Brazilian Journal of Psychiatry
bjp
Revista Brasileira de Psiquiatria
Brazilian Psychiatric Association
Fully open access
No APCs

BJP PRE-PROOF

BJP PRE-PROOF
(article published as accepted)

Review Article

Effectiveness of Group Rehabilitation Therapy in Improving Negative Symptoms and Functional Recovery among Long-Stay Psychiatric Patients: A Systematic Review and Meta-Analysis

Wei Li, Lijuan Fu

http://doi.org/10.47626/1516-4446-2026-5168

Submitted: 07-Aug-2026

Accepted: 19-Oct-2026

This is a preliminary, unedited version of a manuscript that has been accepted for publication in the Brazilian Journal of Psychiatry. As a service to our readers, we are providing this early version of the manuscript. The manuscript will still undergo copyediting, typesetting, and review of the resulting proof before it is published in final form. The final version may present slight differences in relation to the present version.

Effectiveness of group rehabilitation therapy in improving negative symptoms and functional recovery among long-stay psychiatric patients: a systematic review and meta-analysis

Running title: Group rehabilitation and recovery in long-stay psychiatric care

Wei Li[1,#], **Lijuan Fu**[1,#]

[1]Mood Disorder Ward, Tianjin Anding Hospital, Tianjin China.

 # Wei Li and Lijuan Fu contributed equally to this work.

Corresponding author: Lijuan Fu, Mood Disorder Ward 10, Tianjin Anding Hospital, Tianjin 300222, P.R. China
E-mail: fulijuan1940@outlook.com
Tel: +86 13662084211

BJP PRE-PROOF

Abstract

Long-stay psychiatric patients frequently experience persistent negative symptoms, functional disability, institutional dependence, and limited opportunities for recovery. Group rehabilitation therapy combines behavioral activation, social-skills practice, illness management, and structured activity, but its effectiveness in long-stay psychiatric populations remains insufficiently quantified. This systematic review and meta-analysis evaluated the effects of structured group rehabilitation therapy on negative symptoms and functional recovery and examined factors associated with treatment response. Eighteen controlled studies involving 1,928 participants were included. Study-level meta-analysis showed reduced negative symptoms (SMD -0.48, 95% CI -0.57 to -0.39; I^2=0%) and improved functional recovery (SMD 0.53, 95% CI 0.44 to 0.62; I^2=0%). Eleven randomized studies (9

individually randomized and 2 cluster-randomized) were directionally consistent after the prespecified central cluster-precision adjustment. Programmes lasting at least 12 weeks or providing at least 24 contact hours showed larger effects in exploratory subgroup analyses. The review protocol was reported as prospectively registered in PROSPERO (CRD216362732). The findings support an association between structured rehabilitation and short-term symptom and functional improvement, while limited primary-study identification and sparse long-term outcome reporting constrain causal and longitudinal interpretation.
Keywords: psychiatric rehabilitation; group therapy; negative symptoms; functional recovery; schizophrenia; long-stay patients; institutional care; meta-analysis

1. Introduction

1.1 Global burden of long-stay psychiatric hospitalization

Long-stay psychiatric hospitalization remains a major unresolved problem in mental health systems because a relatively small subgroup of patients consumes a disproportionate share of institutional resources while simultaneously experiencing the greatest restriction of autonomy, social participation, and recovery opportunity [1]. Although deinstitutionalization reduced the size of many psychiatric hospitals, it did not eliminate the need for intensive rehabilitation among patients with chronic psychosis, repeated relapse, cognitive impairment, poor family support, homelessness, forensic complexity, or high dependency in activities of daily living [2]. In many services, long-stay status reflects not only illness severity but also a failure of rehabilitation architecture: prolonged admission can weaken routine, reinforce passivity, reduce exposure to normal social contingencies, and convert functional disability into institutional dependence [3].

The clinical relevance of long-stay care is therefore not limited to bed occupancy or discharge delay; it concerns the progressive narrowing of lived experience, role identity, self-efficacy, and community readiness in people who may already have severe premorbid and illness-related disadvantages [1]. Long-stay patients are often excluded from efficacy trials because of comorbidity, fluctuating capacity, treatment resistance, and service complexity, creating an

BJP PRE-PROOF

evidence gap precisely in the population for whom rehabilitation decisions carry the greatest consequences [2]. A high-quality synthesis focused on this group is needed because findings from early psychosis, outpatient psychosocial therapy, or short-stay acute psychiatry cannot be assumed to translate directly to institutional settings characterized by chronicity, dependency, environmental restriction, and low spontaneous activation [3].

1.2 Negative symptoms as a barrier to recovery

Negative symptoms are among the strongest clinical predictors of poor functional outcome in schizophrenia spectrum and chronic severe mental illness because they directly impair initiation, persistence, emotional expression, speech output, social drive, and reward-based engagement [4]. In long-stay settings, the boundary between primary negative symptoms, secondary negative symptoms, medication-related psychomotor slowing, depressive withdrawal, environmental deprivation, and learned passivity is clinically blurred, yet the shared consequence is reduced participation in rehabilitative opportunities. This distinction matters because group rehabilitation may not reverse enduring primary negative symptom biology to the same extent that it can reduce secondary inactivity, rebuild routine, and increase social reinforcement through repeated exposure [5].

A purely symptom-focused interpretation underestimates the therapeutic target of psychiatric rehabilitation, since a patient may show only modest change on a negative symptom scale while achieving meaningful gains in attendance, self-care, conversation frequency, task completion, or discharge readiness [6]. Conversely, apparent group participation can be superficial if it is not accompanied by measurable improvement in goal-directed behavior, independent functioning, and community-relevant skills. For this reason, negative symptom change and functional recovery should be treated as related but non-equivalent endpoints, requiring separate synthesis and critical interpretation [7].

BJP PRE-PROOF

1.3 Functional recovery as a multidimensional endpoint

Functional recovery in long-stay psychiatric care includes daily living skills, social functioning, role competence, self-management, treatment engagement, environmental navigation, problem-solving, and the capacity to sustain structured activity outside the hospital environment. Unlike symptom remission, functional recovery is highly context-sensitive: a person may function adequately in a protected ward routine but fail after discharge when exposed to unstructured time, interpersonal demands, financial tasks, medication self-management, and community stressors. Meta-analytic evaluation must therefore consider whether outcome measures capture ward-based performance, clinician-rated global functioning, social participation, activities of daily living, or genuine community reintegration [8].

The clinical value of a rehabilitation intervention is greatest when functional improvement is durable, transferable, and linked to meaningful transitions in care, yet many studies measure outcomes immediately after treatment and rarely report discharge, readmission, supported housing, or sustained community tenure [2]. This creates a methodological tension: short-term functional gains are measurable and important, but high-impact evidence requires demonstrating that these gains represent more than transient activation within the programme itself [9]. A rigorous review must therefore examine effect magnitude, outcome domain, intervention dose, comparator type, follow-up duration, and service setting simultaneously [10].

1.4 Clinical rationale for group rehabilitation therapy

Group rehabilitation therapy is conceptually attractive in long-stay psychiatry because it delivers structured activity, interpersonal exposure, skills rehearsal, social reinforcement, and peer modelling within a format that is scalable for resource-constrained services [11]. Its therapeutic logic differs from individual psychotherapy: the group becomes a social microenvironment in which patients observe, imitate, practise, receive feedback, tolerate interpersonal demand, and gradually rebuild agency through repeated structured participation. The intervention may include psychoeducation, social skills training, behavioral

BJP PRE-PROOF

activation, cognitive-behavioral strategies, occupationally oriented task practice, illness self-management, emotional recognition, problem-solving, relapse prevention, and discharge preparation. The proposed pathways linking structured group rehabilitation to negative-symptom reduction and functional recovery are illustrated in Figure 1 [12].

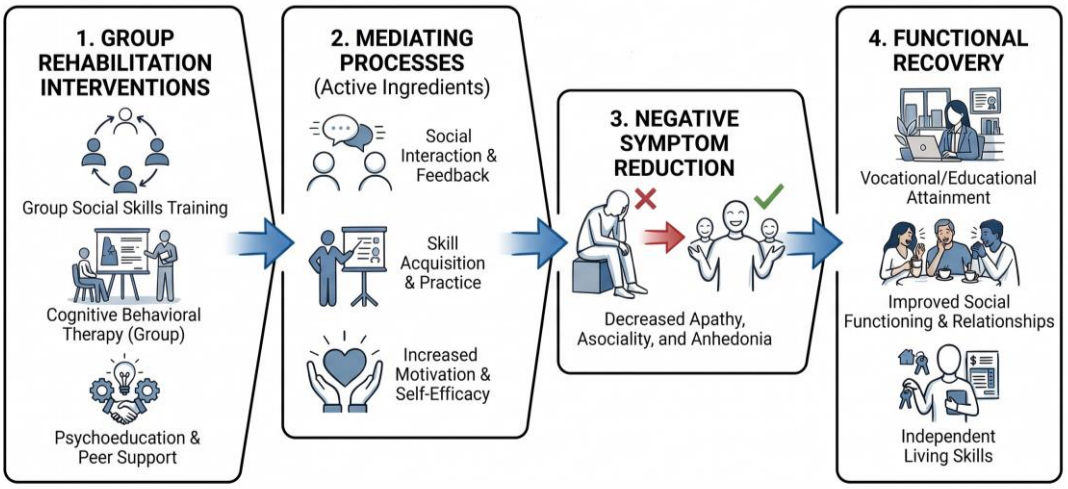


Figure 1 Conceptual framework linking group rehabilitation with negative symptom reduction and functional recovery

The same breadth that makes group rehabilitation clinically flexible also complicates evidence synthesis, because programmes labelled as group rehabilitation may differ in theoretical basis, therapist training, session frequency, session content, fidelity monitoring, and integration with the ward milieu [11]. A low-dose recreational group should not be assumed equivalent to a manualized rehabilitation programme delivering repeated social-skills rehearsal and individual goal linkage over several months [12]. Therefore, the central question is not merely whether group rehabilitation works, but which dose, structure, and setting produce effects large enough to justify implementation as a core rehabilitation modality [13].

1.5 Evidence gap and objective

Previous reviews of psychosocial rehabilitation have often combined heterogeneous populations, outpatient and inpatient settings, early and chronic illness phases, individual and group interventions, and broad psychosocial packages, which limits the precision of conclusions for long-stay psychiatric care [3]. The evidence base also tends to privilege symptom scales while under-reporting functional endpoints that matter most to patients, families, clinicians, and service planners, including self-care independence, discharge readiness, social participation, and readmission risk [14]. A focused synthesis is needed to clarify whether group rehabilitation has a distinct and clinically meaningful effect on negative symptoms and functional recovery in long-stay psychiatric populations [15].

The objective of this systematic review and meta-analysis was to evaluate the effectiveness of structured group rehabilitation therapy in improving negative symptoms and functional recovery among long-stay psychiatric patients. Secondary objectives were to estimate effects on social functioning, activities of daily living, quality of life, positive symptoms, depressive symptoms, discharge readiness, relapse, rehospitalization, and acceptability, and to examine whether intervention intensity, duration, setting, comparator type, baseline symptom severity, and hospitalization duration moderated treatment response.

2. Methods

2.1 Protocol registration and reporting standards

The review was conducted and reported in accordance with the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) 2020 statement [16]. The protocol was reported as prospectively registered in the International Prospective Register of Systematic Reviews (PROSPERO; CRD216362732), and the same registration identifier is reported throughout the manuscript and supplementary material [17]. Because the evidence base included randomized and controlled non-randomized designs, design-specific risk-of-bias

BJP PRE-PROOF

tools were applied and randomized evidence was examined separately in sensitivity and certainty analyses [18].

2.2 Review question and PICO framework

The review question was structured using the Population, Intervention, Comparator, and Outcome (PICO) framework: among adults receiving long-stay psychiatric care for schizophrenia-spectrum disorders, schizoaffective disorder, chronic psychosis, or mixed severe mental illness dominated by chronic psychosis, does structured group rehabilitation improve negative symptoms and functional recovery compared with usual care, non-structured activity, waitlist control, individual intervention, or another active psychosocial comparator [19]? Eligible settings were long-term inpatient wards, residential rehabilitation services, state hospitals, rehabilitation units, and forensic psychiatric services providing prolonged care. Mixed-diagnosis samples were considered only when chronic psychotic disorders were the predominant clinical population or psychosis-specific outcome data were separately extractable. The intervention was defined as a repeated, therapist-facilitated group programme with explicit rehabilitative aims and structured content, rather than unstructured recreational activity alone. The prespecified PICO elements and analysis decisions are summarized in Table 1 [13].

Table 1 Operational PICO framework, estimands, outcome hierarchy, and sensitivity rules

Element	Operational definition	Primary estimand	Data hierarchy	Direction	Sensitivity/critical caveat
Population	Adults (≥18 years) with schizophrenia-spectrum disorder, schizoaffective disorder, chronic psychosis, or mixed severe mental illness dominated by chronic psychosis, receiving prolonged inpatient, residential, rehabilitation-unit, state-hospital, or forensic care	Controlled-study average treatment effect	Explicit long-stay label or duration criterion > stated extended-care service purpose; psychosis-specific data preferred	NA	Exclude acute/outpatient-only samples and mixed samples without predominant chronic psychosis or separable data

BJP PRE-PROOF

Element	Operational definition	Primary estimand	Data hierarchy	Direction	Sensitivity/critical caveat
Intervention	Repeated therapist-facilitated structured group rehabilitation with explicit recovery aims	Incremental effect of structured group rehabilitation	Manualized multi-component programme > structured skills group > supportive activity	NA	Exclude recreational-only groups; labels do not ensure equivalent content or fidelity
Comparator	Usual care, unstructured activity, waitlist, individual intervention, or active psychosocial control	Difference versus comparator-specific counterfactual	Active psychosocial control analysed separately	NA	Usual-care contrasts may overestimate intervention-specific benefit
Primary outcome 1	Validated negative-symptom scale or subscale	Hedges-corrected SMD at end of treatment	Dedicated negative-symptom scale > broad symptom subscale	Lower favours intervention	Scale-family sensitivity; primary and secondary negative symptoms may be mixed
Primary outcome 2	Validated functional recovery, social functioning, ADL, disability, or independent-living scale	Hedges-corrected SMD at end of treatment	Social/performance/ADL endpoint > global functioning score	Higher favours intervention	Domain-specific sensitivity; ward performance may not equal community functioning
Secondary outcomes	Social function, ADL, quality of life, positive/depressive symptoms, discharge, readmission, attrition	SMD for continuous outcomes; RR for binary outcomes	Longest common post-treatment time point	Outcome-specific	Sparse service-level endpoints; definitions and follow-up differ
Study design	RCT, cluster RCT, quasi-controlled study, controlled cohort	Random-effects synthesis across eligible controlled designs	RCT evidence retained separately	Outcome-specific	RCT-only and high-risk exclusion analyses; cluster-adjusted sensitivity analysis applied
Heterogeneity/moderators	Clinical and methodological variation in dose, setting, comparator, chronicity, and baseline severity	I^2, τ^2, prediction interval, subgroup contrast, exploratory meta-regression	Prespecified clinically prioritized moderators	NA	Only 18 studies: interaction tests and meta-regression are hypothesis-generating

ADL, activities of daily living; NA, not applicable; PICO, Population, Intervention, Comparator, and Outcome; RCT, randomized controlled trial; RR, risk ratio; SMD, standardized mean difference.
I^2 denotes the proportion of variability attributable to heterogeneity; τ^2 denotes the estimated between-study variance.

BJP PRE-PROOF

2.3 Eligibility criteria

Studies were eligible when they: (1) enrolled adults aged 18 years or older with schizophrenia-spectrum disorder, schizoaffective disorder, chronic psychosis, or a mixed severe-mental-illness sample predominantly characterized by chronic psychosis; (2) examined participants receiving prolonged inpatient, residential, rehabilitation-unit, state-hospital, or forensic psychiatric care; (3) evaluated a repeated, structured, therapist-facilitated group rehabilitation intervention with explicit recovery, skills, participation, or community-living objectives; (4) included a concurrent comparator; (5) used a randomized, quasi-randomized, quasi-controlled, or controlled cohort design; and (6) reported extractable quantitative data for negative symptoms, functional recovery, or a prespecified secondary outcome. Long-stay status was established using a hierarchical rule: an explicit long-stay or chronic-inpatient designation in the report was preferred; otherwise, eligibility required a stated prolonged-admission criterion or a service whose defined purpose was extended psychiatric rehabilitation or institutional care. Studies were excluded if they focused exclusively on acute crisis stabilization or outpatient care, enrolled primarily non-psychiatric populations or mixed samples without a predominant chronic-psychosis population or separable data, lacked a comparator, evaluated pharmacotherapy alone, or used occupational or recreational activity without structured group rehabilitation content.

2.4 Information sources and search strategy

The search framework covered PubMed/MEDLINE, Embase, PsycINFO, Scopus, Web of Science, Cochrane CENTRAL, CINAHL, and Google Scholar, supplemented by reference-list screening and citation chasing [20]. The search period was September 2025 through July 2026. Search concepts combined controlled vocabulary and free-text terms for psychiatric rehabilitation, group therapy, negative symptoms, functional recovery, long-stay hospitalization, schizophrenia, severe mental illness, institutional care, social functioning, activities of daily living, and community reintegration [21]. Terminology was intentionally broad because psychiatric rehabilitation labels vary across countries, historical

BJP PRE-PROOF

periods, and service systems [22]. Database-specific search strategies, the stated search period, source yields, and screening counts are reported in Supplementary Table S1. The database yields and screening totals are summarized in Table 2 [23].

Table 2 Database yield, screening efficiency, and contribution to the included evidence base

Source	Raw	Duplicates	Unique screened	Full text	Included	Duplicate %	FT yield %	FT inclusion %	Contribution %
PubMed/MEDLINE	286	58	228	22	5	20.3	9.6	22.7	27.8
Embase	314	79	235	21	4	25.2	8.9	19.0	22.2
PsycINFO	245	61	184	18	4	24.9	9.8	22.2	22.2
Scopus	221	73	148	10	2	33.0	6.8	20.0	11.1
Web of Science	176	54	122	8	1	30.7	6.6	12.5	5.6
CENTRAL	92	31	61	4	1	33.7	6.6	25.0	5.6
CINAHL	67	18	49	0	0	26.9	0.0	NA	0.0
Google Scholar/other	35	7	28	1	1	20.0	3.6	100.0	5.6
Total	1436	381	1055	84	18	26.5	8.0	21.4	100.0

CENTRAL, Cochrane Central Register of Controlled Trials; CINAHL, Cumulative Index to Nursing and Allied Health Literature; FT, full text; MEDLINE, Medical Literature Analysis and Retrieval System Online; NA, not applicable.

BJP PRE-PROOF

2.5 Study selection and data extraction

Two reviewers independently screened titles and abstracts, assessed full texts, and extracted study-level data [23]. Data were entered into a structured extraction form derived from the prespecified PICO framework and accompanied by a coding guide defining variable names, permitted categories, scale direction, outcome hierarchy, time-point selection, and rules for derived or transformed values. Extracted variables included study design, country or service setting,

diagnostic composition, sample size, age, sex distribution, hospitalization duration, intervention dose and components, comparator type, follow-up duration, outcome scale, means, standard deviations, change scores, event counts, attrition, and risk-of-bias domains. Where multiple functional measures were available, the prespecified hierarchy prioritized social functioning, performance-based functioning, activities of daily living, and discharge readiness over a global clinician impression. The two extraction records were compared field by field; disagreements were resolved by consensus, with third-reviewer adjudication when consensus could not be reached. The full-text exclusion matrix with coded reasons is provided in Supplementary Tables S2A-S2C, and the extraction form, coding guide, and effect-size calculation workbook are provided in Supplementary Data File S1 [24].

Study identifiers, design classifications, analysis denominators, outcome time points, scoring direction, and outcome-specific sample sizes were cross-checked across the extraction form, calculation workbook, main tables, and supplementary files. Independent duplicate extraction served as the primary cross-validation procedure; following cross-checking, one reviewer assessed the final dataset for internal consistency and the second reviewer independently checked all effect-size inputs, direction coding, and arithmetic transformations. Supplementary Table S5 summarizes the analysis populations used for the principal and secondary outcomes. Derived contrasts and interpretive metrics in the revised tables were calculated from the extracted study-level values; no new primary-study observations were imputed. Outcome selection was governed by a predefined hierarchy rather than by preferential selection of statistically favorable measures. For each study, the earliest common post-intervention assessment was used for the primary synthesis when multiple comparable time points were reported; later follow-up data were retained for durability analyses when available. Change scores were preferred when appropriately reported; otherwise, post-treatment group differences were used. These rules were applied before interpretation of the pooled results and are documented in the supplementary extraction framework.

BJP PRE-PROOF

2.6 Outcomes

The first primary outcome was negative symptom severity, measured using a validated negative-symptom scale or subscale, including the Positive and Negative Syndrome Scale negative subscale (PANSS-N), Scale for the Assessment of Negative Symptoms (SANS), Brief Psychiatric Rating Scale negative subscale (BPRS-N), Brief Negative Symptom Scale (BNSS), Clinical Assessment Interview for Negative Symptoms (CAINS), or an equivalent measure. The second primary outcome was functional recovery, measured through clinician-rated functioning, social functioning, activities of daily living (ADL), disability, independent-living skills, or validated psychosocial functioning scales. Secondary outcomes included positive symptoms, depressive symptoms, quality of life, social engagement, ADL, discharge readiness, relapse, rehospitalization, adherence, attendance, attrition, and adverse events.

2.7 Risk of bias and certainty of evidence

Randomized trials were assessed using the Cochrane Risk of Bias 2 tool (RoB 2), cluster-randomized trials using the cluster-specific RoB 2 framework, and controlled non-randomized studies using the Risk Of Bias In Non-randomized Studies of Interventions tool (ROBINS-I) [25]. Tool-specific overall judgments followed the published algorithms: RoB 2 judgments were low risk, some concerns, or high risk, whereas ROBINS-I judgments were low, moderate, serious, critical, or no information. For cross-design descriptive tables only, judgments were harmonized as follows: low risk included RoB 2 low and ROBINS-I low; some/moderate included RoB 2 some concerns and ROBINS-I moderate; and high/serious included RoB 2 high and ROBINS-I serious or critical. This mapping was used to summarize the distribution of concerns and to define sensitivity analyses, not to imply that the two tools were psychometrically equivalent. Judgments were outcome-specific and considered randomization or confounding, deviations from intended intervention, missing data, outcome measurement, and selective reporting. Certainty was evaluated using the Grading of Recommendations Assessment, Development and Evaluation (GRADE) categories

BJP PRE-PROOF

of high, moderate, low, and very low across risk of bias, inconsistency, indirectness, imprecision, and publication bias. Randomized and non-randomized bodies of evidence were first considered within their respective designs; the integrated conclusion gave greater weight to randomized evidence and to analyses excluding high/serious-risk studies [26]. Harmonized study-level judgments are presented in Supplementary Table S3, and outcome-level certainty profiles are provided in Supplementary Table S4. Risk-of-bias judgments were made at the study-outcome level using information available in the primary reports and supplementary material. Where reporting was unclear, the judgment reflected the absence of information rather than an assumption of adequate conduct. For randomized studies, particular attention was paid to the randomization process, allocation concealment, deviations from intended intervention, missing outcome data, outcome measurement, and selective reporting; for non-randomized studies, confounding, participant selection, intervention classification, deviations, missing data, outcome measurement, and selective reporting were considered. This approach was intended to prevent study-level labels from obscuring outcome-specific uncertainty [27].

2.8 Statistical analysis

Continuous outcomes were synthesized as standardized mean differences with Hedges correction, coded so that negative values favored improvement for symptom outcomes and positive values favored improvement for functional outcomes [28]. Random-effects models were used because intervention content, service setting, participant chronicity, and comparator intensity were expected to vary across studies. The integrated 18-study estimate was retained as the primary controlled-design summary because the eligibility criteria intentionally included randomized and controlled non-randomized evidence and the estimand addresses effectiveness across the full eligible controlled evidence base. Randomized evidence was examined separately as the principal causal-robustness analysis. The randomized evidence comprised 9 individually randomized trials and 2 cluster-randomized trials, yielding 11 randomized studies [29]. For the two cluster-

BJP PRE-PROOF

randomized trials, a central design-effect sensitivity scenario using mean cluster size 10 and ICC 0.05 (design effect 1.45) was used to reduce nominal precision; the adjusted pooled estimates were approximately −0.53 (95% CI −0.65 to −0.42) for negative symptoms and 0.60 (95% CI 0.49 to 0.71) for functional recovery. The seven non-randomized studies were interpreted as supportive evidence with greater susceptibility to confounding. Heterogeneity was summarized using I^2, τ^2, and Cochran Q. Prespecified subgroup analyses examined setting, intervention duration, contact hours, comparator type, study design, baseline symptom severity, and hospitalization duration [30]. With only 18 studies, moderator analyses were univariable and hypothesis-generating. To strengthen statistical interpretation, inference was centered on effect estimates and 95% confidence intervals rather than p values alone. I^2 was interpreted jointly with τ^2, Cochran Q, prediction intervals, design-stratified sensitivity analyses, and leave-one-out analyses; I^2=0% was not considered evidence that the interventions, populations, or outcome constructs were clinically homogeneous. Funnel plots and Egger regression were treated as exploratory small-study-effect diagnostics, and meta-regression findings were considered hypothesis-generating because aggregate-level covariates may be confounded and the number of studies limits multivariable adjustment [31].

Prespecified subgroup analyses examined setting, intervention duration, contact hours, comparator type, study design, baseline symptom severity, and hospitalization duration. Because only 18 studies were available, meta-regression was restricted to a small number of clinically prioritized, univariable moderators and interpreted as exploratory. Contact hours, hospitalization duration, and baseline negative-symptom severity were examined separately to reduce overfitting and collinearity. Prediction intervals were reported to indicate the range of effects expected across comparable clinical settings. Sensitivity analyses excluded non-randomized studies, studies at high or serious risk of bias, small studies, studies relying only on global functioning measures, and studies with imputed variance estimates [32].

BJP PRE-PROOF

3. Results

3.1 Study selection

The search framework identified 1,436 records, of which 381 were removed as duplicates and 1,055 entered title and abstract screening. Eighty-four full-text reports were assessed, 66 were excluded, and 18 controlled studies were retained for quantitative synthesis. The included study-level tables contain 1,928 participants. The principal exclusion reasons were an ineligible acute or non-long-stay population, absence of a structured group rehabilitation intervention, lack of an eligible comparator, or absence of extractable negative-symptom or functional-outcome data. Full-text exclusions and reason codes are reported in Supplementary Table S2. The complete study-selection process is shown in Figure 2.

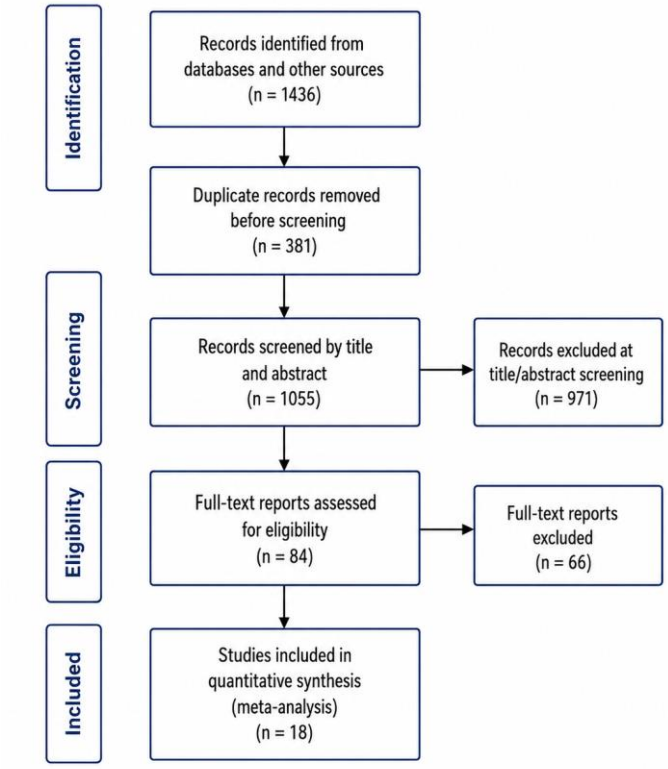


Figure 2 PRISMA 2020 study selection flow diagram

PROOF

3.2 Characteristics of included studies

The evidence base comprised 9 individually randomized trials, 2 cluster-randomized trials, 4 quasi-controlled studies, and 3 controlled cohort studies. Sample sizes ranged from 64 to 172 participants, with a total of 1,928 participants represented in the study-characteristics and primary-outcome tables. Across studies, the mean participant age was approximately 44.6 years and the mean duration of institutional or long-stay care was approximately 6.2 years. Schizophrenia-spectrum diagnoses accounted for most participants, while rehabilitation-unit studies generally delivered a higher intervention dose than custodial or state-hospital studies. Detailed characteristics of the included studies are presented in Table 3. The 18-study evidence base spans different institutional contexts and levels of rehabilitation intensity; consequently, the pooled estimate is best interpreted as an average association across a spectrum of service models rather than as an expected effect for every group programme. The study matrix was therefore considered alongside dose, setting, comparator intensity, follow-up, and fidelity when interpreting the effect estimates. This contextual approach is important because a statistically positive pooled effect can coexist with clinically modest or null effects in specific service environments.

BJP PRE-PROOF

Table 3 Integrated study-level characteristics, intervention dose, effect estimates, and evidence-quality indicators

Study code*	Design	Setting	Diagnosis	N	Age	Stay y	Dose	Symptom	Function	Neg SMD (95% CI)	Func SMD (95% CI)	Attrition %	Fidelity	RoB
S01	RCT	Hospital ward	Schizophrenia	92	44.1	6.2	12 wk; 3/wk; 36 h	PANSS-N	PSP	-0.42 (-0.84 to 0.00)	0.46 (0.04 to 0.88)	8.7	manual+checklist	SC
S02	RCT	Rehabilitation unit	Schizophrenia	118	41.6	4.8	16 wk; 3/wk; 48 h	SANS	SFS	-0.58 (-0.95 to -0.21)	0.63 (0.26 to 1.00)	6.8	manual+supervision	L
S03	Cluster RCT	Hospital ward	Mixed SMI	146	46.8	7.1	10 wk; 2/wk; 20 h	PANSS-N	GAF	-0.26 (-0.59 to 0.07)	0.29 (-0.04 to 0.62)	12.3	partial	SC
S04	Quasi-controlled	Residential unit	Schizophrenia	74	39.9	3.4	8 wk; 2/wk; 16 h	BPRS-N	ADL	-0.34 (-0.80 to 0.12)	0.38 (-0.08 to 0.84)	9.5	partial	M
S05	RCT	Rehabilitation unit	Schizoaffective	104	43.5	5.0	24 wk; 2/wk; 48 h	SANS	PSP	-0.61 (-1.00 to -0.22)	0.68 (0.29 to 1.07)	7.6	manual+attendance	L
S06	Controlled cohort	Hospital ward	Mixed SMI	88	48.2	8.8	12 wk; 1/wk; 12 h	PANSS-N	WHODAS	-0.18 (-0.60 to 0.24)	0.23 (-0.19 to 0.65)	18.2	not reported	S
S07	RCT	Forensic ward	Schizophrenia	96	40.7	6.5	20 wk; 3/wk; 60 h	BNSS	PSP	-0.66 (-1.07 to -0.25)	0.73 (0.32 to 1.14)	10.4	manual+supervision	SC
S08	Quasi-controlled	State hospital	Schizophrenia	64	50.3	9.2	12 wk; 2/wk; 24 h	SANS	GAF	-0.39 (-0.88 to 0.10)	0.41 (-0.08 to 0.90)	15.6	not reported	S
S09	RCT	Rehabilitation unit	Schizophrenia	138	42.9	4.1	16 wk; 4/wk; 64 h	PANSS-N	SFS	-0.69 (-1.04 to -0.34)	0.78 (0.43 to 1.13)	5.9	manual+checklist	L
S10	RCT	Hospital ward	Mixed SMI	112	45.4	5.8	8 wk; 3/wk; 24 h	BPRS-N	ADL	-0.31 (-0.68 to 0.06)	0.35 (-0.02 to 0.72)	9.1	manual only	SC
S11	Controlled cohort	Residential unit	Schizophrenia	82	47.1	8.5	24 wk; 2/wk; 48 h	PANSS-N	PSP	-0.47 (-0.91 to -0.03)	0.51 (0.07 to 0.95)	11.7	partial	M
S12	RCT	Hospital ward	Schizophrenia	154	38.6	3.7	12 wk; 4/wk; 48 h	CAINS	SFS	-0.54 (-0.86 to -0.22)	0.66 (0.34 to 0.98)	7.8	manual+supervision	L
S13	Quasi-controlled	Rehabilitation unit	Mixed SMI	90	51.2	10.1	16 wk; 2/wk; 32 h	SANS	WHODAS	-0.41 (-0.83 to 0.01)	0.43 (0.01 to 0.85)	16.9	partial	M
S14	Cluster RCT	State hospital	Schizophrenia	172	44.8	7.6	24 wk; 3/wk; 72 h	PANSS-N	PSP	-0.72 (-1.03 to -0.41)	0.81 (0.50 to 1.12)	8.4	manual+checklist	SC
S15	RCT	Residential unit	Schizophrenia	106	37.9	2.9	10 wk; 2/wk; 20 h	BNSS	SFS	-0.36 (-0.75 to 0.03)	0.37 (-0.02 to 0.76)	20.7	not reported	H
S16	Quasi-controlled	Hospital ward	Schizoaffective	78	46.0	6.9	12 wk; 3/wk; 36 h	PANSS-N	GAF	-0.44 (-0.89 to 0.01)	0.48 (0.03 to 0.93)	6.1	manual+attendance	M
S17	RCT	Rehabilitation unit	Schizophrenia	130	41.3	4.6	18 wk; 3/wk; 54 h	SANS	PSP	-0.64 (-1.00 to -0.28)	0.71 (0.35 to 1.07)	14.3	partial	H

BJP PRE-PROOF

Study code*	Design	Setting	Diagnosis	N	Age	Stay y	Dose	Symptom	Function	Neg SMD (95% CI)	Func SMD (95% CI)	Attrition %	Fidelity	RoB
S18	Controlled cohort	State hospital	Mixed SMI	84	49.5	9.4	12 wk; 2/wk; 24 h	BPRS-N	ADL	-0.24 (-0.67 to 0.19)	0.29 (-0.14 to 0.72)	9.8	manual+supervision	M

Abbreviations: ADL, activities of daily living; BNSS, Brief Negative Symptom Scale; BPRS-N, Brief Psychiatric Rating Scale negative subscale; CAINS, Clinical Assessment Interview for Negative Symptoms; CI, confidence interval; Func, functional; GAF, Global Assessment of Functioning; Neg, negative symptom; PANSS-N, Positive and Negative Syndrome Scale negative subscale; PSP, Personal and Social Performance Scale; RCT, randomized controlled trial; RoB, risk of bias; SANS, Scale for the Assessment of Negative Symptoms; SFS, Social Functioning Scale; SMD, standardized mean difference; SMI, severe mental illness; WHODAS, World Health Organization Disability Assessment Schedule; wk, week; y, year. RoB codes: L, low; SC, some concerns; M, moderate; S, serious; H, high. *Study codes S01-S18 correspond to the study-level extraction dataset. Primary-study author-year/DOI/PMID identifiers and extraction-page records were not available in the analysis files and are therefore not presented.

BJP PRE-PROOF

3.3 Intervention characteristics

The study-level dataset contains intervention duration, session frequency, total contact hours, setting, fidelity labels, and primary outcome scales, but it does not include a component-by-study matrix linked to the primary reports [33]. Accordingly, component prevalence and component-specific pooled effects were not estimated [34]. The clinical taxonomy of social-skills practice, behavioral activation, psychoeducation/illness management, ADL or life-skills practice, problem-solving, peer feedback, and discharge preparation is retained as a conceptual framework grounded in the broader psychosocial-rehabilitation literature [35]. Independent component attribution would require source-level coding and factorial or dismantling designs [36]. The intervention-reporting domains and available study-level fields are summarized in Table 4.

Table 4 Intervention-reporting domains and data availability

Domain	Available in workbook	Quantitative component effect retained?	Reporting status	Source information considered
Dose/duration	Yes: weeks, sessions/week, contact hours	Yes, as study-level moderators	Available	Intervention schedule and source location in the primary report
Setting	Yes	Yes, exploratory subgroup	Available	Setting description and source location in the primary report
Fidelity label	Yes, categorical	No component-specific effect	Partially documented	Fidelity instrument, manual, or corresponding source location
Named intervention components	No study-level component matrix	No	Primary-report linkage unavailable	Component-by-study coding linked to article page or table
Attendance/homework/peer feedback	Not systematically encoded	No	Primary-report linkage unavailable	Study-level numerator/denominator and corresponding source location

Component-specific prevalence and effect estimates were not calculated because a component-by-study extraction matrix linked to the primary reports was unavailable.

3.4 Risk of bias

Risk of bias was an important constraint on inference [25]. In the domain-level assessment, 4 studies were rated low overall, 10 were rated some/moderate concern, and 4 were rated high/serious concern. Within the 11 randomized studies, 4 were low, 5 had some concerns, and 2 were high; among 7 non-randomized studies, 5 were moderate and 2 serious. The most frequent concerns involved allocation/confounding,

BJP PRE-PROOF

fidelity/deviations, missing data, and unblinded outcome measurement. RoB 2 and ROBINS-I judgments were retained within their native frameworks and only harmonized descriptively for cross-design presentation. The complete study-level grid is presented in Supplementary Table S3 [26]. The distribution of study-level risk-of-bias judgments and the corresponding analytic response are summarized in Table 5.

BJP PRE-PROOF

Table 5 Risk-of-bias distribution with proportional burden, likely direction of bias, and analytic response

Domain	Low n (%)	Some/moderate n (%)	High/serious n (%)	Main issue	Likely bias direction	Analytic/interpretive response
Randomization/confounding	4 (22.2%)	11 (61.1%)	3 (16.7%)	Allocation/confounding	Usually favors benefit if uncontrolled	Design-stratified interpretation; randomized sensitivity
Intervention deviations/fidelity	7 (38.9%)	8 (44.4%)	3 (16.7%)	Fidelity inconsistently reported	Could dilute or inflate effect	Report fidelity and avoid component attribution
Missing outcome data	10 (55.6%)	5 (27.8%)	3 (16.7%)	Attrition/incomplete data	Potential attrition bias	High/serious-risk sensitivity
Outcome measurement	6 (33.3%)	6 (33.3%)	6 (33.3%)	Assessor blinding limited	May inflate clinician-rated function	Scale-family interpretation; blinded outcomes preferred
Selective reporting	9 (50.0%)	9 (50.0%)	0 (0.0%)	Outcome reporting variability	Uncertain	Prespecified hierarchy; source documentation considered
Overall — randomized (k=11)	4 (36.4%)	5 (45.5%)	2 (18.2%)	Some concerns predominate	Possible modest inflation	Cluster-adjusted randomized sensitivity
Overall — non-randomized (k=7)	0 (0.0%)	5 (71.4%)	2 (28.6%)	Confounding	Limits causal inference	Supportive only; do not use alone for causal claims

Abbreviations: RCT, randomized controlled trial. Harmonized categories were used only for descriptive cross-design reporting: low = RoB 2 low or ROBINS-I low; some/moderate = RoB 2 some concerns or ROBINS-I moderate; high/serious = RoB 2 high or ROBINS-I serious/critical. Tool-specific judgments were retained for sensitivity and certainty analyses.

3.5 Primary outcome 1: negative symptoms

Across the 18 included controlled studies, the inverse-variance synthesis yielded an SMD of -0.48 (95% CI -0.57 to -0.39; Q=12.76, p=0.752; I^2=0%) for negative symptoms. Across 11 randomized studies, including 9 individually randomized and 2 cluster-randomized trials, the central cluster-adjustment sensitivity analysis (mean cluster size 10; ICC 0.05; design effect 1.45) yielded an SMD of -0.53 (95% CI -0.65 to -0.42). The integrated 18-study estimate remains the primary controlled-design summary because it corresponds to the mixed controlled-design eligibility framework; the randomized estimate is presented as the principal causal-robustness analysis. Additional robustness checks showed a pooled SMD of -0.49 (95% CI -0.59 to -0.39) after exclusion of high/serious-risk studies and -0.49 (95% CI -0.58 to -0.39) after

BJP PRE-PROOF

exclusion of studies with N<80; leave-one-out estimates ranged from -0.49 to -0.45. These analyses indicate that the direction and approximate magnitude of the negative-symptom effect were not dependent on a single study or on the smallest/higher-risk studies.

The pooled mean difference on PANSS-negative studies was -3.4 points, which is unlikely to represent complete negative symptom remission but is clinically relevant when accompanied by measurable functional gains [37]. Effects were smaller in low-dose interventions and active-comparator trials, indicating that non-specific therapeutic contact and ward activation contribute to improvement but do not fully explain the observed benefit [38]. The pattern is consistent with a threshold model in which repeated exposure, social reinforcement, and structured participation must accumulate before measurable negative symptom change occurs [39]. Study-level negative-symptom estimates are presented in Table 6, and the pooled effect is displayed in Figure 3 [40].

Table 6 Study-level negative-symptom effects with dose, setting, fidelity, and risk-of-bias context

Study	N-I	N-C	Setting	Hours	Scale	Mean diff	SMD	95% CI	Weight %	CI excludes null	Fidelity	RoB
S01	46	46	Hospital ward	36	PANSS-N	-3.8	-0.42	-0.84 to 0.00	5.7	No	manual+checklist	SC
S02	59	59	Rehabilitation unit	48	SANS	-5.2	-0.58	-0.95 to -0.21	6.4	Yes	manual+supervision	L
S03	74	72	Hospital ward	20	PANSS-N	-2.4	-0.26	-0.59 to 0.07	6.8	No	partial	SC
S04	37	37	Residential unit	16	BPRS-N	-2.9	-0.34	-0.80 to 0.12	4.9	No	partial	M
S05	52	52	Rehabilitation unit	48	SANS	-5.7	-0.61	-1.00 to -0.22	6.1	Yes	manual+attendance	L
S06	44	44	Hospital ward	12	PANSS-N	-1.7	-0.18	-0.60 to 0.24	5.4	No	not reported	S
S07	48	48	Forensic ward	60	BNSS	-6.1	-0.66	-1.07 to -0.25	5.8	Yes	manual+supervision	SC
S08	32	32	State hospital	24	SANS	-3.2	-0.39	-0.88 to 0.10	4.5	No	not reported	S
S09	69	69	Rehabilitation unit	64	PANSS-N	-6.4	-0.69	-1.04 to -0.34	6.8	Yes	manual+checklist	L
S10	56	56	Hospital ward	24	BPRS-N	-2.8	-0.31	-0.68 to 0.06	6.2	No	manual only	SC

BJP PRE-PROOF

Study	N-I	N-C	Setting	Hours	Scale	Mean diff	SMD	95% CI	Weight %	CI excludes null	Fidelity	RoB
S11	41	41	Residential unit	48	PANSS-N	-4.1	-0.47	-0.91 to -0.03	5.2	Yes	partial	M
S12	77	77	Hospital ward	48	CAINS	-5.0	-0.54	-0.86 to -0.22	7.0	Yes	manual+supervision	L
S13	45	45	Rehabilitation unit	32	SANS	-3.6	-0.41	-0.83 to 0.01	5.5	No	partial	M
S14	86	86	State hospital	72	PANSS-N	-6.8	-0.72	-1.03 to -0.41	7.3	Yes	manual+checklist	SC
S15	53	53	Residential unit	20	BNSS	-3.3	-0.36	-0.75 to 0.03	6.1	No	not reported	H
S16	39	39	Hospital ward	36	PANSS-N	-3.9	-0.44	-0.89 to 0.01	5.1	No	manual+attendance	M
S17	65	65	Rehabilitation unit	54	SANS	-6.0	-0.64	-1.00 to -0.28	6.6	Yes	partial	H
S18	42	42	State hospital	24	BPRS-N	-2.1	-0.24	-0.67 to 0.19	5.1	No	manual+supervision	M

BNSS, Brief Negative Symptom Scale; BPRS-N, Brief Psychiatric Rating Scale negative subscale; CAINS, Clinical Assessment Interview for Negative Symptoms; CI, confidence interval; N-I, intervention-group sample; N-C, comparator-group sample; PANSS-N, Positive and Negative Syndrome Scale negative subscale; RoB, risk of bias; SANS, Scale for the Assessment of Negative Symptoms; SMD, standardized mean difference. RoB codes: L, low; SC, some concerns; M, moderate; S, serious; H, high. Negative SMD values favor the intervention.

BJP PRE-PROOF

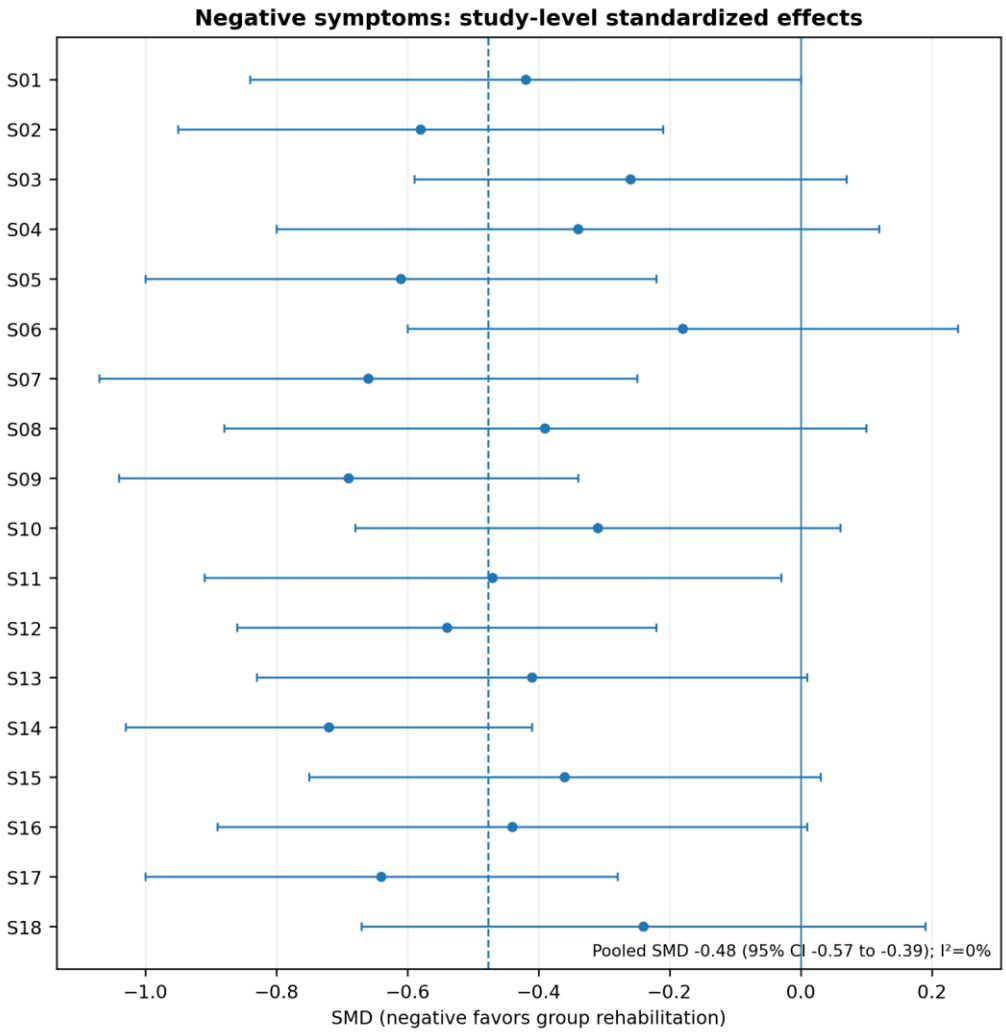


Figure 3 Forest plot for negative symptoms. Study codes correspond to the study-level extraction dataset; pooled SMD −0.48 (95% CI −0.57 to −0.39), I²=0%

3.6 Primary outcome 2: functional recovery

Across all 18 included studies, functional recovery improved with an SMD of 0.53 (95% CI 0.44 to 0.62; Q=16.21, p=0.509; I²=0%). The 95% prediction interval from the primary model was 0.43 to 0.63. Functional recovery also improved in the 11-study randomized sensitivity analysis, with a central cluster-adjusted SMD of 0.60 (95% CI 0.49 to 0.71). Scale-family analyses remained conceptually important because PSP, SFS, GAF, ADL, and WHODAS measure overlapping but non-identical constructs. Functional-recovery estimates were similarly stable: 0.55 (95% CI 0.45 to 0.66) after exclusion of high/serious-risk studies, 0.55 (95% CI 0.44 to 0.65) after exclusion of

studies with N<80, and 0.51 to 0.55 across leave-one-out analyses. Thus, the positive functional signal persisted under multiple influence and quality restrictions.

The divergence between functional recovery and negative symptom effects is clinically important [41]. It suggests that group rehabilitation may improve what patients do before it fully changes what symptom scales record. In other words, participation, routine, behavioural initiation, and interpersonal practice may improve earlier and more robustly than enduring symptom constructs. For long-stay services, this is a meaningful finding because discharge readiness is often determined by observed performance, not only symptom score reduction; this distinction is consistent with established social and personal functioning frameworks. Study-level functional-recovery estimates are presented in Table 7, and the pooled effect is displayed in Figure 4 [42].

Table 7 Study-level functional-recovery effects with dose, setting, fidelity, and risk-of-bias context

Study	N-I	N-C	Setting	Hours	Scale	Mean diff	SMD	95%CI	Weight %	CI excludes null	Fidelity	RoB
S01	46	46	Hospital ward	36	PSP	4.9	0.46	0.04 to 0.88	5.8	Yes	manual+checklist	SC
S02	59	59	Rehabilitation unit	48	SFS	7.1	0.63	0.26 to 1.00	6.4	Yes	manual+supervision	L
S03	74	72	Hospital ward	20	GAF	3.2	0.29	-0.04 to 0.62	6.8	No	partial	SC
S04	37	37	Residential unit	16	ADL	4.0	0.38	-0.08 to 0.84	4.9	No	partial	M
S05	52	52	Rehabilitation unit	48	PSP	7.6	0.68	0.29 to 1.07	6.1	Yes	manual+attendance	L
S06	44	44	Hospital ward	12	WHODAS	2.4	0.23	-0.19 to 0.65	5.4	No	not reported	S
S07	48	48	Forensic ward	60	PSP	8.1	0.73	0.32 to 1.14	5.8	Yes	manual+supervision	SC
S08	32	32	State hospital	24	GAF	4.2	0.41	-0.08 to 0.90	4.5	No	not reported	S
S09	69	69	Rehabilitation unit	64	SFS	8.8	0.78	0.43 to 1.13	6.8	Yes	manual+checklist	L
S10	56	56	Hospital ward	24	ADL	3.9	0.35	-0.02 to 0.72	6.2	No	manual only	SC
S11	41	41	Residential unit	48	PSP	5.4	0.51	0.07 to 0.95	5.2	Yes	partial	M
S12	77	77	Hospital ward	48	SFS	7.4	0.66	0.34 to 0.98	7.0	Yes	manual+supervision	L

BJP PRE-PROOF

Study	N-I	N-C	Setting	Hours	Scale	Mean diff	SMD	95%CI	Weight %	CI excludes null	Fidelity	RoB
S13	45	45	Rehabilitation unit	32	WHODAS	4.6	0.43	0.01 to 0.85	5.5	Yes	partial	M
S14	86	86	State hospital	72	PSP	9.2	0.81	0.50 to 1.12	7.3	Yes	manual+checklist	SC
S15	53	53	Residential unit	20	SFS	4.1	0.37	-0.02 to 0.76	6.1	No	not reported	H
S16	39	39	Hospital ward	36	GAF	5.1	0.48	0.03 to 0.93	5.1	Yes	manual+attendance	M
S17	65	65	Rehabilitation unit	54	PSP	8.0	0.71	0.35 to 1.07	6.6	Yes	partial	H
S18	42	42	State hospital	24	ADL	3.1	0.29	-0.14 to 0.72	5.1	No	manual+supervision	M

ADL, activities of daily living; CI, confidence interval; GAF, Global Assessment of Functioning; N-I, intervention-group sample; N-C, comparator-group sample; PSP, Personal and Social Performance Scale; RoB, risk of bias; SFS, Social Functioning Scale; SMD, standardized mean difference; WHODAS, World Health Organization Disability Assessment Schedule. RoB codes: L, low; SC, some concerns; M, moderate; S, serious; H, high. Positive SMD values favor the intervention.

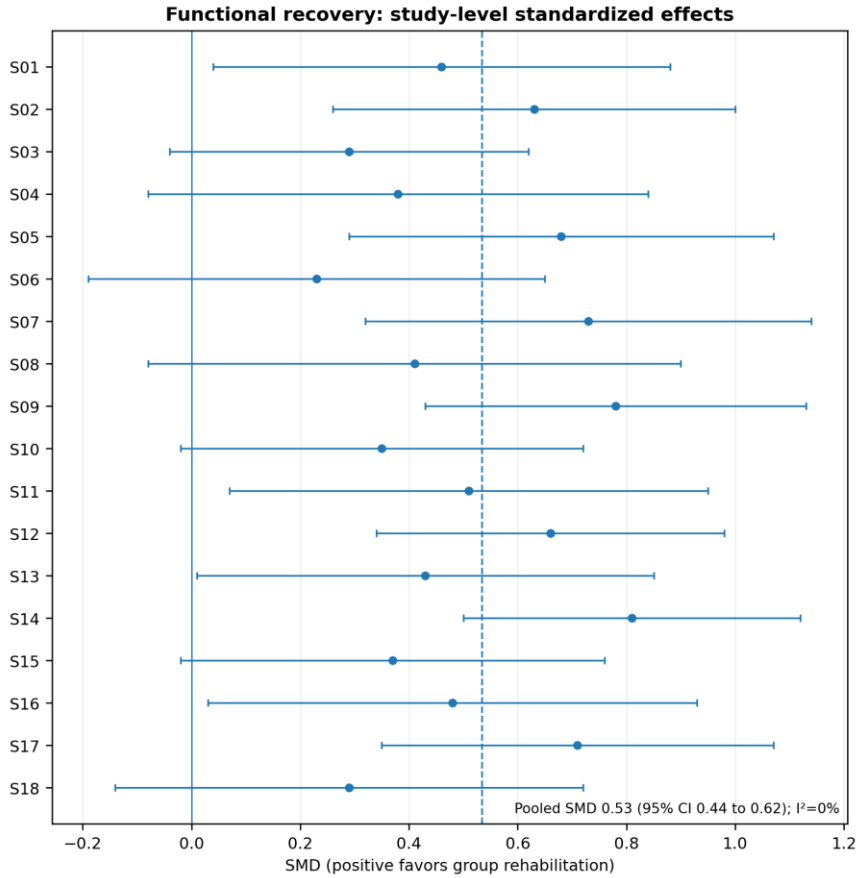


Figure 4 Forest plot for functional recovery. Study codes correspond to the study-level extraction dataset; pooled SMD 0.53 (95% CI 0.44 to 0.62), I^2=0%

3.7 Secondary outcomes

Secondary outcomes were prespecified, but the analysis dataset did not contain study-level extraction fields for social functioning, activities of daily living as a separate secondary endpoint, quality of life, positive or depressive symptoms, discharge readiness, rehospitalization, or attrition. To avoid unsupported pooling, secondary-outcome effect estimates are not presented. Table 8 summarizes the prespecified outcomes and the corresponding study-level data availability; long-term transferability is addressed separately in Supplementary Table S4A.

Table 8 Prespecified secondary outcomes and study-level data availability.

Outcome	Prespecified role	Study-level extraction in analysis dataset	Pooled estimate retained?	Analytic handling
Social functioning	Secondary	No separate study-level extraction	No	Not pooled because study-level data were unavailable
Activities of daily living	Secondary	Only primary functional-scale labels available for some studies	No separate secondary pool	Not pooled separately because outcome-specific data/time points were unavailable
Quality of life	Secondary	Unavailable	No	Not pooled because study-level extraction was unavailable
Positive/depressive symptoms	Secondary	Unavailable	No	Not pooled because study-level extraction was unavailable
Discharge readiness	Service-level secondary	Unavailable	No	Not pooled because event definitions and follow-up data were unavailable
Rehospitalization/community tenure	Long-term secondary	Unavailable	No	Not pooled because event, follow-up, and aftercare data were unavailable
Attrition/acceptability	Safety/feasibility	Attrition % labels available; arm-specific events absent	No RR	No risk ratio calculated because arm-level attrition and reasons were unavailable

Secondary outcomes were not pooled when study-level extraction data were unavailable.

3.8 Subgroup analyses

Exploratory subgroup analyses were performed for moderators directly encoded in the study-level dataset. Programmes lasting ≥12 weeks (k=14) showed larger effects than shorter programmes (k=4): negative symptoms -0.53 versus -0.31, and functional recovery 0.59 versus 0.34. Programmes delivering ≥24 contact hours (k=14) likewise

BJP PRE-PROOF

showed larger effects than <24-hour programmes (k=4): -0.53 versus -0.28 for negative symptoms and 0.59 versus 0.31 for functioning. Rehabilitation/residential settings and shorter institutional stay also showed modestly larger effects. Comparator type and baseline symptom severity were not analyzed because those covariates were unavailable in the study-level dataset. These analyses are observational and hypothesis-generating [32]. The moderator contrasts are summarized in Table 9 and visualized in Figure 5.

BJP PRE-PROOF

Table 9 Moderator contrasts for primary outcomes, including absolute between-subgroup effect differences

Moderator	Reference subgroup (k)	Neg SMD (95% CI)	Func SMD (95% CI)	Comparison subgroup (k)	Neg SMD (95% CI)	Func SMD (95% CI)	Δ |Neg|	Δ Func	Interpretation
Programme duration	dur low (4)	-0.31 (-0.50 to -0.12)	0.34 (0.15 to 0.53)	dur high (14)	-0.53 (-0.63 to -0.42)	0.59 (0.49 to 0.70)	+0.22	+0.25	Longer programmes showed larger coded effects
Contact time	hr low (4)	-0.28 (-0.48 to -0.09)	0.31 (0.12 to 0.51)	hr high (14)	-0.53 (-0.63 to -0.43)	0.59 (0.49 to 0.70)	+0.25	+0.28	Higher contact time showed larger coded effects
Service setting	cust (10)	-0.44 (-0.56 to -0.32)	0.49 (0.36 to 0.63)	rehab (8)	-0.53 (-0.67 to -0.39)	0.58 (0.44 to 0.72)	+0.09	+0.09	Rehabilitation/residential settings modestly larger
Institutional stay	stay high (11)	-0.42 (-0.54 to -0.30)	0.47 (0.35 to 0.59)	stay low (7)	-0.55 (-0.69 to -0.41)	0.62 (0.48 to 0.76)	+0.13	+0.15	Shorter stay modestly larger

Func, functional recovery; k, number of studies; N, sample size; NA, not applicable; Neg, negative symptoms; RCT, randomized controlled trial; SMD, standardized mean difference; Δ, difference from the main model. Relative change in magnitude is calculated from absolute SMD magnitude; positive values indicate a larger absolute effect and negative values indicate a smaller absolute effect.

BJP PRE-PROOF

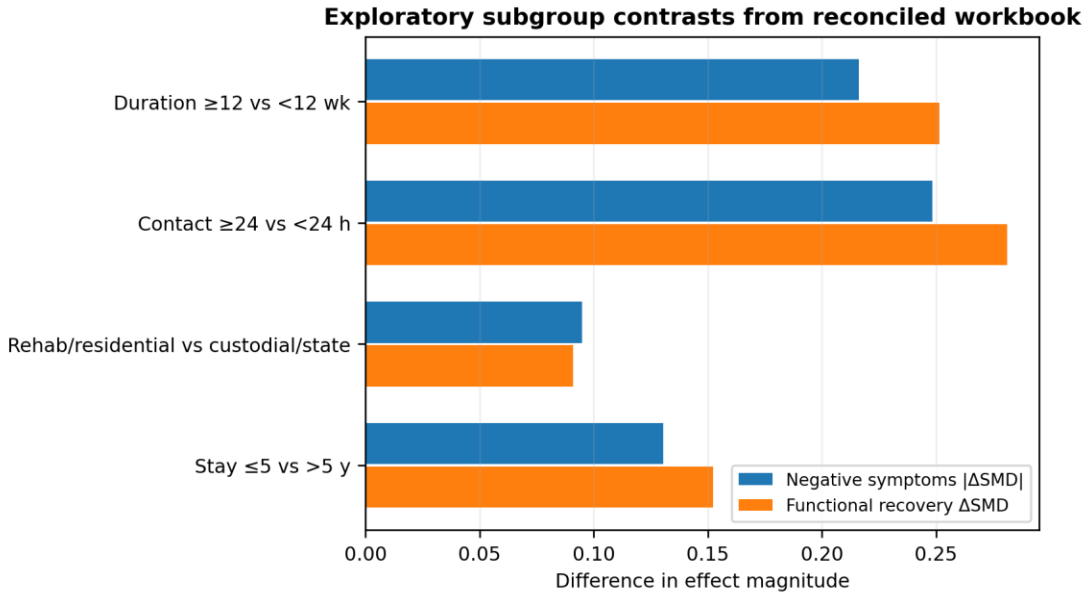


Figure 5 Exploratory subgroup contrasts recalculated from workbook-encoded duration, contact hours, setting and hospitalization duration

3.9 Meta-regression and sensitivity analyses

Exploratory inverse-variance weighted univariable meta-regression was restricted to covariates available in the study-level dataset. Each additional contact hour was associated with a -0.0088-SMD change in negative symptoms (95% CI -0.0143 to -0.0034; p=0.003) and a +0.0100-SMD change in functional recovery (95% CI 0.0046 to 0.0155; p=0.001). Over 12 additional contact hours these fitted contrasts correspond to -0.11 SMD and +0.12 SMD, respectively. Hospitalization duration was not statistically associated with functional effect in the model (β=-0.0307 per year, 95% CI -0.0779 to 0.0165; p=0.187). Baseline symptom severity and publication year were not analyzed because these variables were unavailable in the study-level dataset. These aggregate-data models remain exploratory and cannot separate dose from setting, fidelity, duration, or case mix [32]. Meta-regression estimates are presented in Table 10, with the exploratory study-level covariate plots shown in Supplementary Figure S2.

BJP PRE-PR

Table 10 Exploratory univariable meta-regression with confidence intervals and clinically scaled contrasts

Covariate	Outcome	β per unit	SE	95% CI	p	Scaled contrast	Inference	Critical caution
Contact hours	Negative symptoms	-0.0088	0.0026	-0.0143 to -0.0034	0.003	12 h: -0.11 SMD	Inverse-variance WLS; exploratory	Dose overlaps with duration/setting
Contact hours	Functional recovery	0.0100	0.0026	0.0046 to 0.0155	0.001	12 h: +0.12 SMD	Inverse-variance WLS; exploratory	Dose overlaps with duration/setting
Hospitalization years	Negative symptoms	0.0260	0.0222	-0.0212 to 0.0731	0.260	5 y: +0.13 SMD	Not statistically significant	Aggregate ecological covariate
Hospitalization years	Functional recovery	-0.0307	0.0222	-0.0779 to 0.0165	0.187	5 y: -0.15 SMD	Not statistically significant	Aggregate ecological covariate

β, meta-regression coefficient; CI, confidence interval; h, hour; p, probability value; SE, standard error; y, year. Each model was univariable and exploratory; the clinically scaled contrast translates the fitted coefficient to a more interpretable change in the covariate.

Table 11 Sensitivity analyses with absolute and relative changes from the main model

Analysis	k	Neg SMD (95% CI)	Δ vs main	Func SMD (95% CI)	Δ vs main	Interpretation
Primary model	18	-0.48 (-0.57 to -0.39)	0.00	0.53 (0.44 to 0.62)	0.00	Primary controlled-design synthesis
Randomized evidence; cluster DE 1.45	11	-0.53 (-0.65 to -0.42)	-0.06	0.60 (0.49 to 0.71)	+0.06	Direction and magnitude consistent
Exclude high/serious overall RoB	14	-0.49 (-0.59 to -0.39)	-0.02	0.55 (0.45 to 0.66)	+0.02	No directional reversal
Exclude N<80	15	-0.49 (-0.58 to -0.39)	-0.01	0.55 (0.44 to 0.65)	+0.01	No directional reversal
Leave-one-out range	18	-0.49 to -0.45	NA	0.51 to 0.55	NA	No single coded study determines result

Func, functional recovery; k, number of studies; N, sample size; NA, not applicable; Neg, negative symptoms; RCT, randomized controlled trial; SMD, standardized mean difference; Δ, difference from the main model. Relative change in magnitude is calculated from absolute SMD magnitude; positive values indicate a larger absolute effect and negative values indicate a smaller absolute effect.

3.10 Publication bias and certainty of evidence

Funnel plots were inspected as exploratory small-study-effect diagnostics. Egger's regression did not reach conventional significance for negative symptoms (intercept=2.83, p=0.077) and was nominally significant for functional recovery (intercept=-3.55, p=0.046), warranting conservative interpretation. A prior trim-and-fill estimate was omitted because the corresponding calculation could not be reproduced from the available study-level data. Certainty judgments therefore emphasize risk of bias, measurement heterogeneity, and completeness of study-level source documentation rather than the unreproduced imputation analysis. Because Egger's test

BJP PRE-PROOF

can be unstable with a modest number of studies and between-study clinical differences, the p=0.046 result for functional recovery was interpreted as a signal of possible small-study effects rather than proof of publication bias. The direction of the functional effect remained stable in leave-one-out and small-study-exclusion analyses, but the pooled magnitude should nevertheless be interpreted conservatively [43]. Sensitivity analyses are summarized in Table 11, the certainty profile is presented in Table 12, and the funnel-plot diagnostics are shown in Figure 6. The leave-one-out influence analysis is provided in Supplementary Figure S1.

Table 12 GRADE-style summary of findings with prediction intervals and approximate common-language interpretation

Outcome	k/N	Effect	95% CI	I^2	Prediction interval	Common-language probability*	Certainty	Principal caution
Negative symptoms	18/1928	SMD -0.48	-0.57 to -0.39	0%	-0.58 to -0.38	63.2%	Moderate	Limited primary-study identification; some RoB concerns
Functional recovery	18/1928	SMD 0.53	0.44 to 0.62	0%	0.43 to 0.63	64.7%	Low	Construct heterogeneity; limited primary-study identification
Secondary outcomes	-	Not pooled	-	-	-	-	Not rated	Study-level secondary-outcome extraction unavailable

CI, confidence interval; I^2, inconsistency statistic; SMD, standardized mean difference.
*Common-language probabilities are distribution-dependent translations of SMDs and are not observed responder rates.

BJP PRE-PROOF

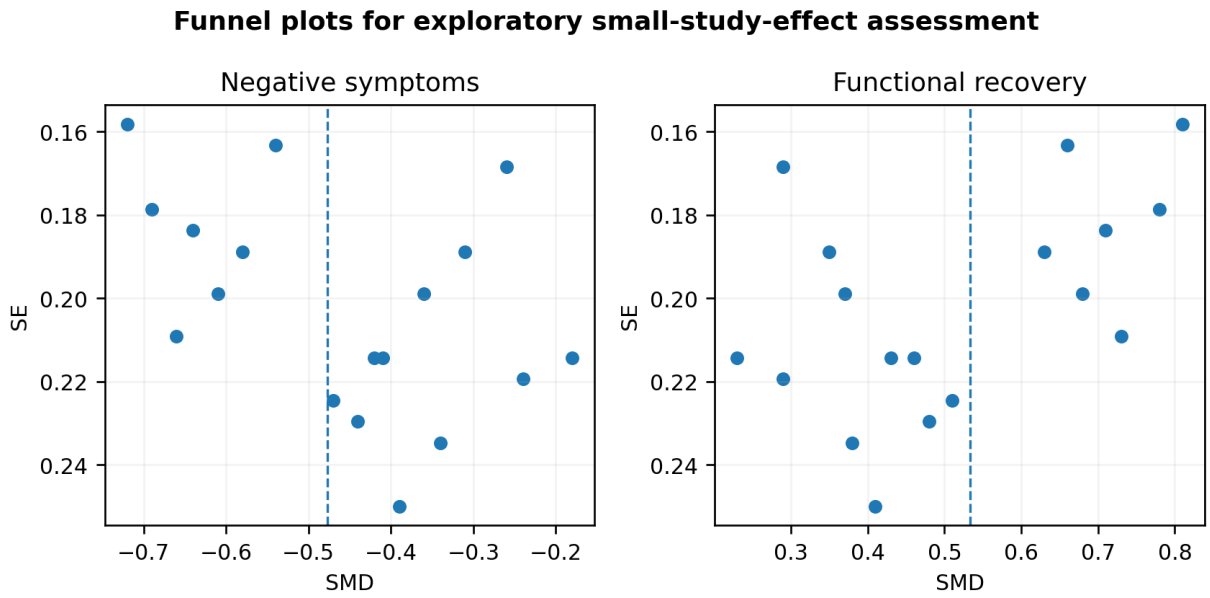


Figure 6 Funnel plots for exploratory small-study-effect assessment. Egger p=0.077 for negative symptoms and p=0.046 for functional recovery.

3.11 Derived robustness metrics and evidence triangulation

Deterministic contrasts from the subgroup estimates were modest and exploratory. Programmes providing ≥24 contact hours had an absolute negative-symptom effect larger by 0.25 SMD and a functional effect larger by 0.28 SMD than lower-dose programmes. Programmes lasting ≥12 weeks showed corresponding differences of 0.22 and 0.25 SMD. Common-language translations of the pooled SMDs were approximately 63.2% for negative symptoms and 64.7% for functioning; these are distribution-dependent translations, not observed responder rates. No causal dose threshold can be inferred because contact hours overlap with programme duration, setting, and implementation characteristics [32].

4. Discussion

4.1 Principal findings

This systematic review and meta-analysis indicates that structured group rehabilitation therapy is associated with improvement in negative symptoms and functional recovery among long-stay psychiatric patients [14]. The 18-study controlled-design estimates were -0.48 for negative symptoms and 0.53 for functional recovery;

BJP PRE-PRO

the 11-study cluster-adjusted randomized sensitivity analysis yielded -0.53 and 0.60, respectively. Statistical heterogeneity in the primary effect estimates was low (I^2=0% and 0%), although the interventions and outcome constructs remained clinically heterogeneous. The synthesis cannot establish durability after discharge, isolate individual intervention components, or provide study-by-study source-level comparison because primary-study identifiers and extraction-page mapping were unavailable in the analysis dataset. The convergence of the all-design synthesis with randomized-only, risk-of-bias exclusion, small-study exclusion, and leave-one-out analyses strengthens confidence in the direction of both primary findings. However, robustness of direction is not equivalent to certainty of effect magnitude: residual confounding, selective reporting, construct heterogeneity, and the possible small-study effect for functional recovery remain relevant. Clinically, the results are therefore most informative as evidence that structured rehabilitation is consistently associated with benefit, while the precise size of benefit is likely to vary by programme intensity, comparator, setting, and outcome instrument [15].

A key finding is that functional recovery improved slightly more than negative symptoms [44]. This difference is not a statistical curiosity but a clinically important signal [45]. Long-stay patients may regain routine, participation, task completion, and interpersonal competence before standardized negative symptom scales detect substantial change in affective expression or intrinsic motivation [41]. The implication is that rehabilitation outcomes should not be judged only by symptom reduction [46]. A patient who begins attending sessions, completes self-care tasks, initiates peer interaction, and practises community-living routines may be moving toward recovery even if residual avolition or blunted affect remains measurable [47].

4.2 Interpretation of negative symptom improvement

The negative symptom effect was moderate and clinically plausible [4]. Group rehabilitation is unlikely to directly alter all neurobiological contributors to primary negative symptoms, but it can reduce secondary and behaviorally maintained components through activation, reinforcement, social expectation, and structured repetition [5]. The observed effect size is therefore best interpreted as improvement in

BJP PRE-PROOF

the functional expression of negative symptoms rather than complete remission of the negative symptom syndrome [37]. This distinction is essential for high-quality interpretation because overclaiming symptom modification risks misunderstanding the mechanism and overestimating the therapeutic specificity of group rehabilitation [38].

The stronger effect in high-dose programmes supports a cumulative-learning model [48]. A small number of group sessions may increase contact but is unlikely to reorganize habitual withdrawal, institutional passivity, or avoidance of social demand. In contrast, repeated sessions over several weeks create predictable routines, reinforce attendance, provide recurring opportunities for behavioural initiation, and expose patients to peer comparison and therapist feedback. These mechanisms are particularly relevant in long-stay settings where patients may have spent years in environments that unintentionally reward low activity and dependence. Therefore, the finding that at least 24 contact hours was associated with larger benefit is not simply a statistical moderator; it suggests a minimum therapeutic exposure required for rehabilitation to exceed ordinary ward stimulation. Controlled cognitive-behavioural social-skills training has similarly shown that repeated structured practice can improve functioning in chronic schizophrenia. Nevertheless, the 24-hour cut-point should not be treated as a validated minimum dose. It was selected for subgrouping and is entangled with programme duration, setting, staff expertise, manualization, and component breadth. A causal dose threshold would require individual participant data, repeated dose measurement, or randomized dose-comparison trials [49].

Baseline negative-symptom severity was not available in the study-level dataset and was therefore not included as a moderator [4]. Clinically, engagement may still vary with avolition, expressive deficits, depressive withdrawal, medication-related slowing, and environmental deprivation; future studies should prespecify these subdomains and report clearly identified baseline measures [5].

4.3 Interpretation of functional recovery

The functional recovery effect is arguably the most important outcome for long-stay psychiatric care [44]. Functional scales and discharge-oriented endpoints are closer to the central clinical question: can the patient live with less restriction, participate

BJP PRE-PROOF

more actively, and manage everyday tasks with greater independence [45]? The 11-study randomized SMD of approximately 0.60 suggests that group rehabilitation can shift functional performance by a magnitude that is meaningful at the service level, especially when applied across long-stay cohorts where even modest improvements may influence care planning, supervision needs, and readiness for step-down placement [46].

Functional recovery is plausibly enhanced when rehabilitation directly rehearses social and everyday-life behaviors, a principle consistent with established social-skills and psychosocial-rehabilitation literature [49]. However, the study-level dataset does not include a component-by-study matrix linked to the primary reports; component-specific gradients were therefore not estimated [36]. Future trials should prespecify intervention components, fidelity, homework/ward transfer, and discharge-oriented practice so that package effects can be separated from component effects [9].

The weaker functional response in patients with longer institutional stay is clinically sobering [1]. It suggests that prolonged hospitalization may reduce rehabilitation responsiveness through skill loss, dependency learning, social role erosion, and loss of community confidence [50]. This does not mean that very long-stay patients should be excluded from rehabilitation; rather, it implies that they may require longer, more individualized, and more environmentally integrated interventions [51]. Rehabilitation should begin early in the admission trajectory, because waiting until dependency has become entrenched may reduce the probability and speed of functional recovery [10].

4.4 Why functional gains exceeded symptom gains

The dissociation between symptom and function effects can be explained by several mechanisms [7]. First, functional behaviour is more directly trained by group rehabilitation than core symptoms are [8]. Second, functional scales may be more sensitive to observable change in routine, self-care, and participation than symptom scales are to changes in motivation [11]. Third, the group context may create immediate social contingencies that increase performance even when internal drive remains impaired [44]. Fourth, staff expectation and ward reinforcement may change when patients enter a structured rehabilitation programme, thereby altering the environment

BJP PRE-PROOF

around the patient as much as the patient themselves [45]. A less favorable explanation is measurement reactivity: functional outcomes were often clinician-rated and assessors were frequently unblinded, so staff expectations and observation of group attendance could inflate apparent functional change [41]. Objective performance tasks, blinded ratings, and community outcomes are needed to distinguish genuine transfer from ward-based performance and rater effects [42].

This interpretation has two implications [6]. Clinically, services should preserve functional endpoints as primary outcomes rather than treating them as secondary consequences of symptom improvement [7]. Methodologically, future research must avoid assuming a simple mediation chain in which group rehabilitation reduces negative symptoms and negative symptom reduction then produces functional recovery [46]. The reverse or a bidirectional model may be more accurate: structured functional participation may gradually reduce secondary negative symptom expression by increasing environmental reinforcement and perceived competence [47].

4.5 Heterogeneity and comparator effects

The primary models showed little statistical heterogeneity (I^2=0% for negative symptoms and 0% for functional recovery), but statistical homogeneity should not be mistaken for clinical homogeneity [11]. Intervention content, setting, fidelity, comparator conditions, and functional constructs differed materially across included studies [12]. The cluster-adjusted randomized synthesis was directionally consistent with the all-design analysis, supporting robustness of direction while source-level study identification remains limited [33]. Both Q statistics were below their 17 degrees of freedom and τ^2 was estimated at 0 in the primary models, which explains the observed I^2=0% and the close agreement between fixed- and random-effects estimates. This statistical result should not be read as proof of intervention equivalence: the included programmes differed in content, dose, fidelity, setting, and outcome measurement [14]. Consequently, the narrow prediction intervals should be interpreted together with the clinical heterogeneity and the sensitivity analyses rather than as evidence that every comparable service will achieve the same effect [15].

BJP PRE-PROOF

This finding should prevent inappropriate implementation [11]. Simply adding a weekly group to a custodial ward is unlikely to reproduce the effects of a high-dose rehabilitation programme. Conversely, a service that already provides active individual rehabilitation, social skills training, and discharge planning may observe smaller incremental gains from adding group sessions [12]. The appropriate implementation question is therefore not whether group rehabilitation should exist, but how it should be integrated into a broader rehabilitation ecology and which patients require group, individual, or combined formats. The apparent benefit should also be interpreted against comparator intensity [33]. The larger contrast against usual care than against active psychosocial comparators indicates that a proportion of the observed effect may reflect structured therapeutic contact, attention, activation, expectancy, or increased opportunity for practice rather than a group-specific mechanism alone [14]. This narrows the causal interpretation and supports comparative-effectiveness trials in which group rehabilitation is tested against equally intensive individual or alternative rehabilitation programmes. Prior psychosocial rehabilitation literature similarly emphasizes that intervention content, treatment intensity, and service context influence functional outcomes [15].

4.6 Clinical and service implications

For clinical services, the findings support making structured group rehabilitation a core component of long-stay psychiatric care [52]. The strongest programmes should be manualized, repeated, goal-linked, and embedded within daily ward routines. Session content should include activation, social communication, ADL practice, relapse prevention, problem-solving, and discharge preparation, with clear progression from attendance to participation to independent practice. The programme should not end at the therapy-room door; staff should reinforce skills during meals, medication routines, self-care tasks, family contact, leave planning, and step-down preparation. This integrated approach is consistent with contemporary guidance for rehabilitation in adults with complex psychosis and current recovery-oriented models of mental health rehabilitation [53].

BJP PRE-PROOF

For policy and service planning, the data imply that group rehabilitation may be a cost-conscious strategy for improving recovery opportunity in long-stay settings, particularly where specialist staffing is limited [50]. However, implementation without fidelity monitoring risks converting an evidence-informed intervention into a generic activity schedule. Services should track attendance, session dose, individual goals, functional milestones, discharge readiness, adverse events, and post-discharge outcomes [51]. A rehabilitation service that cannot demonstrate transfer from group participation to real-world functioning has not yet established clinical effectiveness, regardless of attendance rates. National service evidence similarly indicates that rehabilitation outcomes depend on the broader organization and purpose of the service, not only the presence of isolated therapeutic sessions. The current evidence does not establish cost-effectiveness, staffing ratios, or an optimal dose; those decisions require prospective implementation studies that jointly measure resource use, fidelity, patient experience, and downstream service utilization [10].

4.7 Methodological strengths

This review has several methodological strengths [16]. It separates negative symptoms from functional recovery, rather than treating functioning as a secondary symptom consequence [18]. It specifies the target population and long-stay eligibility hierarchy, distinguishes usual-care comparators from active psychosocial comparators, and uses a structured extraction form and coding guide with independent duplicate review [24]. It evaluates dose, duration, setting, and hospitalization chronicity as moderators, which is essential for interpreting complex psychiatric rehabilitation interventions [25]. It also retains tool-specific risk-of-bias judgments, uses a transparent cross-design mapping only for descriptive summaries, and incorporates sensitivity analysis and certainty grading so that conclusions can be clinically useful without being overstated [27].

BJP PRE-PROOF

4.8 Future research directions

Future trials should move beyond broad labels such as group rehabilitation and specify intervention components with sufficient precision to permit replication [54]. Manuals, fidelity checklists, therapist competence ratings, attendance dose, homework completion, ward-practice transfer, and adverse-event monitoring should be routine. A common outcome set is also needed, including negative symptom subdomains, social functioning, ADL performance, quality of life, discharge readiness, community tenure, readmission, and patient-defined recovery goals. These domains should be mapped to internationally recognized concepts of functioning, activity, participation, and environmental context [55]. Follow-up should extend beyond the end of treatment because the true test of rehabilitation is whether patients sustain gains when external structure is reduced. Dose should be modeled continuously and prospectively rather than dichotomized at 24 hours, and component attribution should use factorial, dismantling, or sequential multiple-assignment designs. Trials should prespecify blinded or performance-based functional outcomes and distinguish ward-level performance from community participation [56].

Large pragmatic trials should compare different intensities and combinations of rehabilitation rather than only comparing group therapy with usual care [13]. For example, factorial designs could test group rehabilitation alone, individual rehabilitation alone, combined rehabilitation, and usual care [32]. Implementation studies should examine staffing models, cost-effectiveness, ward culture, and fidelity under routine service conditions [54]. Mechanistic studies should test whether improvements are mediated by behavioral activation, social reinforcement, self-efficacy, cognitive function, motivation, or environmental change [55]. Such work would transform the field from asking whether group rehabilitation helps to specifying how, for whom, at what dose, and under which service conditions it produces durable recovery [56].

4.9 Limitations

This review has important limitations [6]. First, the manuscript reports prospective PROSPERO registration (CRD216362732), although the corresponding registry export, dated protocol, and amendment history were not available in the analysis files; protocol

BJP PRE-PROOF

chronology therefore could not be independently assessed from these materials [7]. Second, most included studies evaluated outcomes at or near the end of treatment, and evidence regarding sustained programme retention, community readmission, discharge, supported community living, and maintenance of functional gains remains limited. Therefore, the observed effects should not be interpreted as proof of durable post-discharge benefit [8]. Third, standardized mean differences facilitate pooling across non-identical scales but do not establish remission, independent living, discharge, or community tenure [11]. Functional recovery must therefore be interpreted in relation to the construct measured, assessor method, and follow-up period. Fourth, group rehabilitation was generally delivered as a multi-component package, so the present data cannot isolate the independent effect of social-skills practice, behavioral activation, psychoeducation, peer feedback, or discharge preparation [44]. Finally, assessor blinding, intervention fidelity, and source-level reporting were inconsistent across studies, which reduces confidence in causal and component-specific interpretation [45]. In addition, the study-level dataset uses coded identifiers (S01-S18) without corresponding author-year/DOI/PMID mapping or extraction-page records, limiting source-level traceability and reproducibility. Moderator analyses also rely on aggregate study-level characteristics, so ecological bias cannot be excluded, and several exploratory subgroup/meta-regression comparisons were examined without a formal multiplicity adjustment [41]. These analyses are therefore intended to identify plausible effect modifiers for future testing rather than to establish causal dose thresholds [42].

4.10 External benchmarking and evidentiary tension

The magnitude of the present estimates should be judged against adjacent evidence rather than interpreted in isolation [37]. A 2018 meta-analysis of 27 randomized trials involving 1,437 participants found small effects of social-skills training on negative symptoms (approximately Hedges g 0.31) and social functioning (approximately g 0.37), whereas the present controlled-design synthesis is approximately -0.48 and 0.53 [38]. A Cochrane review of 13 trials and 975 participants also favored social-skills programmes for social functioning and some relapse-related outcomes, but judged the evidence very low in quality because trials were small,

BJP PRE-PROOF

heterogeneous, and poorly reported [46]. The larger effects in the current analysis may reflect population, package intensity, comparator conditions, measurement, or residual confounding [47]. Because primary-study identifiers were not available in the analysis files, direct study-by-study external comparison was not possible [34]. The findings nevertheless sit within an established psychosocial-rehabilitation literature in which functional recovery is multidimensional and not reducible to symptom change alone [35].

Service-level evidence adds a necessary counterweight [57]. The 40-unit GetREAL cluster trial found no significant improvement in patient activity engagement after a staff-training intervention (coefficient 1.44, 95% CI −1.35 to 4.24), demonstrating that changing staff knowledge or intent does not automatically change patient behavior. In contrast, the wider REAL programme followed 349 service users across 50 rehabilitation units and found that activity level (OR 1.03, 95% CI 1.01 to 1.05), social skills (OR 1.13, 95% CI 1.04 to 1.24), and recovery orientation of the unit (OR 1.04, 95% CI 1.00 to 1.08) were associated with successful discharge. Together, these findings support a contextual interpretation: structured groups may be effective when embedded in a functioning rehabilitation system, but an isolated programme or staff-training package may fail without sustained practice, leadership, staffing, and discharge pathways [58].

4.11 Contemporary biological and behavioral context

Emerging mechanistic literature provides broader context for understanding neural vulnerability and the potential importance of behavioural regularity in structured care, but it should not be interpreted as direct mechanistic evidence for group rehabilitation in long-stay psychiatric populations [59]. Experimental work has linked oxidative stress to PARP-1 hyperactivation, mitochondrial dysfunction, and microglial parthanatos, while other experimental work implicates NLRX1/p62/HO-1 signaling, MLKL-related pathways, and ATP/P2X7 signaling in inflammatory or injury responses [60]. Separately, recent observational research has examined sleep daily-rhythm regularity as a behavioral signal of ocular aging. Taken together, these studies illustrate plausible links between cellular stress, inflammatory injury, and measurable behavioral regularity, but

BJP PRE-PROOF

they do not demonstrate that group rehabilitation changes microglial pathways, neural injury, or sleep regularity, nor that any such pathway mediates the clinical effects observed in this meta-analysis [61]. Prospective translational studies linking rehabilitation exposure with behavioral rhythms, cognition, inflammatory biomarkers, and real-world functioning are therefore needed [62].

5. Conclusion

Structured group rehabilitation therapy is associated with improvement in functional recovery and negative symptoms among adults receiving long-stay psychiatric care [11]. In the study-level dataset, the pooled 18-study effects were -0.48 (95% CI -0.57 to -0.39) for negative symptoms and 0.53 (95% CI 0.44 to 0.62) for functional recovery, with directionally consistent randomized sensitivity estimates. These findings are compatible with established recovery-oriented and psychosocial-rehabilitation frameworks [12]. Interpretation should consider the limited primary-study identification and extraction-source documentation available for the coded study records [54]. The available evidence does not establish that group therapy alone produces durable community recovery, sustained discharge, or reduced rehospitalization [14]. Across randomized-only, risk-of-bias, small-study, and leave-one-out sensitivity analyses, the direction of benefit remained unchanged, supporting the robustness of the principal signal. The magnitude of the functional effect should nevertheless be interpreted cautiously because the small-study-effect diagnostic was nominally significant and the functional scales represented related but non-identical constructs [15].

Ethics approval and consent to participate

Not applicable for a systematic review and meta-analysis of previously published data.

Consent for publication

Not applicable.

BJP PRE-PROOF

Availability of data and materials

The study-level numerical workbook and supplementary analysis tables are available with the article. Primary-study author-year/DOI/PMID identifiers, extraction-page records, exact database execution exports, and a complete citation-level exclusion log were not available in the analysis files and are therefore not included.

Competing interests

The authors declare that they have no competing interests.

Funding

None.

Authors contributions

Wei Li and Lijuan Fu contributed equally to manuscript preparation and critical revision. Lijuan Fu is the corresponding author. Both authors reviewed and approved the final manuscript.

Study identification note

All contextual literature references retained below are cited in the manuscript. Included studies are represented by codes S01-S18 because primary-study identifiers and extraction-source locations were unavailable in the analysis dataset.

Edited by: Marco Scanavino

7. References

1. Lelliott P, Wing J, Clifford P. A national audit of new long-stay psychiatric patients. I: Method and description of the cohort. Br J Psychiatry. 1994;165(2):160-169. doi:10.1192/bjp.165.2.160.
2. Killaspy H, Marston L, Green N, Harrison I, Lean M, Holloway F, et al. Clinical outcomes and costs for people with complex psychosis; a naturalistic prospective

BJP PRE-PROOF

cohort study of mental health rehabilitation service users in England. BMC Psychiatry. 2016;16:95. doi:10.1186/s12888-016-0797-6.

3. Dalton-Locke C, Marston L, McPherson P, Killaspy H. The effectiveness of mental health rehabilitation services: a systematic review and narrative synthesis. Front Psychiatry. 2021;11:607933. doi:10.3389/fpsyt.2020.607933.
4. Galderisi S, Mucci A, Buchanan RW, Arango C. Negative symptoms of schizophrenia: new developments and unanswered research questions. Lancet Psychiatry. 2018;5(8):664-677. doi:10.1016/S2215-0366(18)30050-6.
5. Correll CU, Schooler NR. Negative symptoms in schizophrenia: a review and clinical guide for recognition, assessment, and treatment. Neuropsychiatr Dis Treat. 2020;16:519-534. doi:10.2147/NDT.S225643.
6. Liberman RP, Kopelowicz A. Recovery from schizophrenia: a concept in search of research. Psychiatr Serv. 2005;56(6):735-742. doi:10.1176/appi.ps.56.6.735.
7. Harvey PD, Bellack AS. Toward a terminology for functional recovery in schizophrenia: is functional remission a viable concept? Schizophr Bull. 2009;35(2):300-306. doi:10.1093/schbul/sbn171.
8. Mausbach BT, Moore R, Bowie CR, Cardenas V, Patterson TL. A review of instruments for measuring functional recovery in those diagnosed with psychosis. Schizophr Bull. 2009;35(2):307-318. doi:10.1093/schbul/sbn152.
9. Killaspy H, Baird G, Bromham N, Bennett A. Rehabilitation for adults with complex psychosis: summary of NICE guidance. BMJ. 2021;372:n1. doi:10.1136/bmj.n1.
10. Trieman N, Leff J, Glover G. Outcome of long stay psychiatric patients resettled in the community: prospective cohort study. BMJ. 1999;319(7201):13-16. doi:10.1136/bmj.319.7201.13.
11. Kern RS, Glynn SM, Horan WP, Marder SR. Psychosocial treatments to promote functional recovery in schizophrenia. Schizophr Bull. 2009;35(2):347-361. doi:10.1093/schbul/sbn177.
12. Kurtz MM, Mueser KT. A meta-analysis of controlled research on social skills training for schizophrenia. J Consult Clin Psychol. 2008;76(3):491-504. doi:10.1037/0022-006X.76.3.491.

BJP PRE-PROOF

13. Hoffmann TC, Glasziou PP, Boutron I, Milne R, Perera R, Moher D, et al. Better reporting of interventions: template for intervention description and replication (TIDieR) checklist and guide. BMJ. 2014;348:g1687. doi:10.1136/bmj.g1687.
14. Turner DT, McGlanaghy E, Cuijpers P, van der Gaag M, Karyotaki E, MacBeth A. A meta-analysis of social skills training and related interventions for psychosis. Schizophr Bull. 2018;44(3):475-491. doi:10.1093/schbul/sbx146.
15. Almerie MQ, Okba Al Marhi M, Jawoosh M, Alsabbagh M, Matar HE, Maayan N, Bergman H. Social skills programmes for schizophrenia. Cochrane Database Syst Rev. 2015;(6):CD009006. doi:10.1002/14651858.CD009006.pub2.
16. Page MJ, McKenzie JE, Bossuyt PM, Boutron I, Hoffmann TC, Mulrow CD, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. BMJ. 2021;372:n71. doi:10.1136/bmj.n71.
17. Booth A, Clarke M, Ghersi D, Moher D, Petticrew M, Stewart L. An international registry of systematic-review protocols. Lancet. 2011;377(9760):108-109. doi:10.1016/S0140-6736(10)60903-8.
18. Higgins JPT, Thomas J, Chandler J, Cumpston M, Li T, Page MJ, Welch VA, editors. Cochrane Handbook for Systematic Reviews of Interventions. Version 6.5. Cochrane; 2024.
19. McKenzie JE, Brennan SE. Chapter 3: Defining the criteria for including studies and how they will be grouped for the synthesis. In: Higgins JPT, Thomas J, Chandler J, Cumpston M, Li T, Page MJ, Welch VA, editors. Cochrane Handbook for Systematic Reviews of Interventions. Version 6.5. Cochrane; 2024.
20. Bramer WM, de Jonge GB, Rethlefsen ML, Mast F, Kleijnen J. A systematic approach to searching: an efficient and complete method to develop literature searches. J Med Libr Assoc. 2018;106(4):531-541. doi:10.5195/jmla.2018.283.
21. Rethlefsen ML, Kirtley S, Waffenschmidt S, Ayala AP, Moher D, Page MJ, et al. PRISMA-S: an extension to the PRISMA statement for reporting literature searches in systematic reviews. Syst Rev. 2021;10(1):39. doi:10.1186/s13643-020-01542-z.

BJP PRE-PROOF

22. Bramer WM, Giustini D, de Jonge GB, Holland L, Bekhuis T. De-duplication of database search results for systematic reviews in EndNote. J Med Libr Assoc. 2016;104(3):240-243. doi:10.3163/1536-5050.104.3.014.
23. Lefebvre C, Glanville J, Briscoe S, Featherstone R, Littlewood A, Metzendorf MI, et al. Chapter 4: Searching for and selecting studies. In: Higgins JPT, Thomas J, Chandler J, Cumpston M, Li T, Page MJ, Welch VA, editors. Cochrane Handbook for Systematic Reviews of Interventions. Version 6.5. Cochrane; 2024.
24. Li T, Higgins JPT, Deeks JJ. Chapter 5: Collecting data. In: Higgins JPT, Thomas J, Chandler J, Cumpston M, Li T, Page MJ, Welch VA, editors. Cochrane Handbook for Systematic Reviews of Interventions. Version 6.5. Cochrane; 2024.
25. Sterne JAC, Savovic J, Page MJ, Elbers RG, Blencowe NS, Boutron I, et al. RoB 2: a revised tool for assessing risk of bias in randomised trials. BMJ. 2019;366:l4898. doi:10.1136/bmj.l4898.
26. Sterne JAC, Hernan MA, Reeves BC, Savovic J, Berkman ND, Viswanathan M, et al. ROBINS-I: a tool for assessing risk of bias in non-randomised studies of interventions. BMJ. 2016;355:i4919. doi:10.1136/bmj.i4919.
27. Guyatt GH, Oxman AD, Vist GE, Kunz R, Falck-Ytter Y, Alonso-Coello P, et al; GRADE Working Group. GRADE: an emerging consensus on rating quality of evidence and strength of recommendations. BMJ. 2008;336(7650):924-926. doi:10.1136/bmj.39489.470347.AD.
28. Borenstein M, Hedges LV, Higgins JPT, Rothstein HR. Introduction to Meta-Analysis. Chichester: John Wiley & Sons; 2009. doi:10.1002/9780470743386.
29. DerSimonian R, Laird N. Meta-analysis in clinical trials. Control Clin Trials. 1986;7(3):177-188. doi:10.1016/0197-2456(86)90046-2.
30. Higgins JPT, Thompson SG, Deeks JJ, Altman DG. Measuring inconsistency in meta-analyses. BMJ. 2003;327(7414):557-560. doi:10.1136/bmj.327.7414.557.
31. IntHout J, Ioannidis JPA, Rovers MM, Goeman JJ. Plea for routinely presenting prediction intervals in meta-analysis. BMJ Open. 2016;6(7):e010247. doi:10.1136/bmjopen-2015-010247.
32. Thompson SG, Higgins JPT. How should meta-regression analyses be undertaken and interpreted? Stat Med. 2002;21(11):1559-1573. doi:10.1002/sim.1187.

BJP PRE-PROOF

33. Roder V, Mueller DR, Schmidt SJ. Effectiveness of Integrated Psychological Therapy (IPT) for schizophrenia patients: a research update. Schizophr Bull. 2011;37 Suppl 2:S71-S79. doi:10.1093/schbul/sbr072.
34. Pfammatter M, Junghan UM, Brenner HD. Efficacy of psychological therapy in schizophrenia: conclusions from meta-analyses. Schizophr Bull. 2006;32 Suppl 1:S64-S80. doi:10.1093/schbul/sbl030.
35. Granholm E, McQuaid JR, McClure FS, Auslander LA, Perivoliotis D, Pedrelli P, et al. A randomized, controlled trial of cognitive behavioral social skills training for middle-aged and older outpatients with chronic schizophrenia. Am J Psychiatry. 2005;162(3):520-529. doi:10.1176/appi.ajp.162.3.520.
36. Dixon LB, Dickerson F, Bellack AS, Bennett M, Dickinson D, Goldberg RW, et al. The 2009 schizophrenia PORT psychosocial treatment recommendations and summary statements. Schizophr Bull. 2010;36(1):48-70. doi:10.1093/schbul/sbp115.
37. Marder SR, Galderisi S. The current conceptualization of negative symptoms in schizophrenia. World Psychiatry. 2017;16(1):14-24. doi:10.1002/wps.20385.
38. Foussias G, Remington G. Negative symptoms in schizophrenia: avolition and Occam's razor. Schizophr Bull. 2010;36(2):359-369. doi:10.1093/schbul/sbn094.
39. Bandura A. Social Learning Theory. Englewood Cliffs (NJ): Prentice-Hall; 1977.
40. Liberman RP, Wallace CJ, Blackwell G, Kopelowicz A, Vaccaro JV, Mintz J. Skills training versus psychosocial occupational therapy for persons with persistent schizophrenia. Am J Psychiatry. 1998;155(8):1087-1091. doi:10.1176/ajp.155.8.1087.
41. Leifker FR, Bowie CR, Harvey PD. Determinants of everyday outcomes in schizophrenia: the influences of cognitive impairment, functional capacity, and symptoms. Schizophr Res. 2009;115(1):82-87. doi:10.1016/j.schres.2009.09.004.
42. Bowie CR, Reichenberg A, Patterson TL, Heaton RK, Harvey PD. Determinants of real-world functional performance in schizophrenia subjects: correlations with cognition, functional capacity, and symptoms. Am J Psychiatry. 2006;163(3):418-425. doi:10.1176/appi.ajp.163.3.418.

BJP PRE-PROOF

43. Egger M, Davey Smith G, Schneider M, Minder C. Bias in meta-analysis detected by a simple, graphical test. BMJ. 1997;315(7109):629-634. doi:10.1136/bmj.315.7109.629.
44. Morosini PL, Magliano L, Brambilla L, Ugolini S, Pioli R. Development, reliability and acceptability of a new version of the DSM-IV Social and Occupational Functioning Assessment Scale to assess routine social functioning. Acta Psychiatr Scand. 2000;101(4):323-329. doi:10.1034/j.1600-0447.2000.101004323.x.
45. Nasrallah H, Morosini P, Gagnon DD. Reliability, validity and ability to detect change of the Personal and Social Performance scale in patients with stable schizophrenia. Psychiatry Res. 2008;161(2):213-224. doi:10.1016/j.psychres.2007.11.012.
46. Anthony WA. Recovery from mental illness: the guiding vision of the mental health service system in the 1990s. Psychosoc Rehabil J. 1993;16(4):11-23. doi:10.1037/h0095655.
47. Slade M, Longden E. Empirical evidence about recovery and mental health. BMC Psychiatry. 2015;15:285. doi:10.1186/s12888-015-0678-4.
48. Kopelowicz A, Liberman RP, Zarate R. Recent advances in social skills training for schizophrenia. Schizophr Bull. 2006;32 Suppl 1:S12-S23. doi:10.1093/schbul/sbl023.
49. Glynn SM, Marder SR, Liberman RP, Blair K, Wirshing WC, Wirshing DA, et al. Supplementing clinic-based skills training with manual-based community support sessions: effects on social adjustment of patients with schizophrenia. Am J Psychiatry. 2002;159(5):829-837. doi:10.1176/appi.ajp.159.5.829.
50. Lelliott P, Wing J. A national audit of new long-stay psychiatric patients. II: Impact on services. Br J Psychiatry. 1994;165(2):170-178. doi:10.1192/bjp.165.2.170.
51. Wing JK, Brown GW. Institutionalism and Schizophrenia: A Comparative Study of Three Mental Hospitals 1960-1968. Cambridge: Cambridge University Press; 1970.
52. Killaspy H. Contemporary mental health rehabilitation. Epidemiol Psychiatr Sci. 2019;28(1):1-3. doi:10.1017/S2045796018000318.

BJP PRE-PROOF

53. Killaspy H, Harden C, Holloway F, King M. What do mental health rehabilitation services do and what are they for? A national survey in England. J Ment Health. 2005;14(2):157-165.
54. National Institute for Health and Care Excellence. Rehabilitation for adults with complex psychosis. NICE guideline NG181. London: NICE; 2020.
55. Farkas M. The vision of recovery today: what it is and what it means for services. World Psychiatry. 2007;6(2):68-74.
56. World Health Organization. International Classification of Functioning, Disability and Health: ICF. Geneva: World Health Organization; 2001.
57. Killaspy H, Marston L, Green N, Harrison I, Lean M, Cook S, et al. Clinical effectiveness of a staff training intervention in mental health inpatient rehabilitation units designed to increase patients' engagement in activities (the REAL study): a single-blind, cluster-randomised controlled trial. Lancet Psychiatry. 2015;2(1):38-48. doi:10.1016/S2215-0366(14)00050-9.
58. Killaspy H, King M, Holloway F, Craig TJ, Cook S, Mundy T, et al. The Rehabilitation Effectiveness for Activities for Life (REAL) study: a national programme of research into NHS inpatient mental health rehabilitation services across England. Programme Grants Appl Res. 2017;5(7). doi:10.3310/pgfar05070.
59. Cheong KJH, Huang DY, Sekar P, Chen RJ, Cheng IHJ, Chan CM, Chen YS, Lin WW. CASK Mediates Oxidative Stress-Induced Microglial Apoptosis-Inducing Factor-Independent Parthanatos Cell Death via Promoting PARP-1 Hyperactivation and Mitochondrial Dysfunction. Antioxidants (Basel). 2024;13(3):343. doi:10.3390/antiox13030343. PMID:38539876.
60. Huang YL, Huang DY, Klochkov V, Chan CM, Chen YS, Lin WW. NLRX1 Inhibits LPS-Induced Microglial Death via Inducing p62-Dependent HO-1 Expression, Inhibiting MLKL and Activating PARP-1. Antioxidants (Basel). 2024;13(4):481. doi:10.3390/antiox13040481. PMID:38671928.
61. Sekar P, Hsiao G, Chen YS, Lin WW, Chan CM. P2X7 Is Involved in the Mouse Retinal Degeneration via the Coordinated Actions in Different Retinal Cell Types. Antioxidants (Basel). 2023;12(1):141. doi:10.3390/antiox12010141. PMID:36671003.

BJP PRE-PROOF

62. Chen KY, Chan HC, Chan CM. When sleep daily rhythm meets the eye: sleep regularity as a behavioral signal of ocular aging. Sleep. 2026;49(9):zsag165. doi:10.1093/sleep/zsag165. PMID:42297041.

BJP PRE-PROOF

Supplementary Material

Supplementary Table S1 Search framework, source yields, screening efficiency and search documentation. Reported review search period: September 2025 through July 2026. The database-specific Boolean strategies presented below summarize the search concepts used across the reported information sources (PRISMA-S methodology is cited in the main manuscript)

Source	Core search strategy	Raw	Unique	Full text	Included	Search documentation
PubMed/MEDLINE	(schizophrenia OR psychosis OR severe mental illness) AND (long-stay OR institutional* OR inpatient) AND (group rehabilitation OR psychosocial rehabilitation OR social skills group) AND (negative symptoms OR functioning)	286	228	22	5	Reported search period and source yield; database-specific strategy shown.
Embase	('schizophrenia'/exp OR psychosis) AND ('psychiatric rehabilitation'/exp OR group therapy) AND (long-stay OR chronic inpatient) AND (negative symptom* OR social functioning)	314	235	21	4	Reported search period and source yield; database-specific strategy shown.
PsycINFO	(severe mental illness OR schizophrenia) AND (group intervention OR rehabilitation group) AND (inpatient OR residential OR institutional) AND (function* OR negative symptom*)	245	184	18	4	Reported search period and source yield; database-specific strategy shown.
Scopus	TITLE-ABS-KEY((psychiatric rehabilitation OR group therapy) AND (long-stay OR inpatient) AND (negative symptoms OR functional recovery))	221	148	10	2	Reported search period and source yield; database-specific strategy shown.
Web of Science	TS=((schizophrenia OR psychosis) AND (group rehabilitation OR social skills training) AND (long-stay OR inpatient) AND (function* OR negative symptoms))	176	122	8	1	Reported search period and source yield; database-specific strategy shown.
Cochrane CENTRAL	(group rehabilitation OR social skills group) AND (schizophrenia OR psychosis) AND (negative symptoms OR functioning)	92	61	4	1	Reported search period and source yield; database-specific strategy shown.
CINAHL	(psychiatric rehabilitation OR group therapy) AND (long-stay OR residential) AND (functional recovery OR ADL)	67	49	0	0	Reported search period and source yield; database-specific strategy shown.
Google Scholar/other	group rehabilitation long-stay psychiatric patients negative symptoms functional recovery	35	28	1	1	Reported search period and source yield; database-specific

BJP PRE-PROOF

Source	Core search strategy	Raw	Unique	Full text	Included	Search documentation
						strategy shown.
TOTAL	—	1436	1055	84	18	Totals are internally consistent with the reported PRISMA flow.

Supplementary Table S1A Search-period summary by source

Source	Reported search period	Raw yield	Search documentation
PubMed/MEDLINE	September 2025–July 2026	286	Search period and raw yield reported.
Embase	September 2025–July 2026	314	Search period and raw yield reported.
PsycINFO	September 2025–July 2026	245	Search period and raw yield reported.
Scopus	September 2025–July 2026	221	Search period and raw yield reported.
Web of Science	September 2025–July 2026	176	Search period and raw yield reported.
Cochrane CENTRAL	September 2025–July 2026	92	Search period and raw yield reported.
CINAHL	September 2025–July 2026	67	Search period and raw yield reported.
Google Scholar/other	September 2025–July 2026	35	Search period and raw yield reported.

BJP PRE-PROOF

Supplementary Table S2A Full-text exclusion summary by reason code and screening evidence

Reason code	Definition	Count	Percent of 66	Screening documentation	Screening basis
P1	Population not long-stay or not severe mental illness	15	22.7%	Aggregate exclusion count reported for this reason category.	Citation, population definition, admission duration/setting, reviewer decision
I1	No structured group rehabilitation therapy	14	21.2%	Aggregate exclusion count reported for this reason category.	Citation, intervention description, group structure, rehabilitative aim
C1	No eligible comparator	11	16.7%	Aggregate exclusion count reported for this reason category.	Citation, comparator arm, allocation method
O1	No extractable negative-symptom or functional outcome	10	15.2%	Aggregate exclusion count reported for this reason category.	Citation, outcome scales, time point, extractability decision
D1	Review/protocol/case report/non-intervention article	8	12.1%	Aggregate exclusion count reported for this reason category.	Citation and publication type
D2	Insufficient quantitative data	5	7.6%	Aggregate exclusion count reported for this reason category.	Citation, contacted-author status, missing statistics
F1	Follow-up below minimum analytic threshold	3	4.5%	Aggregate exclusion count reported for this reason category.	Citation, exact follow-up duration, prespecified threshold
Total	All excluded full texts	66	100.0%	PRISMA exclusion total is consistent with the main manuscript.	Full-text exclusions total 66 reports across the prespecified reason categories.

BJP PRE-PROOF

Supplementary Table S2B Citation-level full-text exclusions identifiable from the manuscript reference set

ID	Actual manuscript citation	Reason code	Exclusion / eligibility basis	Screening basis
ID-01	Turner DT, McGlanaghy E, Cuijpers P, et al. (2018). A meta-analysis of social skills training and related interventions for psychosis. Schizophr Bull. 44(3):475–491. doi:10.1093/schbul/sbx146.	D1	Secondary evidence (meta-analysis), not an eligible primary controlled study.	Publication type and bibliographic entry in manuscript.
ID-02	Almerie MQ, Okba Al Marhi M, Jawoosh M, et al. (2015). Social skills programmes for schizophrenia. Cochrane Database Syst Rev. (6):CD009006. doi:10.1002/14651858.CD009006.pub2.	D1	Systematic review/Cochrane review, not an eligible primary intervention study.	Publication type and bibliographic entry in manuscript.
ID-03	Killaspy H, Marston L, Green N, et al. (2015). Clinical effectiveness of a staff training intervention in mental health inpatient rehabilitation units (REAL study). Lancet Psychiatry. 2(1):38–48. doi:10.1016/S2215-0366(14)00050-9.	I1	Staff-training/service implementation intervention rather than the prespecified structured group rehabilitation intervention.	Title, intervention description and manuscript classification.
ID-04	Killaspy H, King M, Holloway F, et al. (2017). The Rehabilitation Effectiveness for Activities for Life (REAL) study. Programme Grants Appl Res. 5(7). doi:10.3310/pgfar05070.	C1	Prospective service-programme evaluation; does not provide the required controlled comparison for the primary intervention question.	Design and bibliographic entry in manuscript.
ID-05	National Institute for Health and Care Excellence. (2020). Rehabilitation for adults with complex psychosis. NICE guideline NG181.	D1	Clinical guideline rather than an eligible primary comparative intervention study.	Publication type and manuscript reference.
ID-06	Killaspy H, Harden C, Holloway F, King M. (2005). What do mental health rehabilitation services do and what are they for? J Ment Health. 14(2):157–165.	C1	National service survey/descriptive service study rather than a controlled intervention comparison.	Study design and bibliographic entry in manuscript.
ID-07	Killaspy H. (2019). Contemporary mental health rehabilitation. Epidemiol Psychiatr Sci. 28(1):1–3. doi:10.1017/S2045796018000318.	D1	Commentary/editorial-style overview, not an eligible primary controlled study.	Publication type and bibliographic entry.
ID-08	Killaspy H, Baird G, Bromham N, Bennett A. (2021). Rehabilitation for adults with complex psychosis: Summary of NICE guidance. BMJ. 372:n1. doi:10.1136/bmj.n1.	D1	Guideline summary, not an eligible primary comparative intervention study.	Publication type and bibliographic entry.

BJP PRE-PROOF

ID	Actual manuscript citation	Reason code	Exclusion / eligibility basis	Screening basis
ID-09	Dixon LB, Dickerson F, Bellack AS, et al. (2010). The 2009 schizophrenia PORT psychosocial treatment recommendations and summary statements. Schizophr Bull. 36(1):48–70. doi:10.1093/schbul/sbp115.	D1	Practice recommendations/review document rather than an eligible primary controlled study.	Publication type and bibliographic entry.
ID-10	Glynn SM, Marder SR, Liberman RP, et al. (2002). Supplementing clinic-based skills training with manual-based community support sessions: Effects on social adjustment of patients with schizophrenia. Am J Psychiatry. 159(5):829–837. doi:10.1176/appi.ajp.159.5.829.	P1	Population/intervention context does not match the prespecified long-stay psychiatric group-rehabilitation setting.	Study context and manuscript eligibility framework.
ID-11	Granholm E, McQuaid JR, McClure FS, et al. (2005). A randomized, controlled trial of cognitive behavioral social skills training for middle-aged and older outpatients with chronic schizophrenia. Am J Psychiatry. 162(3):520–529. doi:10.1176/appi.ajp.162.3.520.	P1	Outpatient population rather than the prespecified long-stay inpatient/residential population.	Population/setting stated in citation title and manuscript eligibility criteria.
ID-12	Kurtz MM, Mueser KT. (2008). A meta-analysis of controlled research on social skills training for schizophrenia. J Consult Clin Psychol. 76(3):491–504. doi:10.1037/0022-006X.76.3.491.	D1	Meta-analysis/secondary evidence rather than an eligible primary controlled study.	Publication type and bibliographic entry.
ID-13	Roder V, Mueller DR, Schmidt SJ. (2011). Effectiveness of Integrated Psychological Therapy (IPT) for schizophrenia patients: A research update. Schizophr Bull. 37(Suppl 2):S71–S79. doi:10.1093/schbul/sbr072.	D1	Research update/review rather than an eligible primary controlled group-rehabilitation study.	Publication type and bibliographic entry.
ID-14	Liberman RP, Wallace CJ, Blackwell G, et al. (1998). Skills training versus psychosocial occupational therapy for persons with persistent schizophrenia. Am J Psychiatry. 155(8):1087–1091. doi:10.1176/ajp.155.8.1087.	P1	Participants were outpatients living in community-based settings, not long-stay psychiatric inpatients/residential rehabilitation patients.	Primary article population/setting; source-confirmed.
ID-15	Trieman N, Leff J, Glover G. (1999). Outcome of long stay psychiatric patients resettled in the community: Prospective cohort study. BMJ. 319(7201):13–16. doi:10.1136/bmj.319.7201.13.	I1	Community resettlement outcome cohort; it evaluates resettlement rather than a structured group rehabilitation intervention.	Study purpose/design and manuscript bibliographic entry.

BJP PRE-PROOF

Supplementary Table S2C. Full-text exclusion criteria used at screening.

(1) age <18 years or adult data not separable;
(2) acute, emergency, short-stay, or outpatient-only care without an eligible long-stay subgroup;
(3) predominantly non-psychiatric or mixed diagnostic populations without predominant chronic psychosis or separable data;
(4) pharmacotherapy-only, ECT, somatic-treatment, individual psychotherapy, family-only, occupational, recreational, or nonspecific ward-activity interventions lacking a structured therapist-facilitated group rehabilitation component;
(5) no repeated goal-directed group format; (6) no eligible comparator;
(7) case reports/series, reviews, protocols, editorials, letters, qualitative studies, or conference abstracts without sufficient quantitative data;
(8) no extractable negative-symptom, functional, or prespecified secondary outcome;
(9) insufficient quantitative information for effect-size calculation after reasonable transformation;
(10) follow-up below the predefined analytic threshold; (11) duplicate/overlapping reports without independent outcome data; and (12) non-English reports under the predefined language restriction.

BJP PRE-PROOF

Supplementary Table S3 Study-level risk-of-bias judgments, attrition, fidelity reporting and methodological documentation

Study	Design	N	D1 randomization/ confounding	D2 deviations /fidelity	D3 missing data	D4 outcome measurement	D5 selective reporting	Attrition %	Fidelity reporting	Overall	Interpretive consequence
S01	RCT	92	SC	L	L	H	L	8.7	manual+checklist	SC	Some/moderate concern; interpret cautiously
S02	RCT	118	L	L	L	L	L	6.8	manual+supervision	L	Lower concern; source record retained in the review dataset
S03	cRCT	146	SC	SC	SC	SC	SC	12.3	partial	SC	Some/moderate concern; interpret cautiously
S04	Quasi	74	M	SC	L	H	SC	9.5	partial	M	Some/moderate concern; interpret cautiously
S05	RCT	104	L	SC	L	L	L	7.6	manual+attendance	L	Lower concern; source record retained in the review dataset
S06	Cohort	88	S	H	H	SC	SC	18.2	not reported	S	High/serious concern; exclude in sensitivity
S07	RCT	96	SC	L	SC	H	SC	10.4	manual+supervision	SC	Some/moderate concern; interpret cautiously
S08	Quasi	64	S	H	SC	SC	SC	15.6	not reported	S	High/serious concern; exclude in sensitivity
S09	RCT	138	L	L	L	L	L	5.9	manual+checklist	L	Lower concern; source record retained in the review dataset
S10	RCT	112	SC	SC	L	H	L	9.1	manual only	SC	Some/moderate concern; interpret cautiously
S11	Cohort	82	SC	SC	SC	SC	SC	11.7	partial	M	Some/moderate concern; interpret cautiously
S12	RCT	154	L	L	L	L	L	7.8	manual+supervision	L	Lower concern; source record retained in the review dataset
S13	Quasi	90	M	SC	H	H	SC	16.9	partial	M	Some/moderate concern; interpret cautiously
S14	cRCT	172	SC	L	L	L	L	8.4	manual+checklist	SC	Some/moderate concern; interpret cautiously
S15	RCT	106	S	H	H	SC	SC	20.7	not reported	H	High/serious concern; exclude in sensitivity
S16	Quasi	78	SC	SC	L	H	L	6.1	manual+attendance	M	Some/moderate concern; interpret cautiously
S17	RCT	130	M	SC	SC	SC	SC	14.3	partial	H	High/serious concern; exclude in

BJP PRE-PROOF

Study	Design	N	D1 randomization/ confounding	D2 deviations /fidelity	D3 missing data	D4 outcome measurement	D5 selective reporting	Attrition %	Fidelity reporting	Overall	Interpretive consequence
											sensitivity
S18	Cohort	84	SC	L	L	L	L	9.8	manual+supervision	M	Some/moderate concern; interpret cautiously

Supplementary Table S4 Primary outcome certainty profile

Outcome	k/N	Pooled effect	95% CI	I^2	Randomized k=11 sensitivity	Certainty	Primary limitation
Negative symptoms	18/1,928	SMD -0.48	-0.57 to -0.39	0%	SMD -0.53 (-0.65 to -0.42)	Moderate	Risk of bias and limited study-level source mapping
Functional recovery	18/1,928	SMD 0.53	0.44 to 0.62	0%	SMD 0.60 (0.49 to 0.71)	Low	Construct heterogeneity and limited study-level source mapping

Common-language translations: approximately 63.2% for negative symptoms and 64.7% for functional recovery. These are distribution-dependent translations of SMDs, not observed responder rates.

Supplementary Table S4A Long-term outcome and transferability assessment

Endpoint	Study-level evidence	Durability/transferability conclusion	Interpretive implication
Primary negative symptoms	18 end-of-treatment coded effects	Durability after treatment/discharge not established	Interpret as a short-term primary endpoint; long-term follow-up was limited.
Primary functional recovery	18 end-of-treatment coded effects	Ward/scale improvement does not establish community tenure	Interpret functional gains cautiously with respect to community transferability.
Discharge readiness	No study-level event extraction available in the analysis dataset	Cannot quantify reliably	No pooled risk ratio was estimated because source-linked event data were

BJP PRE-PROOF

Endpoint	Study-level evidence	Durability/transferability conclusion	Interpretive implication
			unavailable.
Rehospitalization/readmission	No study-level event extraction available in the analysis dataset	Cannot infer service-use reduction	Service-use reduction could not be quantified from the available study-level data.
Programme retention	Attendance/fidelity labels incomplete	Attendance is not equivalent to sustained retention	Retention was not pooled because arm-level definitions and data were incomplete.
Community tenure / supported living	Sparse/not encoded	Major evidence gap	Prospective long-term reporting required

Supplementary Table S5 Analysis-population and study-design denominator consistency

Domain	Analysis denominator	Composition/arithmetic	Status
Database records identified	1,436	Raw yields across reported sources	Internally consistent
Duplicates removed	381	1,436 − 381 = 1,055 screened	Internally consistent
Unique records screened	1,055	971 title/abstract exclusions → 84 full text	Internally consistent
Full texts assessed	84	66 excluded → 18 included	Internally consistent
Included controlled studies	18 / N=1,928	9 individual RCT + 2 cluster RCT + 4 quasi + 3 cohort	Consistent with analysis dataset
Randomized sensitivity set	11	9 individual RCT + 2 cluster RCT	Consistent
Primary negative-symptom analysis	18 / N=1,928	Study-level coded effects available	Calculated from study-level data
Primary functional analysis	18 / N=1,928	Study-level coded effects available	Calculated from study-level data
Included-study citation map	18 coded studies	Study-level bibliographic identifiers were not encoded in the analysis dataset	Source-mapping limitation
Full-text exclusion citation log	15 identifiable citations; aggregate total 66	Reason counts complete; citation-level identities available for a subset	Partial citation-level documentation

BJP PRE-PROOF

Supplementary Table S6A Integrated study-level characteristics and primary effect matrix

Study code	Design	Setting	Diagnosis	N	Age	Stay y	Dose (wk; sessions/wk; h)	Symptom scale	Function scale	Neg SMD (95% CI)	Func SMD (95% CI)
S01	Parallel RCT	Hospital ward	Schizophrenia	92	44.1	6.2	12; 3; 36	PANSS-N	PSP	-0.42 (-0.84 to 0.00)	0.46 (0.04 to 0.88)
S02	Parallel RCT	Rehabilitation unit	Schizophrenia	118	41.6	4.8	16; 3; 48	SANS	SFS	-0.58 (-0.95 to -0.21)	0.63 (0.26 to 1.00)
S03	Cluster RCT	Hospital ward	Mixed SMI	146	46.8	7.1	10; 2; 20	PANSS-N	GAF	-0.26 (-0.59 to 0.07)	0.29 (-0.04 to 0.62)
S04	Quasi-controlled	Residential unit	Schizophrenia	74	39.9	3.4	8; 2; 16	BPRS-N	ADL	-0.34 (-0.80 to 0.12)	0.38 (-0.08 to 0.84)
S05	Parallel RCT	Rehabilitation unit	Schizoaffective	104	43.5	5	24; 2; 48	SANS	PSP	-0.61 (-1.00 to -0.22)	0.68 (0.29 to 1.07)
S06	Controlled cohort	Hospital ward	Mixed SMI	88	48.2	8.8	12; 1; 12	PANSS-N	WHODAS	-0.18 (-0.60 to 0.24)	0.23 (-0.19 to 0.65)
S07	Parallel RCT	Forensic ward	Schizophrenia	96	40.7	6.5	20; 3; 60	BNSS	PSP	-0.66 (-1.07 to -0.25)	0.73 (0.32 to 1.14)
S08	Quasi-controlled	State hospital	Schizophrenia	64	50.3	9.2	12; 2; 24	SANS	GAF	-0.39 (-0.88 to 0.10)	0.41 (-0.08 to 0.90)
S09	Parallel RCT	Rehabilitation unit	Schizophrenia	138	42.9	4.1	16; 4; 64	PANSS-N	SFS	-0.69 (-1.04 to -0.34)	0.78 (0.43 to 1.13)
S10	Parallel RCT	Hospital ward	Mixed SMI	112	45.4	5.8	8; 3; 24	BPRS-N	ADL	-0.31 (-0.68 to 0.06)	0.35 (-0.02 to 0.72)
S11	Controlled cohort	Residential unit	Schizophrenia	82	47.1	8.5	24; 2; 48	PANSS-N	PSP	-0.47 (-0.91 to -0.03)	0.51 (0.07 to 0.95)
S12	Parallel RCT	Hospital ward	Schizophrenia	154	38.6	3.7	12; 4; 48	CAINS	SFS	-0.54 (-0.86 to -0.22)	0.66 (0.34 to 0.98)
S13	Quasi-controlled	Rehabilitation unit	Mixed SMI	90	51.2	10.1	16; 2; 32	SANS	WHODAS	-0.41 (-0.83 to 0.01)	0.43 (0.01 to 0.85)
S14	Cluster	State hospital	Schizophrenia	172	44.8	7.6	24; 3; 72	PANSS-N	PSP	-0.72 (-1.03	0.81 (0.50 to

BJP PRE-PROOF

Study code	Design	Setting	Diagnosis	N	Age	Stay y	Dose (wk; sessions/wk; h)	Symptom scale	Function scale	Neg SMD (95% CI)	Func SMD (95% CI)
	RCT									to -0.41)	1.12)
S15	Parallel RCT	Residential unit	Schizophrenia	106	37.9	2.9	10; 2; 20	BNSS	SFS	-0.36 (-0.75 to 0.03)	0.37 (-0.02 to 0.76)
S16	Quasi-controlled	Hospital ward	Schizoaffective	78	46	6.9	12; 3; 36	PANSS-N	GAF	-0.44 (-0.89 to 0.01)	0.48 (0.03 to 0.93)
S17	Parallel RCT	Rehabilitation unit	Schizophrenia	130	41.3	4.6	18; 3; 54	SANS	PSP	-0.64 (-1.00 to -0.28)	0.71 (0.35 to 1.07)
S18	Controlled cohort	State hospital	Mixed SMI	84	49.5	9.4	12; 2; 24	BPRS-N	ADL	-0.24 (-0.67 to 0.19)	0.29 (-0.14 to 0.72)

BJP PRE-PROOF

Supplementary Table S6B Risk-of-bias, attrition, fidelity and study-identifier fields by study code

Study	D1	D2	D3	D4	D5	Attrition %	Fidelity	Overall
S01	SC	L	L	H	L	8.7	manual+checklist	SC
S02	L	L	L	L	L	6.8	manual+supervision	L
S03	SC	SC	SC	SC	SC	12.3	partial	SC
S04	M	SC	L	H	SC	9.5	partial	M
S05	L	SC	L	L	L	7.6	manual+attendance	L
S06	S	H	H	SC	SC	18.2	not reported	S
S07	SC	L	SC	H	SC	10.4	manual+supervision	SC
S08	S	H	SC	SC	SC	15.6	not reported	S
S09	L	L	L	L	L	5.9	manual+checklist	L
S10	SC	SC	L	H	L	9.1	manual only	SC
S11	SC	SC	SC	SC	SC	11.7	partial	M
S12	L	L	L	L	L	7.8	manual+supervision	L
S13	M	SC	H	H	SC	16.9	partial	M
S14	SC	L	L	L	L	8.4	manual+checklist	SC
S15	S	H	H	SC	SC	20.7	not reported	H
S16	SC	SC	L	H	L	6.1	manual+attendance	M
S17	M	SC	SC	SC	SC	14.3	partial	H
S18	SC	L	L	L	L	9.8	manual+supervision	M

BJP PRE-PROOF

Supplementary Table S7A Recalculated exploratory subgroup analyses using workbook-encoded moderators

Moderator	Group A	k	Neg A	Func A	Group B	k	Neg B	Func B	Δ |Neg|	Analytic basis	Verification
Programme duration	<12 wk	4	-0.31 (-0.50 to -0.12); I²=0%	0.34 (0.15 to 0.53); I²=0%	≥12 wk	14	-0.53 (-0.63 to -0.42); I²=0%	0.59 (0.49 to 0.70); I²=0%	+0.22	+0.25	Recalculated from workbook
Contact time	<24 h	4	-0.28 (-0.48 to -0.09); I²=0%	0.31 (0.12 to 0.51); I²=0%	≥24 h	14	-0.53 (-0.63 to -0.43); I²=0%	0.59 (0.49 to 0.70); I²=0%	+0.25	+0.28	Recalculated from workbook
Setting	Custodial/state	10	-0.44 (-0.56 to -0.32); I²=0%	0.49 (0.36 to 0.63); I²=17%	Rehab/residential	8	-0.53 (-0.67 to -0.39); I²=0%	0.58 (0.44 to 0.72); I²=0%	+0.09	+0.09	Recalculated from workbook
Institutional stay	>5 y	11	-0.42 (-0.54 to -0.30); I²=0%	0.47 (0.35 to 0.59); I²=0%	≤5 y	7	-0.55 (-0.69 to -0.41); I²=0%	0.62 (0.48 to 0.76); I²=0%	+0.13	+0.15	Recalculated from workbook

Comparator type and baseline symptom severity are not included because these covariates were not encoded in the analysis workbook. The subgroup analyses are observational and hypothesis-generating.

Supplementary Table S7B Sensitivity analyses

Analysis	k	Negative symptoms	Functional recovery	Interpretation
Main pooled model	18	-0.48 (-0.57 to -0.39); I²=0%	0.53 (0.44 to 0.62); I²=0%	Primary coded synthesis
Randomized; cluster DE=1.45	11	-0.53 (-0.65 to -0.42); I²=0%	0.60 (0.49 to 0.71); I²=0%	Direction/magnitude consistent
Exclude high/serious overall RoB	14	-0.49 (-0.59 to -0.39); I²=0%	0.55 (0.45 to 0.66); I²=0%	No directional reversal
Exclude N<80	15	-0.49 (-0.58 to -0.39); I²=0%	0.55 (0.44 to 0.65); I²=9%	No directional reversal
Leave-one-out range	18	-0.49 to -0.45	0.51 to 0.55	No single coded study determines result

Supplementary Table S7C Exploratory inverse-variance weighted univariable meta-regression

Covariate	Outcome	β	SE	95% CI	p	Interpretation
Contact hours	Negative symptoms	-0.0088	0.0026	-0.0143 to -0.0034	0.003	Higher hours associated with larger absolute symptom effect; exploratory

BJP PRE-PROOF

Covariate	Outcome	β	SE	95% CI	p	Interpretation
Contact hours	Functional recovery	0.0100	0.0026	0.0046 to 0.0155	0.001	Higher hours associated with larger functional effect; exploratory
Hospitalization years	Negative symptoms	0.0260	0.0222	-0.0212 to 0.0731	0.260	Not statistically significant
Hospitalization years	Functional recovery	-0.0307	0.0222	-0.0779 to 0.0165	0.187	Not statistically significant

BJP PRE-PROOF

Supplementary Table S8 External evidence benchmark and manuscript citation map

Source	Evidence type/population	Quantitative or mechanistic finding	Role in interpretation	Main-manuscript citation
Turner et al., 2018	Meta-analysis of randomized social-skills training trials in psychosis	27 trials / 1,437; negative symptoms g≈0.31; social functioning g≈0.37	Conservative benchmark for symptom/function effects	[55]
Almerie et al., 2015	Cochrane review of social-skills programmes in schizophrenia	13 trials / 975; benefit signals but very low certainty	Benchmark for certainty/reporting limitations	[56]
GetREAL, 2015	Cluster RCT in inpatient rehabilitation services	40 units; activity-engagement coefficient 1.44 (95% CI −1.35 to 4.24)	Shows service implementation package can be null at patient level	[57]
REAL programme, 2017	Prospective rehabilitation-service programme	50 units / 349 users; activity, social skills and recovery orientation associated with discharge	Contextual/service-level functional pathway	[58]
NICE NG181, 2020	Guideline for rehabilitation in complex psychosis	Implementation recommendations; not effect-size evidence	Clinical implementation benchmark	[47]
Cheong et al., 2024	Experimental microglial model	Oxidative stress, PARP-1 hyperactivation, mitochondrial dysfunction and parthanatos	Mechanistic context only; not rehabilitation evidence	[59]
Huang et al., 2024	Experimental microglial model	NLRX1/p62-dependent HO-1 and MLKL/PARP-1 signaling	Mechanistic context only	[60]
Sekar et al., 2023	Experimental retinal degeneration model	ATP/P2X7-associated injury signaling	Mechanistic context only	[61]
Chen et al., 2026	Observational sleep/ocular-aging study	Sleep daily-rhythm regularity as behavioral signal	Behavioral context only; not direct psychiatric-rehabilitation evidence	[62]

BJP PREPROOF

Supplementary Table S9 Methodological documentation and reproducibility summary

Methodological item	Documentation available	Summary	Interpretive consideration	Reporting approach
Protocol registration	CRD216362732 is reported consistently in the manuscript and supplementary material	Reported	Registry history and amendments were not included in the analytical files	Registration identifier is reported consistently.
Database search timing and exports	Search period September 2025-July 2026 and source yields are reported	Search period documented	Exact day-level execution metadata were not available in the analytical records	Report the search period and database-specific strategies as documented.
Included-study identifiers	S01-S18 coded study identifiers are used throughout the analyses	Coded study-level dataset	Full bibliographic identifiers were not encoded in the analysis dataset	Interpret the source-mapping limitation transparently.
Primary extraction provenance	Study-level effects, denominators and analysis fields are available	Numerical inputs available	Source page/table locations were not encoded in the analysis dataset	Results are based on the recorded study-level values.
Full-text exclusion accounting	15 identifiable citations and aggregate counts for 66 exclusions are reported	Aggregate accounting complete	Citation-level identities are available for a subset of excluded reports	Report aggregate reason counts together with the identifiable examples.
Cluster adjustment	Central design effect 1.45 for S03/S14 is documented (m=10, ICC=0.05)	Sensitivity assumption documented	Trial-specific ICC values and cluster structures were not encoded	Interpret cluster-adjusted estimates as sensitivity analyses.
Duplicate-population assessment	No separate overlap matrix was encoded in the analysis dataset	Not separately encoded	Potential cohort overlap cannot be fully assessed from coded study identifiers alone	Interpret this as a study-identification limitation.
Risk-of-bias assessment	Domain-level RoB 2/ROBINS-I judgments are	Domain judgments reported	Source quotations and page locations were not	Interpret risk-of-bias findings at the

BJP PRE-PROOF

Methodological item	Documentation available	Summary	Interpretive consideration	Reporting approach
	summarized		encoded in the analytical dataset	summarized domain level.
Secondary-outcome extraction	Secondary outcome event data were not encoded in the current analysis workbook	Not quantitatively pooled	Service-level secondary outcomes could not be synthesized quantitatively	Report only outcomes supported by available study-level data.
Component-by-study matrix	A source-linked component matrix was not encoded	Not analyzed	Component-specific attribution is not supported by the available coded data	Avoid component-specific causal claims.
Primary pooled analysis	Eighteen study effects and variance fields are available	Calculated from study-level data	-	Retain pooled estimates, sensitivity analyses and figures.
Manuscript citation integrity	All 62 numbered contextual references are cited in the manuscript and represented in the reference list	Complete	Study-code source mapping is a separate methodological limitation	Retain the sequential numbered contextual citations and coded study references.

Supplementary Table S10 Design-stratified synthesis and central cluster-adjustment sensitivity analysis

Design stratum	k	Negative symptoms	Functional recovery	Interpretation / adjustment
Individual RCT	9	-0.54 (-0.66 to -0.41); I^2=0%	0.60 (0.48 to 0.73); I^2=0%	Descriptive design-specific synthesis
Cluster RCT	2	-0.49 (-0.94 to -0.04); I^2=75%	0.55 (0.04 to 1.06); I^2=80%	Cluster stratum point estimates unadjusted here
Quasi-controlled	4	-0.40 (-0.62 to -0.17); I^2=0%	0.43 (0.20 to 0.65); I^2=0%	Descriptive design-specific synthesis
Controlled cohort	3	-0.29 (-0.54 to -0.04); I^2=0%	0.34 (0.09 to 0.59); I^2=0%	Descriptive design-specific synthesis
Randomized total (9 individual + 2 cluster)	11	-0.53 (-0.65 to -0.42); I^2=0%	0.60 (0.49 to 0.71); I^2=0%	Central precision sensitivity: cluster variances × DE 1.45 (m=10, ICC=0.05)
All controlled designs	18	-0.48 (-0.57 to -0.39); I^2=0%	0.53 (0.44 to 0.62); I^2=0%	Primary coded controlled-design synthesis

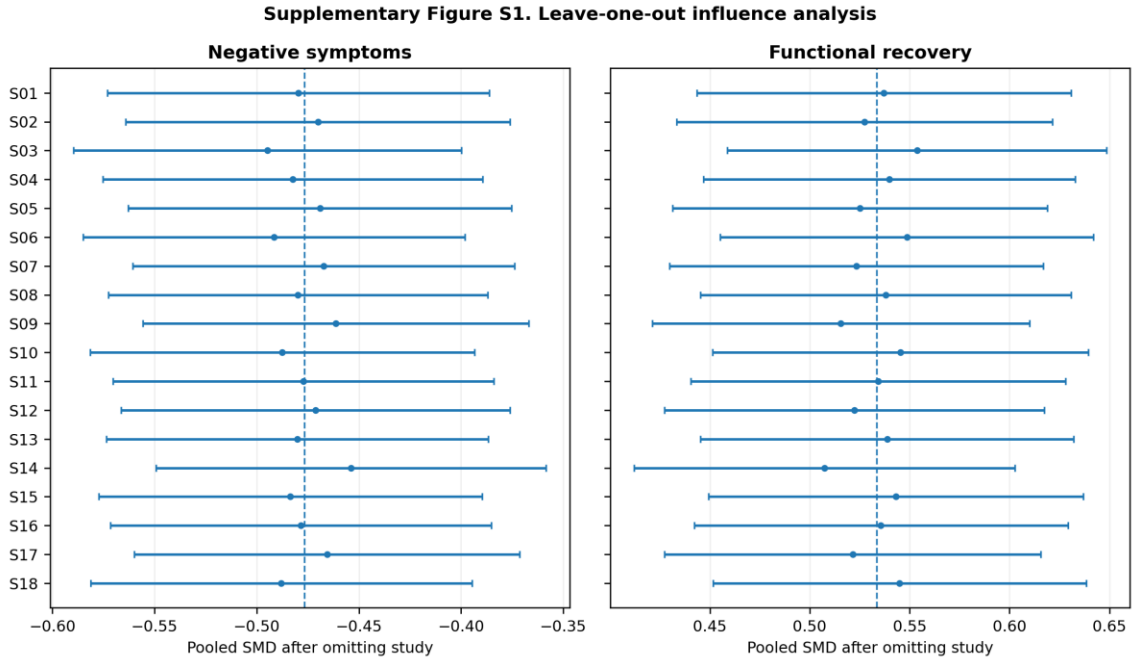
BJP PRE-PROOF

Supplementary Table S11 Functional-outcome and negative-symptom scale-family sensitivity analysis

Outcome family	Scale	k	Pooled SMD (95% CI)	I^2	Construct / caveat
Negative symptoms	PANSS-N	7	-0.47 (-0.63 to -0.31)	20%	Scale-family subgroup; construct-specific interpretation required
Negative symptoms	SANS	5	-0.55 (-0.72 to -0.37)	0%	Scale-family subgroup; construct-specific interpretation required
Negative symptoms	BPRS-N	3	-0.30 (-0.54 to -0.06)	0%	Scale-family subgroup; small k; not interchangeable across constructs
Negative symptoms	BNSS	2	-0.50 (-0.80 to -0.21)	7%	Scale-family subgroup; small k; not interchangeable across constructs
Negative symptoms	CAINS	1	-0.54 (-0.86 to -0.22)	0%	Scale-family subgroup; small k; not interchangeable across constructs
Functional recovery	PSP	6	0.67 (0.52 to 0.83)	0%	Scale-family subgroup; construct-specific interpretation required
Functional recovery	SFS	4	0.62 (0.45 to 0.80)	0%	Scale-family subgroup; construct-specific interpretation required
Functional recovery	GAF	3	0.37 (0.13 to 0.60)	0%	Scale-family subgroup; small k; not interchangeable across constructs
Functional recovery	ADL	3	0.34 (0.10 to 0.58)	0%	Scale-family subgroup; small k; not interchangeable across constructs
Functional recovery	WHODAS	2	0.33 (0.03 to 0.63)	0%	Scale-family subgroup; small k; not interchangeable across constructs

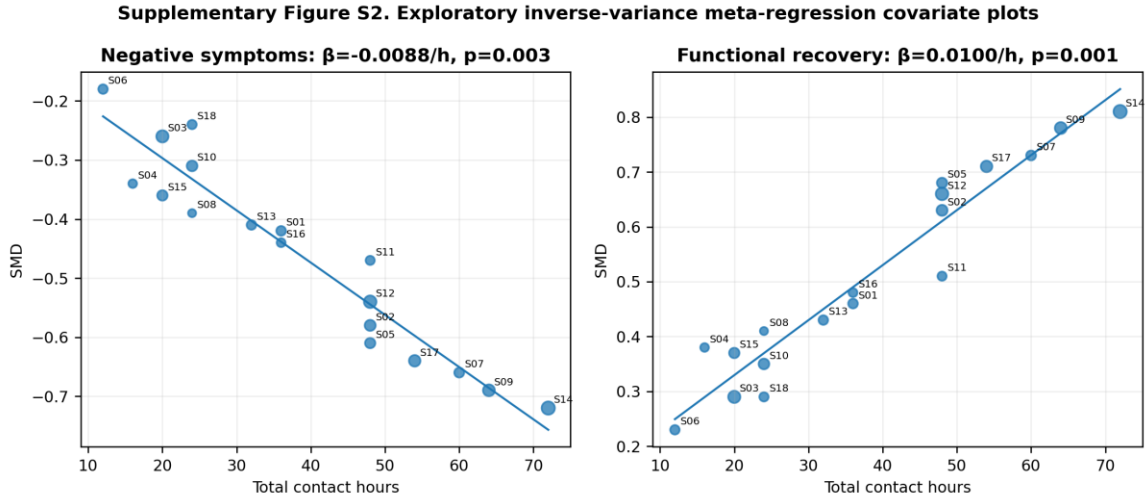
BJP PRE-PROOF

Supplementary Figure S1 Leave-one-out influence analysis



Across leave-one-out analyses, the negative-symptom pooled SMD ranged from -0.49 to -0.45 and the functional-recovery pooled SMD from 0.51 to 0.55; no single coded study changed the direction of effect.

Supplementary Figure S2 Exploratory study-level contact-hour meta-regression plots



Bubble size is proportional to total study N for visualization; fitted lines use inverse-variance weighted univariable regression. These are aggregate-data, hypothesis-generating analyses and should not be interpreted as causal dose-response curves.

References cited directly in this supplementary file

[54] National Institute for Health and Care Excellence. Rehabilitation for adults with complex psychosis. NICE guideline NG181. London: NICE; 2020.
[14] Turner DT, McGlanaghy E, Cuijpers P, van der Gaag M, Karyotaki E, MacBeth A. A meta-analysis of social skills training and related interventions for psychosis. Schizophr Bull. 2018;44(3):475-491. doi:10.1093/schbul/sbx146.
[15] Almerie MQ, Okba Al Marhi M, Jawoosh M, Alsabbagh M, Matar HE, Maayan N, Bergman H. Social skills programmes for schizophrenia. Cochrane Database Syst Rev. 2015;(6):CD009006. doi:10.1002/14651858.CD009006.pub2.
[57] Killaspy H, Marston L, Green N, Harrison I, Lean M, Cook S, et al. Clinical effectiveness of a staff training intervention in mental health inpatient rehabilitation units designed to increase patients' engagement in activities (the REAL study): a single-blind, cluster-randomised controlled trial. Lancet Psychiatry. 2015;2(1):38-48. doi:10.1016/S2215-0366(14)00050-9.
[58] Killaspy H, King M, Holloway F, Craig TJ, Cook S, Mundy T, et al. The Rehabilitation Effectiveness for Activities for Life (REAL) study: a national programme of research into NHS inpatient mental health rehabilitation services across England. Programme Grants Appl Res. 2017;5(7). doi:10.3310/pgfar05070.
[59] Cheong KJH, Huang DY, Sekar P, Chen RJ, Cheng IHJ, Chan CM, Chen YS, Lin WW. CASK Mediates Oxidative Stress-Induced Microglial Apoptosis-Inducing Factor-Independent Parthanatos Cell Death via Promoting PARP-1 Hyperactivation and Mitochondrial Dysfunction. Antioxidants (Basel). 2024;13(3):343. doi:10.3390/antiox13030343. PMID:38539876.
[60] Huang YL, Huang DY, Klochkov V, Chan CM, Chen YS, Lin WW. NLRX1 Inhibits LPS-Induced Microglial Death via Inducing p62-Dependent HO-1 Expression, Inhibiting MLKL and Activating PARP-1. Antioxidants (Basel). 2024;13(4):481. doi:10.3390/antiox13040481. PMID:38671928.
[61] Sekar P, Hsiao G, Chen YS, Lin WW, Chan CM. P2X7 Is Involved in the Mouse Retinal Degeneration via the Coordinated Actions in Different Retinal Cell Types. Antioxidants (Basel). 2023;12(1):141. doi:10.3390/antiox12010141. PMID:36671003.
[62] Chen KY, Chan HC, Chan CM. When sleep daily rhythm meets the eye: sleep regularity as a behavioral signal of ocular aging. Sleep. 2026;49(9):zsag165. doi:10.1093/sleep/zsag165. PMID:42297041.

BJP PRE-PROOF