that family psychoeducational support represents an elective complementary resource, aimed not at the direct resolution of symptoms, but rather at optimizing communication, deconstructing blame-inducing attributions, and strengthening the family's problem-solving skills, while preserving the agency of the person receiving care. Such an intervention does not replace the standard clinical treatments for DDM nor does it constitute a delegation of therapeutic responsibilities to caregivers; its value lies in grounding the treatment journey within a relational network capable of fostering its sustainability over time. The exploratory convergence between the patient's progress in personal recovery, the reduction in the family member's expressed emotional distress, and the maintenance of the family member's quality of life suggests a positive systemic impact that warrants rigorous evaluation through large-scale trials.

Implications for Research and Future Directions

The preliminary evidence emerging from this pilot study, critically analyzed in light of the methodological limitations described, allows us to outline several priority research directions for future studies.

Validation through experimental designs with high statistical power: Since the primary outcome of this study is purely hypothesis-generating in nature, the top priority is to plan a multicenter RCT. This design should incorporate a sample size calculation based on the preliminary effect sizes observed here and extend the follow-up period to 12 and 24 months. A longer longitudinal follow-up period is essential in the DDM to distinguish transient effects associated with the active treatment phase from therapeutic gains that are stably maintained over time, as well as to examine the durability of benefits for outcomes near the threshold of statistical significance.

Optimization and Formal Specification of the Components of Self-Agency: The enhancement of self-agency is a cornerstone of personal recovery models, serving as a catalyst for taking an active role in therapeutic decisions, pursuing meaningful life goals, and developing adaptive and resilient responses (Francis et al., 2024). Although Falloon's program inherently stimulates these dimensions through goal planning and problem solving, future research should aim for an even more clearly defined systematization of these components. Structuring interventions that intensively and specifically target self-efficacy could transform the positive signals and descriptive trends observed in this study into statistically robust changes.

Integration of multi-component anti-stigma modules: The longitudinal stability of internalized stigma documented in this trial highlights the need to implement targeted therapeutic strategies. The international literature indicates that this construct does not respond well to an informational/psychoeducational component alone, requiring approaches explicitly aimed at deconstructing internalized stereotypes. Systematic reviews and meta-analyses confirm the superior efficacy of multi-component models in which psychoeducation is combined with cognitive restructuring techniques, social skills training, and, ideally, peer support programs (Tsang et al., 2016; Moreno Dias et al., 2024). The inclusion of such structured modules represents a methodologically consistent and necessary expansion.

Integration with network-based interventions and models of social inclusion: The invariance observed in the relational and community dimensions of empowerment, together with the stability of the PSP, reflects the structural limitations of an intervention focused primarily on the immediate family unit. To have a significant impact on the social roles of those receiving support, future research must include a synergistic integration between family psychoeducation and community-based network interventions. The literature documents the effectiveness of community-based models oriented toward inclusion and social participation on functioning and relational empowerment (Killaspy et al., 2022), supporting the argument that empowerment cannot be viewed as a mere individual attribute, but rather as an intrinsically transactional and contextual process (Christens, 2012).

Adoption of observational and externally assessed measures of the family climate: In terms of measurement tools, the exclusive use of the LEE captures only the subjective representation of the emotional climate (Kazarian, 1992; Van Humbeeck et al., 2002). The combined, multi-method use of self-report instruments and observational or third-party assessment protocols would make it possible to distinguish changes in internal perception from actual changes in the dyad's daily communication patterns, offering an empirical framework for interpreting the discrepancies in assessments between patients and family members that emerged in this study.

Modeling Mechanisms of Change: Finally, the development of this research must move beyond the mere estimation of overall effects toward an analytical explanation of causal processes. The application of structural mediation and moderation models to samples with adequate statistical power will allow us to verify whether a reduction in the caregiver's expressed emotionality or an increase in self-management act as effective psychotherapeutic mediators of the patient's personal recovery. At the same time, the analysis of moderators will help identify which clinical, anamnestic, or psychosocial variables influence the differential response to treatment.

In summary, the future research agenda must address three converging imperatives: to consolidate the replicability of the promising effects observed on proximal outcomes; to scale up the modifiability of complex constructs that have remained stable through targeted therapeutic modules; and to identify the internal mediators that explain the clinical and rehabilitative efficacy of the intervention in the real-world context of healthcare services.

Declarations

Conflict of interest:

The authors do not have any conflicts of interest to report.

References

1. Barlati S, Nibbio G, Vita A (2024) Evidence-based psychosocial interventions in schizophrenia: a critical review. Curr Opin Psychiatry 37(3):131–139. https://doi.org/10.1097/YCO.0000000000000925
2. Boardman J (2010) Personal Recovery and Mental Illness: A Guide for Mental Health Professionals. Br J Psychiatry. https://doi.org/10.1192/BJP.BP.109.069997
3. Boevink W, Kroon H, Delespaul P, Van Os J (2016) Empowerment according to persons with severe mental illness: development of the Netherlands empowerment list and its psychometric properties. Open J Psychiatry 7(1):18–30. 10.4236/ojpsych.2017.71002
4. Boggian I, Lamonaca D, Ghisi M, Bottesi G, Svettini A, Basso L, Bernardelli K, Merlin S, Liberman RP, I.R. 2 group (2020) The Italian Study on Recovery 2 Phase 1: Psychometric Properties of the Recovery Assessment Scale (RAS), Italian Validation of the Recovery Assessment Scale. Front Psychiatry 10:1000. https://doi.org/10.3389/fpsyt.2019.01000
5. Boyd JE, Adler EP, Otilingam PG, Peters T (2014) Internalized Stigma of Mental Illness (ISMI) scale: a multinational review. Compr Psychiatr 55(1):221–231. https://doi.org/10.1016/j.comppsych.2013.06.005
6. Brady P, Kangas M, McGill K (2017) Family Matters: A Systematic Review of the Evidence For Family Psychoeducation For Major Depressive Disorder. J Marital Fam Ther 43(2):245–263. https://doi.org/10.1111/jmft.12204. T., Maeda, T., Katsuki, F., Takeuchi, H., Inagaki
7. Butzlaff RL, Hooley JM (1998) Expressed emotion and psychiatric relapse: a meta-analysis. Arch Gen Psychiatry 55(6):547–552. https://doi.org/10.1001/archpsyc.55.6.547
8. Carrozzino D, Patierno C, Fava GA, Guidi J (2020) The Hamilton Rating Scales for Depression: A Critical Review of Clinimetric Properties of Different Versions. Psychother Psychosom 89(3):133–150. https://doi.org/10.1159/000506879
9. Christens BD (2012) Toward relational empowerment. Am J Community Psychol 50(1–2):114–128. https://doi.org/10.1007/s10464-011-9483-5
10. Di Lorenzo R, Dardi A, Serafini V, Amorado MJ, Ferri P, Filippini T (2024) Psychoeducational Intervention for Caregivers of Adolescents and Young Adults with Psychiatric Disorders: A 7-Year Systematic Review. J Clin Med 13(23):7010. https://doi.org/10.3390/jcm13237010
11. Di Paola F, Faravelli C, Ricca V (2008) The Level of Expressed Emotion Scale: Validazione italiana preliminare. Giornale Italiano di Psicopatologia/Italian J Psychopathol

12. Fahrer J, Brill N, Dobener LM, Asbrand J, Christiansen H (2022) Expressed Emotion in the Family: A Meta-Analytic Review of Expressed Emotion as a Mechanism of the Transgenerational Transmission of Mental Disorders. Front Psychiatry 12:721796. https://doi.org/10.3389/fpsyt.2021.721796
13. Francis CJ, Hazelton M, Wilson RL (2024) Supported Decision-Making Interventions in Mental Healthcare: A Systematic Review of Evidence on the Outcomes for People With Mental Ill Health. Health expectations: Int J public participation health care health policy 27(6):e70134. https://doi.org/10.1111/hex.70134
14. Girdhar A, Patil R, Bezalwar A (2024) Understanding the Dynamics: A Comprehensive Review of Family Therapy's Impact on Expressed Emotions in Schizophrenia Patients. Cureus 16(5):e59491. https://doi.org/10.7759/cureus.59491
15. Guarino A, Negrogno L, Compare C, Madeo A, Bolognini P, Degli Esposti L, Filippi M, Lamberini F, Morrone M, Masetti M, Serra AM, Albanesi C (2024) A Feeling of Safeness and Freedom: The Promotion of Mental Health Recovery Through Co-Production in an Italian Community Organization. Commun Ment Health J 60(8):1452–1463. https://doi.org/10.1007/s10597-024-01279-2
16. HAMILTON M. (1959) The assessment of anxiety states by rating. Br J Med Psychol, *32*(1), 50–55. https://doi.org/10.1111/j.2044-8341.1959.tb00467.x
17. Hamilton M (1967) Development of a rating scale for primary depressive illness. Br J Soc Clin Psychol 6(4):278–296. https://doi.org/10.1111/j.2044-8260.1967.tb00530.x
18. Hayhurst H, Cooper Z, Paykel ES, Vearnals S, Ramana R (1997) Expressed emotion and depression. A longitudinal study. Br J psychiatry: J mental Sci 171:439–443. https://doi.org/10.1192/bjp.171.5.439
19. Hormazábal-Salgado R, Whitehead D, Osman AD, Hills D (2024) Person-Centred Decision-Making in Mental Health: A Scoping Review. Issues Ment Health Nurs 45(3):294–310. https://doi.org/10.1080/01612840.2023.2288181
20. Katsuki F, Watanabe N, Yamada A, Hasegawa T (2022) Effectiveness of family psychoeducation for major depressive disorder: systematic review and meta-analysis. BJPsych open 8(5):e148. https://doi.org/10.1192/bjo.2022.543
21. Kazarian SS (1992) The measurement of expressed emotion: a review. Can J Psychiatry 37(1):51–56. https://doi.org/10.1177/070674379203700112
22. Killaspy H, Harvey C, Brasier C, Brophy L, Ennals P, Fletcher J, Hamilton B (2022) Community-based social interventions for people with severe mental illness: a systematic review and narrative synthesis of recent evidence. World psychiatry: official J World Psychiatric Association (WPA) 21(1):96–123. https://doi.org/10.1002/wps.20940
23. Kubota Y, Shiraishi N, Tabuse H, Kato T, Yamada A, Watanabe N, Akechi T, Furukawa TA (2018) Brief multifamily Psychoeducation for family members of patients with chronic major depression: a randomized controlled trial. BMC Psychiatry 18(1):207. https://doi.org/10.1186/s12888-018-1788-6
24. Leamy M, Bird V, Le Boutillier C, Williams J, Slade M (2011) Conceptual framework for personal recovery in mental health: systematic review and narrative synthesis. Br J psychiatry: J mental Sci

199(6):445–452. https://doi.org/10.1192/bjp.bp.110.083733

25. Lean M, Fornells-Ambrojo M, Milton A, Lloyd-Evans B, Harrison-Stewart B, Yesufu-Udechuku A, Kendall T, Johnson S (2019) Self-management interventions for people with severe mental illness: systematic review and meta-analysis. Br J psychiatry: J mental Sci 214(5):260–268. https://doi.org/10.1192/bjp.2019.54
26. Maier W, Buller R, Philipp M, Heuser I (1988) The Hamilton Anxiety Scale: reliability, validity and sensitivity to change in anxiety and depressive disorders. J Affect Disord 14(1):61–68. https://doi.org/10.1016/0165-0327(88)90072-9
27. Manzar MD, Kashoo FZ, Salahuddin M, Nureye D, Addo HA, Pandi-Perumal SR, Pakpour AH, Bahammam AS (2025) Hamilton Rating Scale for Anxiety: exploring validity with robust measures of classical theory parameters and a rating scale model in university students. BJPsych open 11(5):e176. https://doi.org/10.1192/bjo.2025.10055
28. Miklowitz DJ, Weintraub MJ, Ichinose MC, Denenny DM, Walshaw PD, Wilkerson CA, Frey SJ, Morgan-Fleming GM, Brown RD, Merranko JA, Arevian AC (2023) A Randomized Clinical Trial of Technology-Enhanced Family-Focused Therapy for Youth in the Early Stages of Mood Disorders. JAACAP open 1(2):93–104. https://doi.org/10.1016/j.jaacop.2023.04.002
29. Moreno Dias B, Giunta da Silva LM, Filho H, Bollela PT, V. R., Gabriel CS (2024) Development of qualified items for nursing education assessment: The progress testing experience. Nurse Educ Pract 81:104199. https://doi.org/10.1016/j.nepr.2024.104199
30. Morisky DE, Green LW, Levine DM (1986) Concurrent and predictive validity of a self-reported measure of medication adherence. Med Care 24(1):67–74. https://doi.org/10.1097/00005650-198601000-00007
31. Neckar P, Boxall N (2026) Pathways to empowerment: integrating psychoeducation, peer support, and co-production for lasting recovery. Eur Psychiatry 69(Suppl 1):S46. https://doi.org/10.1192/j.eurpsy.2026.10295
32. Obieche O, Tan JB, Sharma S, Bressington D, Wang T, Li M (2026) Family Psychoeducation for Major Depressive Disorder: A Systematic Review. Int J Ment Health Nurs 35(1):e70205. https://doi.org/10.1111/inm.70205
33. Öksüz E, Karaca S, Özaltın G, Ateş MA (2017) The Effects of Psychoeducation on the Expressed Emotion and Family Functioning of the Family Members in First-Episode Schizophrenia. Commun Ment Health J 53(4):464–473. https://doi.org/10.1007/s10597-017-0086-y
34. Petkari E, Giacco D, Priebe S (2020) Factorial structure of the Manchester Short Assessment of Quality of Life in patients with schizophrenia-spectrum disorders. Qual life research: Int J Qual life aspects Treat care rehabilitation 29(3):833–841. https://doi.org/10.1007/s11136-019-02356-w
35. Rossi R, Socci V, Rossi A (2022) Personal recovery in schizophrenia: A narrative review. In B. Carpiniello, A. Vita, & C. Mencacci (Eds.), *Recovery and major mental disorders* (pp. 45–56). Springer Nature Switzerland AG. https://doi.org/10.1007/978-3-030-98301-7_3

36. Saavedra J, Brzeska J, Matías-García JA, Arias-Sánchez S (2023) Quality of life and psychiatric distress in people with serious mental illness, the role of personal recovery. Psychol Psychother 96(2):525–541. https://doi.org/10.1111/papt.12451
37. Sánchez-Guarnido AJ, Ruiz-Granados MI, Garrido-Cervera JA, Herruzo J, Herruzo C (2024) Implementation of the Recovery Model and Its Outcomes in Patients with Severe Mental Disorder. Healthc (Basel Switzerland) 12(9):952. https://doi.org/10.3390/healthcare12090952
38. Sarkhel S, Singh OP, Arora M (2020) Clinical Practice Guidelines for Psychoeducation in Psychiatric Disorders General Principles of Psychoeducation. Indian J psychiatry 62(Suppl 2):S319–S323. https://doi.org/10.4103/psychiatry.IndianJPsychiatry_780_19)
39. Shetty KV, Math SB, Marimuthu P, Rawat VS (2023) Family intervention for expressed emotion in schizophrenia: Indian context. Indian J psychiatry 65(4):453–459. https://doi.org/10.4103/indianjpsychiatry.indianjpsychiatry_526_22
40. Stawnychy MA, Teitelman AM, Riegel B (2021) Caregiver autonomy support: A systematic review of interventions for adults with chronic illness and their caregivers with narrative synthesis. J Adv Nurs 77(4):1667–1682. https://doi.org/10.1111/jan.14696
41. Stewart V, McMillan SS, Hu J, Ng R, El-Den S, O'Reilly C, Wheeler AJ (2022) Goal planning in mental health service delivery: A systematic integrative review. Front Psychiatry 13:1057915. https://doi.org/10.3389/fpsyt.2022.1057915
42. Tan XI, Patel I, Chang J (2014) Review of the four item Morisky medication adherence scale (MMAS-4) and eight item Morisky medication adherence scale (MMAS-8). INNOVATIONS Pharm 5(3):5. https://doi.org/10.24926/iip.v5i3.347
43. Tsang HW, Ching SC, Tang KH, Lam HT, Law PY, Wan CN (2016) Therapeutic intervention for internalized stigma of severe mental illness: A systematic review and meta-analysis. Schizophr Res 173(1–2):45–53. https://doi.org/10.1016/j.schres.2016.02.013
44. Tse S, Tsoi EW, Hamilton B, O'Hagan M, Shepherd G, Slade M, Whitley R, Petrakis M (2016) Uses of strength-based interventions for people with serious mental illness: A critical review. Int J Soc Psychiatry 62(3):281–291. https://doi.org/10.1177/0020764015623970
45. Van Humbeeck G, Van Audenhove Ch, De Hert M, Pieters G, Storms G (2002) Expressed emotion: a review of assessment instruments. Clin Psychol Rev 22(3):323–343. https://doi.org/10.1016/s0272-7358(01)00098-8
46. van Weeghel J, van Zelst C, Boertien D, Hasson-Ohayon I (2019) Conceptualizations, assessments, and implications of personal recovery in mental illness: A scoping review of systematic reviews and meta-analyses. Psychiatr Rehabil J 42(2):169–181. https://doi.org/10.1037/prj0000356
47. Vauth R, Carpiniello B, Turczyński J, Ivanov M, Cherubin P, Lahaye M, Schreiner A (2021) Relationship between clinical outcomes measures and personal and social performance functioning in a prospective, interventional study in schizophrenia. Int J Methods Psychiatr Res 30(2):e1855. https://doi.org/10.1002/mpr.1855

48. Vita A, Barlati S, Ceraso A, Nibbio G, Durante F, Facchi M, Deste G, Wykes T (2024) Durability of Effects of Cognitive Remediation on Cognition and Psychosocial Functioning in Schizophrenia: A Systematic Review and Meta-Analysis of Randomized Clinical Trials. Am J Psychiatry 181(6):520–531. https://doi.org/10.1176/appi.ajp.20230396
49. White S, Dominise C, Naik D, Killaspy H (2016) The reliability of the Personal and Social Performance scale - informing its training and use. Psychiatry Res 243:312–317. https://doi.org/10.1016/j.psychres.2016.06.047
50. Winsper C, Crawford-Docherty A, Weich S, Fenton SJ, Singh SP (2020) How do recovery-oriented interventions contribute to personal mental health recovery? A systematic review and logic model. Clin Psychol Rev 76:101815. https://doi.org/10.1016/j.cpr.2020.101815
51. Yavari S, Radfar M, Faraji N, Goli R, Bagherzadi A, Atharifar R, Kazemi S, Dadashi M (2023) Self-esteem and optimism in patients with major depression disorder: a cross-sectional study. Annals Med Surg (2012) 85(11):5379–5383. https://doi.org/10.1097/MS9.0000000000001263

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- Appendix.docx