

RESEARCH

Open Access



A secondary exploratory study of associations between patient- and clinician-reported clinical outcomes and fidelity for four evidence-based psychosis treatments

Torleif Ruud^{1,2*} , Jūratė Šaltytė Benth^{2,3}, Eva Biringer^{4,5}, Miriam Hartveit^{4,6}, Karin Drivenes⁷, Kristin Heiervang^{1,8}, Tordis S. Høifødt⁹, Vegard Ø. Haaland^{7,10}, Inge Joa^{11,12}, Jan Olav Johannessen¹¹, Karl Johan Johansen¹³, Bjørn Stensrud¹⁴, Espen Woldsen Haugom^{2,14}, Hanne Clausen^{1,14} and Gary R. Bond¹⁵

Abstract

Background Implementation of evidence-based practices (EBPs), measured as fidelity to the EBP model, is generally expected to yield significant positive clinical outcomes. However, this association has only partially been established for EBPs used in psychosis treatment. From a cluster-randomized controlled trial (CRCT), we previously reported on the effects of implementation support for four EBPs for psychosis, using fidelity as the measure of implementation success. The current secondary, exploratory study used data from the non-blinded CRCT parent study to investigate the associations between patient- and clinician-reported outcomes and fidelity for these four EBPs.

Methods Clinical outcomes were measured in a cohort of 325 patients over three six-month periods. Primary outcomes were BASIS-24 (patient-reported) and HoNOS (clinician-reported). Secondary outcomes were selected subscales of these two measures. The EBPs were Physical Health Care, Antipsychotic Medication Management, Family Psychoeducation, and Illness Management and Recovery. In the CRCT, each of 39 clinical units across six health trusts selected two EBPs for implementation. Units were randomized to the intervention group (implementation support) for one EBP and the control group (written manual) for the other. Fidelity of the four EBPs was measured at baseline and every six months for 18 months. We analyzed the associations between outcomes and fidelity using linear mixed models.

Results BASIS-24 and HoNOS showed improvements for the total sample at 6 and 12 months, and two patient-reported subscales, Symptoms and Relationships, showed improvement at 6 months within two different EBP subsamples. However, no positive associations were found between secondary outcomes and EBP fidelity.

Conclusions Despite some improvements in primary and secondary outcomes over the first 6 to 12 months, we found no positive associations between outcomes and fidelity. Sample size, attrition, trial design, variance in variables, measurement properties, and low exposure, as well as interaction between such factors, might have contributed to

*Correspondence:

Torleif Ruud
torleif.ruud@medisin.uio.no

Full list of author information is available at the end of the article



© The Author(s) 2025. **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

our failure to find positive associations between outcomes and fidelity. Future studies of the association between outcomes and fidelity should involve large samples, use outcome and exposure measures closely related to the EBPs, and track cohorts from the beginning of treatment.

Trial registrations ClinicalTrials NCT03271242, retrospectively registered 31 August 2017.

Keywords Evidence-based treatments, Fidelity, Psychosis, Patient-reported outcome, Clinician-reported outcome, Physical health care, Antipsychotic medication management, Family psychoeducation, Illness management and recovery

Background

Evidence-based practices (EBPs) are defined as interventions that consistently lead to improved outcomes, as established by randomized controlled trials [1]. Within implementation science, fidelity measures the degree to which an implementation of an EBP adheres to the specific standards set by the EBP model [2]. Because fidelity is recognized as an intermediate measure of implementation success, it is widely used in implementation studies [3, 4]. Consequently, mental health services that maintain high fidelity to EBPs in psychosis treatment are expected to achieve positive patient outcomes. For several EBPs, specific fidelity scales have been developed to measure the essential components that account for the practice's effectiveness [1].

Several biological and psychosocial EBPs have been documented to effectively reduce symptoms and problems associated with psychosis [5–7]. To investigate the real-world implementation of some EBPs, we conducted a cluster-randomized controlled trial (CRCT) assessing the impact of implementation support for four specific EBPs in psychosis treatment [8]: Physical Health Care, Antipsychotic Medication Management, Family Psychoeducation, and Illness Management and Recovery. Each participating clinical unit selected two of these EBPs and was randomized to serve as an intervention site for one EBP and a control site for the other. The primary outcome measure of the CRCT was the implementation success of these EBPs, quantified using a specific fidelity scale for each practice.

Prior evidence for the four EBPs is mixed. Physical Health Care is well-supported, with systematic reviews showing that physical exercise leads to significant improvements in negative psychosis symptoms, though effects on positive symptoms are often negligible [9–11]. Conversely, the literature on Antipsychotic Medication Management is scarce regarding multi-component, comprehensive approaches, as most studies treat its elements separately. However, limited evidence suggests that training in medication management can improve symptoms [12, 13]. Evidence for Family Psychoeducation is also limited, but improvement in clinician-reported symptoms have been reported [7]. Finally, Illness Management and Recovery has yielded mixed results, with some studies

finding significant psychiatric symptom improvements [14–16] and others reporting non-significant findings [17–20].

Among 14 available fidelity scales for EBPs in psychosis treatment, eight have been reported to predict outcome [1]. However, the fidelity scales for the four specific EBPs utilized in our CRCT have not demonstrated this predictive capacity. Furthermore, to our knowledge, no study has yet examined the association between fidelity and outcome for any of these four EBPs. Therefore, we conducted this secondary exploratory study using data from the CRCT.

Aims

The aim of this 18-month secondary exploratory study was to investigate whether patient- and clinician-reported outcomes in a cohort of patients with psychosis were associated with the fidelity of the implementing clinical units to four respective EBPs across intervention and control sites. To facilitate this analysis, we first examined outcomes for the total sample across the 6-, 12-, and 18-month follow-up periods, and explored selected relevant secondary outcomes for specific EBP subsamples across the three six-month follow-up assessments.

Methods

Design

This secondary, exploratory study followed a cohort of patients for 18 months to examine whether clinical changes were associated with EBP fidelity across three consecutive six-month periods. We utilized data from an 18-month CRCT, which originally assessed the effect of implementation support on the implementation of four EBPs in psychosis treatment [8]. The trial was retrospectively registered (ClinicalTrials NCT03271242, 31 August 2017). We adhered to the CONSORT Extension guidelines for CRCTs [21]. In the non-blinded CRCT, 39 clinical units each selected two of four core EBPs for implementation. A pairwise randomization design was used: each unit was assigned to be an intervention site (receiving implementation support) for one EBP and a control site (receiving only a written manual) for the other. Consequently, every unit served as an intervention site for one of the four EBPs.

Table 1 Clinician-reported patient baseline characteristics and distribution of sites for each EBP

Samples	Baseline (n = 325)
Variables	n (%) / Mean (SD)
Sex, n (%)	
Male	191 (58.8)
Female	134 (41.2)
Age, mean (SD)	39.9 (12.7)
Age group, n (%)	
16–19	10 (3.1)
20–29	71 (21.8)
30–39	87 (26.8)
40–49	78 (24.0)
50–59	54 (16.6)
60–69	22 (6.8)
70+	3 (0.9)
Completed education, n (%)	
Primary school uncompleted	9 (2.8)
Primary school 7–10 years	96 (29.5)
Secondary school	82 (25.2)
Occupational education	54 (16.6)
Higher education 1–3 years	40 (12.3)
Higher education 4–6 years	23 (7.1)
Other education	10 (3.1)
Unknown (missing data)	11 (3.4)
Total time of provided mental health services, n (%)	
1 Less than 6 months	20 (6.2)
2 7–23 months	28 (8.6)
3 2–5 years	50 (15.4)
4 6–10 years	66 (20.3)
5 More than 10 years	147 (45.2)
Unknown (missing data)	14 (4.3)
Severity of mental illness (CGI), mean (SD)	4.03 (1.43)
Unknown (missing data), n (%)	16 (4.9)
Severity of mental illness (CGI), n (%)	
1 No mental illness	4 (1.2)
2 Minimal mental illness	58 (17.8)
3 Mild mental illness	47 (14.5)
4 Moderate mental illness	77 (23.7)
5 Marked mental illness	67 (20.6)
6 Severe mental illness	52 (16.0)
7 Very severe mental illness	4 (1.2)
Unknown (missing data)	16 (4.9)
Functioning (PSF), mean (SD)	4.02 (1.84)
Unknown (missing data), n (%)	8 (2.5)
Diagnosis ICD-10, n (%)	
Schizophrenia/paranoid psychosis	186 (57.2)
Schizoaffective disorder	60 (18.5)
Bipolar disorder	15 (4.6)
Other	41 (12.6)
Not reported	23 (7.1)
Patients in intervention/control sites for each EBP, n / n	
Physical Health Care	108 / 122
Antipsychotic Medication Management	64 / 57
Family Psychoeducation	39 / 74
Illness Management and Recovery	114 / 72

CGI: Clinical Global Impression, PSF: Practical and Social Functioning, EBP: Evidence based practice

Context

Norwegian mental health services are predominantly public services provided by 19 health trusts that also manage general hospital services for all age groups [22, 23]. Adult mental health and substance abuse services consist of acute and other inpatient hospital units, and community mental health centers (CMHCs) with outpatient clinics, mobile teams, and local inpatient units [24]. Each CMHC serves a local catchment area and collaborates with general practitioners and municipal primary health and social care providers. All public inpatient and outpatient mental health services are free of charge, with the exception of an outpatient consultation fee. This fee is subject to an annual maximum, after which all subsequent consultations for the remainder of the year are free.

Recruitment, samples and data collection

The leadership of the health trusts' mental health services determined the participation of all CMHCs and other relevant clinical units. The study involved mental health clinics from six of the 19 Norwegian health trusts, which served 38% of the country's population across both urban and rural settings. The 39 participating clinical units provided services to adults or adolescents with psychosis. These units included 26 CMHCs, 10 inpatient departments for adults with psychosis, and three combined outpatient and inpatient departments for adolescents.

We prospectively followed a total sample of 325 patients over the 18-month period, obtaining outcome data from both patients and clinicians at six-month intervals. This sample was recruited by clinicians across 32 of the 39 clinical units during a nine-month inclusion period. Seven clinical units did not recruit any patients. Local coordinators observed that in some units, the focus during the first months was mostly on learning and starting to implement the intervention EBP, while the parallel patient study with patient recruitment and data collection received less attention and efforts. Only patients who provided written informed consent were included. Sub-samples for each EBP were comprised of patients receiving treatment at clinical units that had selected that specific EBP for implementation and had been randomized to either the intervention or control site for that EBP.

Table 1 presents patient characteristics and the distribution of intervention and control sites for EBP subsample at baseline. The sample size was progressively reduced at the 6, 12, and 18-month follow-up periods due to patient discharges, the death of 13 patients (confirmed by the Norwegian Cause of Death Registry), and other dropouts. Attrition rates for patient-completed questionnaires were 30%, 27%, and 25% compared to the last available data collection, while the reductions for clinician-reported ratings were 37%, 30%, and 22% at 6, 12, and 18 months, respectively. Descriptive statistics

revealed only minor variations in characteristics across the remaining samples, though the rate of loss to follow-up was higher among males than females.

Evidence-based practices, randomization and implementation support

Evidence-based practices. The four EBPs chosen for the trial were among those recommended in the Norwegian national clinical guidelines for psychosis treatment [25]. Two were biological practices (Physical Health Care, Antipsychotic Medication Management) that the clinical units were already providing without measuring quality. Two were psychosocial practices (Family Psychoeducation, Illness Management and Recovery) that were new to nearly all participating units.

The four EBPs were defined by the following core components: Physical Health Care involved monitoring cardiovascular risk factors, treating physical illnesses, and promoting physical fitness, a healthy diet, smoking cessation, and dental and oral health. Antipsychotic Medication Management encompassed shared decision-making, somatic assessment, medication choice and dosage, limiting polypharmacy, maintaining a current medication list, monitoring adherence and symptoms, monitoring side effects, and managing medication discontinuation. Family Psychoeducation provided families with education, skills training, and support through meetings and workshops over the course of a year or longer. Illness Management and Recovery focused on psychoeducation (to enhance illness understanding), prevention (to reduce relapses and hospitalizations), behavioral training (to improve medication adherence), coping skills training (to mitigate symptoms), and social training (to strengthen social support). Previous papers have elaborated on these four EBPs in greater detail [26–29].

Randomization. In the CRCT, all clinical units received detailed descriptions of the four EBPs and selected two for implementation. They agreed to be randomized to serve as intervention site (receiving implementation support) for one practice and control site (receiving only written guidelines) for the other. Of the units, 26 selected Physical Health Care, 17 selected Antipsychotic Medication Management, 14 selected Family Psychoeducation, and 21 selected Illness Management and Recovery. Stratified randomization, performed by an external statistician, ensured an even distribution between the intervention and control conditions for each of the six possible pairs of practices.

The implementing support provided to intervention sites included a toolkit for the EBP. Clinicians participated in EBP-specific training, and implementation facilitators conducted frequent site visits. Additional telephone supervision was provided for the two psychosocial practices. Furthermore, clinicians received feedback

every six months based on the fidelity assessment and the results from a questionnaire to the staff about their experiences with the implementation process [30]. The intervention period spanned 18 months, lasting from September 2, 2016, to February 28, 2018. More details about the implementation support are available in the paper discussing its impact [8].

Measures

Outcome measures

For *primary patient-reported outcome measure*, we used the mean total score from the Behavior and Symptom Identification Scale 24 (BASIS-24). BASIS-24 has demonstrated acceptable reliability and validity for patients experiencing psychosis, as well as sensitivity to change [31–34]. As *secondary patient-reported outcome measures* for the EBP subsamples and for analyzing associations with fidelity, we used two BASIS-24 subscales: the Psychosis subscale (5 items), which relates to Antipsychotic Medication Management, and the Relationships subscale (6 items), which relates to Family Psychoeducation and Illness Management and Recovery. These subscales were more relevant to the respective EBPs than the total BASIS-24, which consists of several dimensions. The BASIS-24 contains 24 items rated on a scale of 0 to 4, with higher scores indicating more serious symptoms or issues. The Norwegian translation was approved by the license holder following adjustments based on back-translation. We calculated mean scores for the total scale and the selected subscales according to the instructions and item weights specified in the BASIS-24 manual [35]. In our study, the BASIS-24 total scale showed fair internal consistency (Cronbach's alpha) across all time points: 0.71 at baseline, 0.72 at 6 months, 0.71 at 12 months, and 0.69 at 18 months. The alpha range was 0.75 to 0.76 for the Psychosis subscale and 0.72 to 0.80 for the Relationships subscale.

For the *primary clinician-reported outcome measure*, we used the mean total score of the Health of the Nation Outcome Scale (HoNOS) [36]. HoNOS is a widely recognized rating scale to address core psychiatric problems while remaining brief enough for clinicians to complete amid their busy schedules. As *secondary clinician-reported outcome measures* for the EBP subsamples and for analyzing associations with fidelity, we used two HoNOS subscales: the Symptoms subscale (3 items), related to Antipsychotic Medication Management, and the Functioning subscale (2 items), related to Family Psychoeducation and Illness Management and Recovery. These subscales were more relevant to the respective EBPs than the total HoNOS scale which consists of several dimensions. HoNOS was familiar to the clinicians, as it constituted part of the standard data set used by health trusts for routine reporting to the Norwegian

Patient Register. Clinicians were required to undertake 45-minutes e-learning training on proper HoNOS rating. Developed especially for patients with serious mental illness, HoNOS consists of 12 items rated on a scale from 0 = “No problem” to 4 = “Severe or very severe problem” [36].

Because the selected secondary outcome measures were more relevant to the content of the EBPs and their fidelity scales, we chose to use these secondary outcomes—rather than the primary outcomes (total scales)—for the exploratory analyses of associations between patient- and clinician-reported outcomes and fidelity.

Missing values for the BASIS-24 and HoNOS scores at 6 months ($n=19$ for BASIS-24 and $n=13$ for HoNOS) and at 12 months ($n=15$ for BASIS-24 and $n=11$ for HoNOS) were interpolated from existing values at adjacent time points. The interpolation method used the previous time point's value plus half of the difference between the previous and subsequent time points. Overall, the interpolated data constituted 8% of the BASIS-24 data and 7% of the HoNOS data.

Explanatory variables

Fidelity of the four EBPs was assessed by trained fidelity assessors using a specific fidelity scale for each EBP. Two of these scales (Family Psychoeducation Fidelity Scale, Illness Management and Recovery Fidelity Scale) had been developed and utilized in prior studies [26, 27]. The other two scales (Physical Health Care Fidelity Scale, Antipsychotic Medication Management Fidelity Scale) were developed for our CRCT due to the absence of established fidelity scales for these EBPs [28, 29]. All four fidelity scales exhibited adequate psychometric properties (interrater reliability, internal consistency, sensitivity to change, feasibility) as corroborated in the articles cited above.

Fidelity assessment at 6 months was conducted in March–April 2017, at 12 months in September–October 2017, and at 18 months in March–April 2018. Two trained assessors rated fidelity for the two practices being implemented in each clinical unit. The assessors conducted in-person site visits, rating fidelity independently and resolving any discrepancies through consensus. The fidelity visits included interviews with managers and clinicians, reviews of written material for each EBPs, and review of documentations found in 10 randomly selected patient records for Physical Health Care and Antipsychotic Medication Management.

The fidelity achieved during the CRCT is reported in detail in a previous paper [8]. Summarized, the increase in fidelity scores (within range 1–5) over 18 months was significantly greater for intervention sites compared to control sites for three of the four EBPs. Effect sizes

(Cohen's d) for increase in difference between intervention and control sites of mean fidelity scores were 2.24 for Illness Management and Recovery, 0.68 for Physical Health Care, 0.71 for Antipsychotic Medication Management, and 0.27 for Family Psychoeducation. Most improvements occurred during the first 12 months.

Clinician-recorded patients' exposure to each EBP.

The two biological interventions are administered to most patients with psychosis, while the two psychosocial interventions are provided by a limited number of clinicians to a restricted selection of patients. We therefore identified the need for a measure of how much each patient was exposed to the respective intervention to adjust for this in the analysis of associations between outcome and fidelity. Penetration (exposure) has been defined as one of eight implementation outcomes [3].

Exposure for the 0–6 months period was registered by clinicians as time-period and number of sessions attended. Because clinicians found it too time-consuming to extract these data from patient records, this registration was modified at 12 and 18 months. Instead, clinicians then rated exposure to components of each EBP on a five-point ordinal scale considering the scope, regularity and systematicity of the EBP: 0='Nothing in this period', 1='a little/occasionally, not systematic', 2='Partially systematic/short time (weeks)', 3='Quite systematic/longer time (months)', and 4='Systematic/regularly, all the time'. To ensure uniformity, data reported at 6 months were transformed to align with the ordinal scale used at 12 and 18 months. This measure with the ordinal scale was developed by the research group for this study but it was not tested for validity or interrater reliability. An English translation of the registration form for exposure is found in the Supplementary Material, and the calculation of the estimated exposure for each EBP (mean of the relevant items) is shown below the registration form in the Supplementary Material. Table A in Supplementary Material shows descriptive statistics of the estimated exposure to each of the four EBPs during each six-months period as reported by the clinicians. The exposure was highest for the two biological EBPs and lowest for the two psychosocial EBPs.

Patient characteristics. From data collected at patient inclusion in the trial, we selected five baseline patient characteristics to adjust for in our analyses of associations between secondary outcomes and fidelity: age, sex, severity of mental illness, functioning, and duration of contact with mental health services. Severity was rated by clinicians using the Clinical Global Impression (CGI) [37]. CGI is a single seven-point scale ranging from 1 = “normal, not ill” to 7 = “among the most severely ill”, and it is shown to be reliable and valid [37]. Functioning was rated by clinicians interviewing patients using the revised Practical and Social Functioning (PSF) scale version

2 with 32 items and a five-point scale ranging from 1 = “completely incorrect” to 5 = “completely correct”. PSF has been shown to have adequate internal consistency [38], and Cronbach’s alpha was 0.69 at baseline in our study. The total duration of mental health services provided to the patient was rated by clinicians on an ordinal variable with a six-point scale defined for the trial as a proxy measure of mental illness duration (see Table 1 for the ordinal scale).

Data collection

Patients reported on their mental health by completing the BASIS-24 questionnaire. Clinicians assessed the patients’ mental problems using the HoNOS and collected additional baseline information on patient characteristics, as well as rated the exposure to the EBPs. Clinical unit personnel collected the data at the time of patient inclusion (baseline, or 0 months) and again at 6, 12, and 18 months thereafter. The inclusion period spanned nine months, running from June 1, 2016, to February 28, 2017.

Data analysis

We reported descriptive statistics of clinician-reported baseline patient characteristics for the total sample, as well as the distribution of each EBP subsample in intervention and control sites. Primary and secondary outcomes were presented by means and standard deviations (SDs) at each time point and each period within intervention and control sites, respectively.

To analyze the changes in BASIS-24 and HoNOS (primary outcomes) for the total sample from baseline to 6, 12, and 18 months, we estimated linear mixed models with fixed effects for time periods (coded as dummies 0–6, 0–12, and 0–18). Random effects for patients nested within unit and health trust were included if necessary, according to Akaike’s Information Criterion (AIC). As a higher score in the outcome variables indicated more severe problems, a positive difference from the start to the end of the period indicated improvement, while a negative difference indicated deterioration. We reported regression coefficients (RCs) with 95% confidence intervals (CIs) and effect sizes (Cohen’s *d*), adjusting the models for baseline patient characteristics including age, sex, duration of treatment, severity of mental illness, and level of functioning. To test if those lost to follow-up at 6, 12 and 18 months were different in primary outcomes than those remaining in the study (“drop-out analyses”), we used independent samples *t*-test on the primary outcomes for these groups six months earlier.

To analyze changes in the subscales of BASIS-24 and HoNOS (secondary outcomes) across the six-month periods for the EBP subsamples, we applied an independent samples *t*-test at each period and reported the

results as mean differences between the two groups with corresponding 95% CIs and *p*-values. As with the primary outcomes, a positive difference from the start to the end of a period indicated an improvement and a negative difference a deterioration.

To analyze associations between the subscales of BASIS-24 and HoNOS (secondary outcomes) and the clinical unit’s fidelity to the EBP model, we estimated linear mixed models with random intercepts for patients nested within unit and health trust whenever needed according to AIC. We included fixed effects for periods (coded as dummies 0–6, 6–12, and 12–18), unit group (intervention vs. control) and fidelity at the end of each period, and all two- and three-way interactions. We applied AIC to reduce the models for excessive interactions. Random effects for health trust and clinical unit level were considered, but not included as it did not enhance model fit according to AIC. All adjusted models included baseline patient characteristics (age, sex, duration of treatment, severity of mental illness, level of functioning), and patient’s exposure to the practice during the previous six months. Standard residual diagnostics were performed. The results were presented as regression coefficients and standard errors.

The inference results were presented along with *p*-values for the sake of transparency. Given the exploratory nature of the study, the reported *p*-values should be viewed as providing preliminary insights rather than definitive evidence of significance. As such, they warrant cautious interpretation and calls for further confirmation in future studies. We used SPSS for Windows version 28 for descriptive analyses and STATA version 17 for linear mixed model analyses.

Results

Patient- and clinician-reported outcomes

Primary patient- and clinician-reported outcomes are presented descriptively in Table 2. According to the results of the linear mixed model, also presented in Table 2, there was an improvement with small effect size in BASIS-24 and HoNOS scores for the total patient sample at 6 and 12 months but not at 18 months. The additional *t*-test revealed that for HoNOS, the 117 patients remaining in the study at 12 months showed more severe problems ($p < 0.001$) at 6 months than the 58 patients lost to follow-up at 12 months. However, no such differences were observed 6 months earlier between patients remaining or lost to follow-up at either 6 or 18 months. For BASIS-24, no differences were found at any time point between remaining patients and those lost to follow-up.

As shown for secondary outcomes for EBP subsamples in Table 3, patients at intervention sites for the Anti-psychotic Medication Management reported greater improvement in the BASIS-24 Symptoms subscale

Table 2 Primary outcomes (BASIS-24 and HoNOS) for the total sample at 6, 12 and 18 months. Descriptive statistics by time point and results from adjusted linear mixed models

Descriptive statistics in primary outcomes at each time point						
Time point	BASIS-24 mean total score			HoNOS mean total score		
	N	Mean (SD)		N	Mean (SD)	
Month 0	309	1.23 (0.61)		274	0.78 (0.47)	
Month 6	224	1.12 (0.65)		175	0.70 (0.47)	
Month 12	153	1.06 (0.66)		125	0.66 (0.45)	
Month 18	116	1.09 (0.66)		87	0.66 (0.44)	
Results of adjusted linear mixed model for trend in primary outcome variables						
Time point	BASIS-24 mean total score ¹			HoNOS mean total score ²		
	RC (95% CI)	p-value	Effect size: Cohen's d	RC (95% CI)	p-value	Effect size: Cohen's d ³
Intercept	0.91 (0.56; 1.25)	< 0.001		0.39 (0.13; 0.65)	0.004	
Month 0 – ref.	0	-		0		
Month 6	-0.10 (-0.18; -0.02)	0.010	0.16	-0.07 (-0.13; -0.01)	0.022	0.21
Month 12	-0.14 (-0.23; -0.05)	0.003	0.10	-0.07 (-0.13; -0.0004)	0.049	0.22
Month 18	-0.08 (-0.18; 0.02)	0.129	0.13	-0.07 (-0.14; 0.01)	0.086	0.22

The numbers are adjusted for patient characteristics at baseline (age, sex, years in treatment in mental health services, severity of mental illness, and functioning)

BASIS-24: Behavior and Symptom Identification Scale. HoNOS: Health of the Nation Outcome Scales

RC: Regression coefficient, CI: Confidence interval

¹Model with random effects for patients nested within HF

²Model with random effects for patients nested within unit

compared to those at control sites during the initial six months. Likewise, patients receiving the Illness Management and Recovery intervention showed better outcomes on the BASIS-24 Relationships subscale over the first six-month period but with a deterioration between six and twelve months. Clinician-rated HoNOS Symptoms and Functioning subscales did not show any improvement for the Antipsychotic Medication Management or the Illness Management and Recovery groups, respectively.

Association between secondary outcomes and fidelity

Table 4 shows the associations between the BASIS-24 and HoNOS subscales (secondary outcomes) and fidelity to the EBP models. Overall, fidelity was not associated with the BASIS-24 or HoNOS subscales. The only exception was the Symptoms subscale in relation to Antipsychotic Medication Management, where higher fidelity was associated with less improvement in the intervention group.

Discussion

The results seen in relation to other studies

The improvements in primary outcomes for the total sample observed in the first 12 months and the improvement of secondary outcomes for two EBP subsamples in the first 6 months align with mixed results from previous studies on similar treatments in mental health services. For example, two pragmatic trials on training providers in medication management showed improved symptoms for patients with schizophrenia after 6 to 9 months [12, 13]. Of three RCTs on family psychoeducation, only one

reported symptom improvement at the end of treatment, with another showed improvement at a 12-month follow-up [39–41]. Studies on Illness Management and Recovery also yielded inconsistent findings: two pre-post studies and one RCT found evidence of improvement [14–16], whereas three separate RCTs, all with moderate to large sample sizes, reported non-significant results [18–20].

Studies investigating associations between clinical outcomes in psychosis and fidelity to EBP models are limited. We are not aware of any such studies specifically addressing any of the four EBPs in our CRCT. However, the lack of associations between clinical outcomes and fidelity in our study partially aligns with the limited results in studies on such associations in other EBPs for severe mental disorders. For instance, one study found no significant differences in psychiatric symptoms between high- and low-fidelity programs [42]. Another study on dual disorder programs showed limited associations between fidelity and psychosis symptoms [43]. Finally, a recent study on case management with modest sample size and low variance in fidelity failed to identify a significant association between fidelity and psychiatric symptoms [44]. We have no explanation to the negative association found between the Symptom subscale and Antipsychotic Medication Management, and this may be a random finding due to factors discussed below. The association between outcome and fidelity is one of the important areas of research on EBP fidelity, and studies should measure both outcomes and fidelity [1].

Table 3 Secondary outcomes for EBP subsamples. Descriptive statistics and results of independent samples t-test of the differences between intervention and control sites

Secondary patient-reported outcomes: BASIS-24 subscales													
Outcome	Psychosis subscale for Antipsychotic Medication Management				Relationship subscale for Illness Management and Recovery				Functioning subscale for Illness Management and Recovery				
	Intervention sites		Control sites		Intervention sites		Control sites		Intervention sites		Control sites		
Period	n	Mean (SD)	N	Mean (SD)	n	Mean (SD)	n	Mean (SD)	n	Mean (SD)	n	Mean (SD)	p-value
0-6	47	0.19 (0.83)	37	0.19 (0.83)	75	0.47 (1.00)	53	0.47 (1.00)	26	0.21 (0.65)	26	0.17 (0.81)	0.001 (0.93)
6-12	32	0.08 (0.70)	28	0.08 (0.70)	50	0.37 (0.99)	38	-0.19 (1.21)	18	0.07 (1.11)	18	0.06 (0.77)	0.34 (0.88)
12-18	21	-0.41 (0.90)	22	-0.41 (0.90)	36	0.01 (0.95)	28	0.25 (1.07)	16	-0.01 (0.80)	16	-0.23 (1.17)	-0.11 (1.43)
		Mean diff. (95% CI)				Mean diff. (95% CI)				Mean diff. (95% CI)			
0-6		0.47 (0.10; 0.83)				0.47 (0.12; 0.81)				0.03 (-0.30; 0.37)			0.009
6-12		-0.29 (-0.73; 0.15)				-0.52 (-0.99; -0.06)				0.01 (-0.57; 0.59)			0.027
12-18		-0.41 (-0.98; 0.16)				0.36 (-0.26; 0.99)				0.22 (-0.39; 0.84)			0.248
Secondary clinician-reported outcomes: HoNOS subscales													
Outcome	Symptoms subscale for Antipsychotic Medication Management				Relationship subscale for Illness Management and Recovery				Functioning subscale for Illness Management and Recovery				
	Intervention sites		Control sites		Intervention sites		Control sites		Intervention sites		Control sites		p-value
Period	n	Mean (SD)	N	Mean (SD)	n	Mean (SD)	N	Mean (SD)	n	Mean (SD)	N	Mean (SD)	p-value
0-6	36	0.08 (0.64)	34	0.08 (0.64)	53	0.21 (0.65)	26	0.21 (0.65)	26	0.17 (0.81)	26	0.17 (0.81)	0.839
6-12	30	0.15 (0.91)	23	0.15 (0.91)	40	0.07 (1.11)	18	0.07 (1.11)	18	0.06 (0.77)	18	0.06 (0.77)	0.964
12-18	13	0.10 (0.72)	16	0.10 (0.72)	26	-0.01 (0.80)	16	-0.01 (0.80)	16	-0.23 (1.17)	16	-0.23 (1.17)	0.463
		Mean diff. (95% CI)				Mean diff. (95% CI)				Mean diff. (95% CI)			
0-6		0.08 (-0.20; 0.35)				0.03 (-0.30; 0.37)				0.03 (-0.30; 0.37)			0.839
6-12		0.20 (-0.26; 0.66)				0.01 (-0.57; 0.59)				0.01 (-0.57; 0.59)			0.964
12-18		-0.16 (-0.73; 0.41)				0.22 (-0.39; 0.84)				0.22 (-0.39; 0.84)			0.463

BASIS-24: Behavior and Symptom Identification Scale. HoNOS: Health of the Nation Outcome Scales

CI: Confidence interval

Table 4 Associations between subscales of secondary outcomes and fidelity, stratified by unit groups

Patient-reported outcomes for BASIS-24 subscales and evidence-based practices. Adjusted linear mixed models *						
Variables / Periods	Psychosis and Antipsychotic Medication Management (n = 183, 10–11 sites)		Relationships and Family Psychoeducation (n = 161, 9–11 sites)		Relationships and Illness Management and Recovery (n = 235, 12–16 sites)	
	RC (SE)	p	RC (SE)	p	RC (SE)	p
Intercept	-2.26 (1.21)	0.061	0.08 (1.05)	0.941	0.34 (0.61)	0.576
Period (0–6 – ref.)						
6–12	0.71 (0.22)	0.001	-0.29 (0.23)	0.215	0.16 (0.25)	0.527
12–18	0.21 (0.24)	0.384	-0.29 (0.26)	0.276	-0.03 (0.28)	0.921
Unit group (Control – ref.)						
Intervention sites	0.31 (0.21)	0.128	-0.35 (0.29)	0.226	0.22 (0.29)	0.450
Period x Group						
6–12	-0.98 (0.31)	0.002	0.42 (0.39)	0.281	-0.72 (0.32)	0.026
12–18	-0.95 (0.33)	0.004	0.19 (0.44)	0.655	-0.16 (0.36)	0.662
Fidelity	0.49 (0.30)	0.099	-0.01 (0.09)	0.880	0.07 (0.08)	0.381
Clinician-reported outcomes for HoNOS subscales and evidence-based practices. Adjusted linear mixed models*						
Variables / Periods	Symptoms and Antipsychotic Medication Management (n = 146, 6–8 sites)		Functioning and Family Psychoeducation (n = 114, 8–10 sites)		Functioning and Illness Management and Recovery (n = 154, 10–13 sites)	
	RC (SE)	p	RC (SE)	p	RC (SE)	p
Intercept	-1.58 (1.41)	0.264	-1.95 (1.11)	0.079	0.32 (0.70)	0.655
Period (0–6 – ref.)						
6–12	0.02 (0.22)	0.937	-0.26 (0.26)	0.310	-0.32 (0.28)	0.250
12–18	-0.05 (0.24)	0.834	-0.52 (0.30)	0.085	-0.06 (0.30)	0.829
Unit group (Control – ref.)						
Intervention sites	6.81 (2.56)	0.008	-0.17 (0.30)	0.558	0.10 (0.25)	0.687
Period x Group						
6–12	0.59 (0.34)	0.083	0.29 (0.39)	0.460	0.13 (0.34)	0.699
12–18	0.90 (0.51)	0.078	0.60 (0.46)	0.197	-0.27 (0.38)	0.473
Fidelity	0.40 (0.37)	0.279	0.03 (0.09)	0.717	-0.02 (0.07)	0.824
Group x Fidelity						
Intervention sites	-2.19 (0.84)	0.009				

RC: Regression coefficient, SE: Standard error. The numbers are adjusted for patient characteristics at baseline (age, sex, years in treatment in mental health services, severity of mental illness, and functioning) and exposures to EBPs that are not shown in the table

Factors potentially contributing to the lack of positive results

Several factors might have contributed to our failure to find significant positive associations between clinical outcomes and fidelity, as well as the limited results in patient- and clinician-reported outcomes.

Sample sizes and attrition. The subsamples used to analyze secondary outcomes were limited to patients from units designated as intervention or control sites for specific EBPs. Both these subsamples and the total sample diminished over time, potentially impacting the results [45]. In line with CONSORT recommendations, we examined variations in changes over time and found mostly small variations. This suggests that the limited improvement might be due to the low statistical power inherent in small samples [46]. While linear mixed models partially helped control for attrition, influence on the outcomes likely persisted [47].

Patients may have been lost to follow-up for various reasons, such as discharge or non-completion of

questionnaires. We did not collect data on the reasons, which hindered a thorough assessment of these. Furthermore, attrition effects might have varied based on differences between patients who remain in treatment and those lost to follow-up [47]. It is probable that patients who remained in mental health services for a longer duration had more severe and enduring mental health issues. This is supported by the observed differences in illness severity at 6 months between patients who remained in the study and those who were lost to follow-up at 12 months. A paper on social factors and personal recovery, showed that the patients remaining in our study at 18 months were more often women and older than those lost to follow-up, and most likely represented patients with more severe or chronic problems that had not been referred to primary care due to improvement [48].

Exposure to the EBPs. A central explanation for the lack of associations may be the limited exposure of patients to the EBP. Inclusion in the data analyses did not require that patients had received the specific

intervention implemented at their site. As a result, a considerable number of patients in the EBP subsamples had little or no exposure to the EBPs, as shown in the descriptive data in Table A in the Supplementary Material. Even though our analyses adjusted for exposure, the measurement of exposure may have been hampered by uneven reporting and underreporting by clinicians with varying knowledge of treatments provided during the last six months. This misalignment means that fidelity assessed at the service level may primarily reflect organizational implementation capacity rather than the actual treatment delivered to the included patients. The absence of a more systematic linkage between individual exposure and clinical outcomes likely diluted any fidelity–outcome associations and represents a methodological limitation of the study.

In addition to the limited and uneven exposure, the small EBP subsamples and the attrition over time further reduced the statistical power to detect associations. The combination of modest exposure, varying sample sizes, and attrition likely contributed to random variation overshadowing potential intervention effects. These issues are not unique to the present study but highlight the challenges of evaluating complex, multicomponent interventions under real-world service conditions. Among 400 implementation studies reported in a scoping review, only 15% had used a measure of penetration (exposure) [4].

Our exposure measure was a general ordinal scale trying to integrate scope, regularity and systematicity of the delivery of the EBP to make it easier for clinicians to provide information. However, exposure measures more specific to elements of each EBP are probably needed. As an example, patients' attendance to Illness Management and Recovery session are recorded, and a recent review found that attendance represented a recurring determinant for outcome across the studies [49]. As shown in the Supplementary Material Table A and known from clinical practice, there are some differences in provision of the four EBPs. The two biological EBPs are provided by a large group of mental health workers to many patients with psychosis, while the two psychosocial EBPs typically are provided by a few specially trained clinicians to a smaller, selected group of patients. Scheduled specific sessions provided by clinicians in specific roles or with specific competencies are easier to register and measure than ad hoc informal conversations between patients and several staff members providing motivation and support. Furthermore, our ordinal exposure scale, which has unknown psychometric properties, may not have provided a valid and reliable measurement of the actual patient exposure. A possible solution for future studies may be specific systematic information recorded regularly in electronic patient records that may be extracted

as clinical feedback and for research without burden for clinicians.

Inclusion process and inclusion period. The inclusion process may have influenced clinical outcomes because we included both patients in ongoing treatment and newly referred patients. This approach differs from studies tracking cohorts from treatment initiation to completion. Many were recruited during early treatment phases, where treatment effects on psychiatric issues are often more pronounced, although we lack supporting data to confirm this effect [50]. A study on proportion of patients with psychosis who was willing to take part in research found that the patients who were approached had better clinical outcomes than those who were not, suggesting that clinicians were more likely to approach patients who were less ill [51]. This might be the same in our CRCT.

All patients were followed up every six months for 18 months, provided they were not discharged and remained willing to participate. However, the nine months inclusion period introduced variability in the timing of patient outcome data relative to fidelity data. This meant that outcome data for patients recruited late in the period were analyzed against a fidelity assessment that had been completed several months earlier. With variations in the clinical course of different patients, it is difficult to know how this might have influenced the possibility to detect positive associations between outcomes and fidelity.

Informants and outcome measures. Differences between informants and characteristics of the outcome measures may have influenced the results. Combining patient- and clinician-reported outcomes is standard because their roles, perspectives, and access to information may describe different and supplementary aspects of the patient's condition. Consequently, discrepancies in reported results may be expected. A recent study found that the correlation pattern between patient- and clinician-reported measures showed a three-layer structure, representing a continuum from inner experience (e.g., subjective quality of life, only known to the patient) to the external presentation of experiences (e.g., symptom severity, known to both patient and clinician) [52]. The lack of changes in secondary clinician-reported outcomes might also be partially due to a lower sensitivity to change of the HoNOS subscales, which have fewer items, providing lower resolution than the BASIS-24 subscales [53, 54]. We could not find any other study using both BASIS-24 and HoNOS as outcome measures. In our study, the correlation between the two total scales was moderately strong, while correlations between their subscales ranged from weak to strong [55].

The low variance in fidelity likely influenced the results by limiting the possibility of detecting a positive association with outcomes. This limitation persisted even

during the initial 6–12 months, when the changes in both outcome and fidelity were greatest. Furthermore, even for Illness Management and Recovery, which demonstrated a large effect size for the increase in fidelity at intervention sites compared to control sites, we found no association between outcome and fidelity. This finding is likely attributable to the low exposure of patients to the EBP, as discussed above. We cannot rule out that outcome might have been associated with subgroups of fidelity scale items with larger variance than the total scale and more closely related to the selected secondary outcomes.

The complex design of the CRCT, where each site served as both an intervention site for one EBP and a control site for another, was chosen to facilitate the inclusion of four EBPs in the CRCT with EBP fidelity as the primary outcome. As noted previously, comparing four distinct EBPs within the same study revealed differences among them that impact implementation and fidelity. However, for this current secondary exploratory study, this complex design may have reduced the ability to detect positive associations between outcomes and fidelity. Specifically, the fact that all sites were intervention sites for one EBP might have led to greater overall engagement among both clinicians and patients than if some units had been purely passive control sites. Although three of the EBPs demonstrated a significantly higher increase in fidelity at intervention sites compared to control sites, these differences did not translate into an association with outcome. Additionally, we did not account for other treatments provided to patients, which may have had a more substantial impact on clinical outcomes than the designated EBPs.

Strengths and limitations

The study's strengths include the use of data from a CRCT, and the application of well-established outcome measures with adequate psychometric properties. The EBPs selected were core practices recommended in national guidelines for treating individuals with psychosis. Furthermore, the fidelity of these practices was assessed by independent raters using specific fidelity scales, which also demonstrated adequate psychometric properties. Finally, data analysis utilized linear mixed models, as recommended for longitudinal studies to limit the impact of participant attrition.

However, several limitations were present. A substantial proportion of patients was lost to follow-up between measurement points, and this attrition may have influenced the results. Patient recruitment and data collection from both patients and clinicians were carried out by clinicians and health personnel as part of their busy clinical duties, and with unknown inter-rater reliability. Furthermore, the EBP subsamples were small. The measure used for exposure to EBPs had unknown measurement

properties. Clinicians' assessments regarding patients' exposure to each EBP required modifications after the data collection at six months. Patients were included during various illness phases, and variations existed in the timing of outcome data collection relative to fidelity assessments. The design of the parent CRCT was less suitable for the needs of the current secondary exploratory study. Finally, data was not collected on the reasons for patients being lost to follow-up.

Conclusions and implications

There was improvement in both primary outcomes for the total sample over the first 12 months, along with improvement in two patient-reported secondary outcomes for EBP subsamples over the first 6 months. Crucially, we found no association between outcomes and fidelity for any of the EBP subsamples. Our failure to find associations between outcomes and fidelity might be attributed to several factors, including sample size, attrition, trial design, variance in variables and measurement properties, as well as interactions between these factors.

These findings emphasize the importance of aligning fidelity assessments with actual patient-level exposure when evaluating implementation outcomes. Future studies investigating the association between outcomes and EBP fidelity should involve large samples, utilize measurement scales with strong psychometric properties, track patient cohorts from the beginning of treatment, and ensure a close correspondence between the delivery of EBPs and the measurement of both fidelity and clinical change.

Abbreviations

AIC	Akaike's Information Criterion
BASIS-24	Behavior and Symptom Identification Scale
CAMHS	Child and adolescent mental health services
CGI	Clinical Global Impression scale
CI	Confidence interval
CMHC	Community mental health center
CONSORT	Consolidated Standards of Reporting Trials
CRCT	Cluster-randomized controlled trial
EBP	Evidence-based practice
HoNOS	Health of the Nation Outcome Scales
PSF	Practical and Social Functioning Scale
RC	Regression coefficient
RCT	Randomized controlled trial

Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s12888-025-07566-w>.

Supplementary Material 1

Acknowledgements

We are grateful to all patients and clinicians participating in the study providing data by completing questionnaires and assessments, to fidelity assessors providing data on fidelity, and to personnel registering data in the electronic database.

Author contributions

TR was principal investigator, designed the study together with the co-authors, planned the article, and drafted and revised the manuscript. JSB did the data analyses and drafted parts on data analyses and results. EB, MH, KD, KH, TSH, VØH, IJ, JOJ, KJJ, BS, EWH and HC participated in designing the study, organized data collections locally, and participated in revising the manuscript. GRB participated in designing the study and in revising the manuscript. All authors approved the final manuscript.

Funding

Open access funding provided by Akershus University Hospital (AHUS). The study was funded by the South-Eastern Regional Health Authority in Norway (Grant No. 2015106) and by the participating health trusts through the work of researchers, clinicians and other personnel contributing in various ways.

Data availability

The dataset used and analyzed during the current study is available from the corresponding author on reasonable request.

Declarations

Ethics approvals

Only patients giving informed written consent were included in the study. The Regional Committee for Medical and Health Research Ethics in Southeastern Norway (reg.no. REK 2015/2169) and the Data Protection Officer for each health trust approved the study, which followed the principles in the Declaration of Helsinki.

Consent for publication

Not applicable.

Competing interests

The authors declare no competing interests.

Author details

¹Division of Mental Health Services, Akershus University Hospital, Lørenskog, Norway

²Institute of Clinical Medicine, University of Oslo, Oslo, Norway

³Health Services Research Unit, Akershus University Hospital, Lørenskog, Norway

⁴Section of Research and Innovation, Helse Fonna Health Trust, Haugesund, Norway

⁵Department of Health and Caring Sciences, Western Norway University of Applied Sciences, Stord, Norway

⁶Department of Global Public Health and Primary Care, University of Bergen, Bergen, Norway

⁷Division of Mental Health, Sørlandet Hospital, Kristiansand, Norway

⁸Centre for Medical Ethics, Institute for Health and Society, University of Oslo, Oslo, Norway

⁹University Hospital of Northern Norway, Tromsø, Norway

¹⁰Clinical Neuroscience Research Group, Department of Psychology, University of Oslo, Oslo, Norway

¹¹TIPS Centre for Clinical Research in Psychosis, Psychiatric Division, Stavanger University Hospital, Stavanger, Norway

¹²Faculty of Health, Department of Public Health, University of Stavanger, Stavanger, Norway

¹³Mental Health Norway, Oslo, Norway

¹⁴Division of Mental Health, Innlandet Hospital Trust, Brumunddal, Norway

¹⁵Department of Psychology, Indiana University Indianapolis, Indianapolis, IN, USA

Received: 29 June 2025 / Accepted: 27 October 2025

Published online: 28 November 2025

References

- Bond GR, Drake RE. Assessing the fidelity of Evidence-Based practices: history and current status of a standardized measurement methodology. *Adm Policy Ment Health*. 2020;47(6):874–84.

- Bond GR, Evans L, Salyers MP, Williams J, Kim H-W. Measurement of fidelity in psychiatric rehabilitation. *Ment Health Serv Res*. 2000;2(2).
- Proctor E, Silmere H, Raghavan R, Howmand P, Aarons G, Bunger A, et al. Outcomes for implementation research: conceptual distinctions, measurement challenges, and research agenda. *Adm Policy Ment Health*. 2011;38(2):65–76.
- Proctor EK, Bunger AC, Lengnick-Hall R, Gerke DR, Martin JK, Phillips RJ, et al. Ten years of implementation outcomes research: a scoping review. *Implement Sci*. 2023;18(31):1–19.
- Barlatti S, Nibbio G, Vita A. Evidence-based psychosocial interventions in schizophrenia: a critical review. *Curr Opin Psychiatry*. 2024;37(3):131–9.
- Keepers GA, Fochtmann LJ, Anzia JM, Benjamin S, Lyness JM, Mojtabai R, et al. The American psychiatric association practice guideline for the treatment of patients with schizophrenia. *Am J Psychiatry*. 2020;177:868–72.
- McDonagh MS, Dana T, Kopelovich SL, Monroe-DeVita M, Blazina I, Bougatso C et al. Psychosocial interventions for adults with schizophrenia: an overview and update of systematic reviews. *Psychiatr Serv*. (Washington, DC). 2022;73(3):299–312.
- Ruud T, Drake RE, Salyte Benth J, Drivenes K, Hartveit M, Heiervang K, et al. The effect of intensive implementation support on fidelity for four Evidence-Based psychosis treatments: A cluster randomized trial. *Adm Policy Ment Health*. 2021;48(5):909–20.
- Gallardo-Gomez D, Noetel M, Alvarez-Barbosa F, Alfonso-Rosa RM, Ramos-Munell J, Del Pozo Cruz B, et al. Exercise to treat psychopathology and other clinical outcomes in schizophrenia: A systematic review and meta-analysis. *Eur Psychiatry*. 2023;66(1):e40.
- Sabe M, Kaiser S, Sentissi O. Physical exercise for negative symptoms of schizophrenia: systematic review of randomized controlled trials and meta-analysis. *Gen Hosp Psychiatry*. 2020;62:13–20.
- Ziebart C, Bobos P, MacDermid JC, Furtado R, Sobczak DJ, Doering M. The efficacy and safety of exercise and physical activity on psychosis: A systematic review and meta-analysis. *Front Psychiatry*. 2022;13:807140.
- Gray R, Wykes T, Edmonds M, Leese M, Gournay K. Effect of a medication management training package for nurses on clinical outcomes for patients with schizophrenia: cluster randomised controlled trial. *Br J Psychiatry: J Mental Sci*. 2004;185:157–62.
- Harris N, Lovell K, Day J, Roberts C. An evaluation of a medication management training programme for community mental health professionals; service user level outcomes: a cluster randomised controlled trial. *Int J Nurs Stud*. 2009;46(5):645–52.
- Egeland KM, Ruud T, Ogden T, Fardig R, Lindstrom JC, Heiervang KS. How to implement illness management and recovery (IMR) in mental health service settings: evaluation of the implementation strategy. *Int J Ment Health Syst*. 2017;11:13.
- Levitt AJ, Mueser KT, DeGenova J, Lorenzo J, Bradford-Watt D, Barbosa A, et al. Randomized controlled trial of illness management and recovery in Multiple-Unit supportive housing. *Psychiatric Serv*. 2009;60(12):1629–36.
- Mueser KT, Meyer PS, Penn DL, Clancy R, Clancy DM, Salyers MP. The illness management and recovery program: rationale, development, and preliminary findings. *Schizophr Bull*. 2006;32(Suppl 1 Suppl 1):S32–43.
- Jensen SB, Dalum HS, Korsbek L, Hjorthoj C, Mikkelsen JH, Thomsen K, et al. Illness management and recovery: one-year follow-up of a randomized controlled trial in Danish community mental health centers: long-term effects on clinical and personal recovery. *BMC Psychiatry*. 2019;19(1):65.
- Monroe-DeVita M, Morse G, Mueser KT, McHugo GJ, Xie H, Hallgren KA et al. Implementing illness management and recovery within assertive community treatment: a pilot trial of feasibility and effectiveness. *Psychiatr Serv*. (Washington, DC). 2018;69(5):562–71.
- Roosenschoon BJ, van Weeghel J, Deen ML, van Esveld EW, Kamperman AM, Mulder CL. Effects of illness management and recovery: A multicenter randomized controlled trial. *Front Psychiatry*. 2021;12:723435.
- Salyers MP, McGuire AB, Kukla M, Fukui S, Lysaker PH, Mueser KT. A randomized controlled trial of illness management and recovery with an active control group. *Psychiatric Serv* (Washington DC). 2014;65(8):1005–11.
- Campbell MK, Piaggio G, Elbourne DR, Altman DG, Group C. Consort 2010 statement: extension to cluster randomised trials. *BMJ*. 2012;345:e5661.
- OECD. OECD Reviews of Health Care Quality: Norway 2014: Raising Standards. 2014.
- Ringard Å, Sagan A, Perre Saunes I, Lindahl AK. Norway. Health system review. Copenhagen: European Observatory on Health Systems and Policies; 2013. p. 8. Contract No.
- Ruud T, Friis S. Community-based mental health services in Norway. *Consortium Psychiatricum*. 2021;2(1):47–54.

25. Helsedirektoratet. Nasjonal faglig retningslinje for utredning, behandling og oppfølging av personer med Psykoselidelser [National guidelines for assessment, treatment and follow-up of people with psychosis]. Helsedirektoratet [Norwegian Directorate of Health] Oslo; 2013. Report No.: IS-1957.
26. Egeland KM, Heiervang KS, Landers M, Ruud T, Drake RE, Bond GR. Psychometric properties of a fidelity scale for illness management and recovery. *Adm Policy Ment Health*. 2020;47:885–93.
27. Joa I, Johannessen JO, Heiervang KS, Sviland AA, Nordin HA, Landers M, et al. The family psychoeducation fidelity scale: psychometric properties. *Adm Policy Ment Health*. 2020;47:894–900.
28. Ruud T, Drivenes K, Drake RE, Haaland VO, Landers M, Stensrud B, et al. The antipsychotic medication management fidelity scale: psychometric properties. *Adm Policy Ment Health*. 2020a;47:911–9.
29. Ruud T, Hoifodt TS, Hendrick DC, Drake RE, Høye A, Landers M, et al. The physical health care fidelity scale: psychometric properties. *Adm Policy Ment Health*. 2020b;47:901–10.
30. Hartveit M, Hovlid E, Nordin MHA, Ovretveit J, Bond GR, Biringner E, et al. Measuring implementation: development of the implementation process assessment tool (IPAT). *BMC Health Serv Res*. 2019;19(1):721.
31. Cameron IM, Cunningham L, Crawford JR, Eagles JM, Eisen SV, Lawton K, et al. Psychometric properties of the BASIS-24 (Behaviour and symptom identification Scale-Revised) mental health outcome measure. *Int J Psychiatry Clin Pract*. 2007;11(1):36–43.
32. Eisen SV, Gerena M, Ranganathan G, Esch D, Idiculla T. Reliability and validity of the BASIS-24© mental health survey for Whites, African-Americans, and Latinos. *J Behav Health Serv Res*. 2006;33(3):304–23.
33. Eisen SV, Normand SL, Belanger AJ, Spiro A, Esch D. The revised behavior and symptom identification scale (BASIS-R): reliability and validity. *Med Care*. 2004;42(12):1230–41.
34. Niv N, Cohen AN, Mintz J, Ventura J, Young AS. The validity of using patient self-report to assess psychotic symptoms in schizophrenia. *Schizophr Res*. 2007;90(1–3):245–50.
35. McLean_Hospital. Behavior and Symptom Identification Scale. BASIS-24 Instruction Manual. Belmont MA, USA; 2006.
36. Wing JK, Beevor AS, Curtis RH, Park SB, Hadden S, Burns A. Health of the Nation outcome scales (HoNOS). Research and development. *Br J Psychiatry: J Mental Sci*. 1998;172:11–8.
37. Haro JM, Kamath SA, Ochoa S, Novick D, Rele K, Fargas A et al. The clinical global Impression-Schizophrenia scale: a simple instrument to measure the diversity of symptoms present in schizophrenia. *Acta Psychiatr Scand Suppl*. 2003(416):16–23.
38. Clausen H, Ruud T. The practical and social functioning (PSF) scale: development and measurement properties of an instrument for assessing activity and social participation among people with serious mental illness. *BMC Psychiatry*. 2024;24(1):693.
39. Garety PA, Fowler DG, Freeman D, Bebbington P, Dunn G, Kuipers E. Cognitive-behavioural therapy and family intervention for relapse prevention and symptom reduction in psychosis: randomised controlled trial. *Br J Psychiatry: J Mental Sci*. 2008;192(6):412–23.
40. Merinder L-B, Viuff AG, Laugesen HD, Clemmensen K, Misfelt S, Espensen B. Patient and relative education in community psychiatry: a randomized controlled trial regarding its effectiveness. *Soc Psychiatry Psychiatr Epidemiol*. 1999;34:287–94.
41. Bradley GM, Couchman GM, Perlesz A, Nguyen AT, Singh B, Riess C. Multiple-Family group treatment for English- and Vietnamese-Speaking families living with schizophrenia. *Psychiatric Serv*. 2006;57(4):521–30.
42. McHugo GJ, Drake RE, Teague GB, Xie H. Fidelity to assertive community treatment and client outcomes in the new Hampshire dual disorders study. *Psychiatric Serv*. 1999;50(6):818–24.
43. Chandler DW. Fidelity and outcomes in six integrated dual disorders treatment programs. *Community Ment Health J*. 2011;47(1):82–9.
44. Latimer E, Aubry T, Barrett B, Bergeron-Leclerc C, Briand C, Durbin J, et al. Association between fidelity to the strengths model of case management and client outcomes: A quasi-experimental study. *Psychiatr Rehabil J*. 2022;45(3):280–90.
45. Dumville JC, Torgerson DJ, Hewitt CE. Reporting attrition in randomised controlled trials. *BMJ*. 2006;332:969–71.
46. Moher D, Hopewell S, Schulz KF, Montori V, Gotzsche PC, Devereaux PJ, et al. CONSORT 2010 explanation and elaboration: updated guidelines for reporting parallel group randomised trials. *J Clin Epidemiol*. 2010;63(8):e1–37.
47. Bell ML, Kenward MG, Fairclough DL, Horton NJ. Differential dropout and bias in randomised controlled trials: when it matters and when it May not. *BMJ*. 2013;346(e8668):1–7.
48. Linde J, Schmid MT, Ruud T, Skar-Froding R, Biringner E. Social factors and recovery: A longitudinal study of patients with psychosis in mental health services. *Community Ment Health J*. 2023;59(2):294–305.
49. Goh YSS, Ow Yong JQY, Li AZ. Effectiveness of illness management and recovery program on people with severe mental illnesses: a systematic review and meta-analysis. *Front Psychiatry*. 2023;14:1162288.
50. de Winter L, Vermeulen JM, Couwenbergh C, van Weeghel J, Hasson-Ohayon I, Mulder CL, et al. Short- and long-term changes in symptom dimensions among patients with schizophrenia spectrum disorders and different durations of illness: A meta-analysis. *J Psychiatr Res*. 2023;164:416–39.
51. Patel R, Oduola S, Callard F, Wykes T, Broadbent M, Stewart R, et al. What proportion of patients with psychosis is willing to take part in research? A mental health electronic case register analysis. *BMJ Open*. 2017;7(3):e013113.
52. Mendlovic S, Roe D, Markusfeld G, Mainz J, Kristensen S, Goldzweig G. Exploring the relation between clinician ratings and patient-reported experience and outcomes. *Int J Qual Health Care: J Int Soc Qual Health Care / ISQua*. 2022;34(Supplement1):ii98–104.
53. Cameron IM, Cunningham L, Crawford JR, Eagles JM, Eisen SV, Lawton K, et al. Psychometric properties of the BASIS-24© (Behaviour and symptom identification Scale-Revised) mental health outcome measure. *Int J Psychiatry Clin Pract*. 2007;11(1):36–43.
54. Pirkis JE, Burgess PM, Kirk PK, Dodson S, Coombs TJ, Williamson MK. A review of the psychometric properties of the health of the Nation outcome scales (HoNOS) family of measures. *Health Qual Life Outcomes*. 2005;3:76.
55. Skar-Froding R, Clausen H, Saltyte Benth J, Ruud T, Slade M. Associations between personal recovery and service user-rated versus clinician-rated clinical recovery, a cross-sectional study. *BMC Psychiatry*. 2022;22(1):42.

Publisher's note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.