Journal of Psychiatric Research 164 (2023) 416–439

Contents lists available at ScienceDirect

Journal of Psychiatric Research

journal homepage: www.elsevier.com/locate/jpsychires

Short- and long-term changes in symptom dimensions among patients with schizophrenia spectrum disorders and different durations of illness: A meta-analysis

Lars de Winter [a,b,*], Jentien M. Vermeulen [b], Chrisje Couwenbergh [a], Jaap van Weeghel [a,c], Ilanit Hasson-Ohayon [d], Cornelis L. Mulder [e,i], Nynke Boonstra [f,g], Wim Veling [h], Lieuwe de Haan [b]

[a] Phrenos Center of Expertise, Utrecht, the Netherlands
[b] Department of Psychiatry, Amsterdam UMC Location AMC, Amsterdam, the Netherlands
[c] Tranzo, Tilburg University, Tilburg, the Netherlands
[d] Department of Psychology, Bar-Ilan University, Ramat-Gan, Israel
[e] Epidemiological and Social Psychiatric Research Institute, Erasmus MC, Rotterdam, the Netherlands
[f] NHL Stenden University of Applied Science, Leeuwarden, the Netherlands
[g] Utrecht University Medical Center, Utrecht, the Netherlands
[h] University of Groningen, University Medical Center Groningen, Groningen, the Netherlands
[i] Parnassia Psychiatric Institute, the Netherlands

ARTICLE INFO

Keywords:
Schizophrenia spectrum disorders
Symptoms
Longitudinal
Meta-analysis
Course
Psychosis
Epidemiology

ABSTRACT

In schizophrenia spectrum disorders, improvement in symptoms varies between patients with short and long durations of illness. In this meta-analysis we provided an overview of both short- and long-term symptomatic improvement for patients with schizophrenia spectrum disorders with distinct durations of illness. We included 82 longitudinal studies assessing the course of positive, negative, depressive and disorganization symptoms. We analyzed effect sizes of change in four subgroups based on durations of illness at baseline: <2 years, 2–5 years, 5–10 years, >10 years. Potential moderators were explored using meta-regression and sensitivity analyses. Overall, we found large improvements of positive symptoms and small improvements of negative, depressive, and disorganization symptoms. Positive and disorganization symptoms improved relatively stronger for patients earlier in the course of illness, whereas negative and depressive symptoms showed modest improvement regardless of duration of illness. Improvement of symptoms was associated with higher baseline severity of positive symptoms, a younger age, a smaller subsample with schizophrenia, and, specifically for negative symptoms, higher baseline severity of depressive symptoms. Future research should focus on exploring ways to optimize improvement in negative and depressive symptoms for patients with schizophrenia spectrum disorders.

1. Introduction

Schizophrenia spectrum disorders, have a lifetime prevalence of approximately 3% (Sullivan et al., 2020) and are characterized by a diverse set of symptoms, including distortions of thinking and perception, cognitive impairments, motor abnormalities, avolition and apathy, difficulties in communication, and restricted affective expression (Tandon et al., 2009). It includes the following diagnoses: schizophrenia, schizophreniform disorder, schizoaffective disorder, delusional disorder, brief psychotic disorder, psychotic disorder due to another medical condition, substance or medication-induced psychotic disorder, and psychotic disorder not otherwise specified (American Psychiatric Association, 2013). The clinical features can be roughly divided into six symptom dimensions: positive symptoms, disorganization symptoms, negative symptoms, depressive symptoms, motor symptoms and cognitive symptoms (Tandon et al., 2009). The severity and course of these symptom dimensions vary substantially between patients (Andreasen, 1995; Schnack, 2019).

Previous studies have tried to capture the clinical heterogeneity of schizophrenia spectrum disorders in staging models, describing

* Corresponding author. Phrenos Center of Expertise, Utrecht, the Netherlands.
E-mail address: Lwinter@kcphrenos.nl (L. de Winter).

https://doi.org/10.1016/j.jpsychires.2023.06.031
Received 6 December 2022; Received in revised form 29 May 2023; Accepted 23 June 2023
Available online 4 July 2023

categories of illness severity ranging from subthreshold transient early symptoms to chronic unremitting symptoms with functional and cognitive disabilities (Lieberman et al., 2001; McGorry et al., 2010). However, these models remained largely heuristic as, due to the clinical diversity within each stage, it remains complex to grasp the development and course of schizophrenia spectrum disorders. Changes in psychotic symptom dimensions over the course of a schizophrenia spectrum disorder depend on a wide variety of clinical, social, and personal characteristics of patients with schizophrenia spectrum disorders (Díaz et al., 2013; Lally et al., 2017; Lambert et al., 2008, 2010). In addition, changes in any symptom dimension depend on the duration of schizophrenia spectrum disorder and how long these patients are followed as symptoms seem to improve more substantially for persons with a short duration of illness and a short length of follow-up (Lally et al., 2017; Boonstra et al., 2012; Salazar de Pablo et al., 2021; Häfner, 2019; McGinty et al., 2018). The degree of improvement of symptoms also seem to differ between symptom dimensions. Previous studies indicated that severity of positive, depressive, and disorganization symptoms decrease over time, whereas negative symptoms, cognitive symptoms and motor symptoms remain relatively stable over the course of a schizophrenia spectrum disorder (Emsley et al., 2006; Lefebvre et al., 2020; Rowe et al., 2015; Salazar de Pablo et al., 2021).

In order to grasp all above mentioned aspects to assess the clinical diversity of schizophrenia spectrum disorders, we aimed to evaluate and quantify the course of different symptom dimensions in schizophrenia spectrum disorders over time, while controlling for differences in patient characteristics, the duration of illness of patients (i.e. the time since the first diagnosis of schizophrenia spectrum disorder), and the length of follow-up. Therefore, we conducted a meta-analysis of longitudinal studies investigating changes in symptom dimensions for patients with schizophrenia spectrum disorders by comparing subgroups of studies that evaluated patients with different durations of illness at baseline and different lengths of follow-up. Furthermore, we investigated which factors moderated these changes in different symptom dimensions within each subgroup of studies. Although previous meta-analyses already investigated the influence of different factors on the course of psychotic symptoms (Lally et al., 2017; Salazar de Pablo et al., 2021), this is the first meta-analysis that evaluated changes and moderators of change in symptoms over time, while taking duration of illness and length of follow up into account. We aimed to answer the following questions: 1) To what extent do different symptom dimensions change over the course of schizophrenia spectrum disorders? 2) Which moderators at baseline are associated with changes in symptom dimensions over time?

2. Methods

The meta-analysis followed PRISMA guidelines (Moher et al., 2009). Our protocol was preregistered in PROSPERO (CRD42020192015).

2.1. Search strategy

Records were identified through searches in PubMed, PsycInfo, CINAHL, and Cochrane of peer-reviewed journals until June 2020. The search was based on terms related to schizophrenia spectrum disorders, chronicity, course, recovery, and remission (see Supplementary Material 1). Additional references were traced through reference lists of identified studies and systematic reviews.

2.2. Eligibility criteria

Three assessors (LdW, KK & RM) independently executed study selection. Disagreements were resolved by consensus. The included studies meet the following criteria:

1) **Patient population:** Studies including adults (age ≥18) who are all diagnosed with a DSM or ICD diagnosis falling under what is currently indicated as schizophrenia spectrum or other psychotic disorders (American Psychiatric Association, 2013) were included. Studies including children or adolescents with a mean age lower than 18 years old, and studies in which a part of the study sample was not diagnosed with schizophrenia spectrum or other psychotic disorders were excluded from the meta-analysis.
2) **Study design:** Longitudinal cohort study or randomized controlled trial, with a follow-up larger than or equal to 1 year, assuring long-term follow-up evaluations, were included. Other study designs and longitudinal studies with a follow-up of less than 1 year were excluded from the meta-analysis.
3) **Outcomes:** Studies reporting uncorrected quantitative assessments of any symptoms for at least two time points were included. Qualitative studies were excluded. Also studies that did not report symptomatic outcomes, and studies that only reported data which could not be calculated into effect sizes of change were excluded from the meta-analysis.
4) **Publication:** Studies published in English in peer-reviewed journals were included. Studies that are published in another language or in other sources were excluded from the meta-analysis.

2.3. Outcome domains

After study selection, we categorized each study outcome based on the six symptom dimensions, as suggested in a previous study (Tandon et al., 2009): positive symptoms, disorganization symptoms, negative symptoms, depressive symptoms, motor symptoms and cognitive symptoms. We were not able to report on outcomes of motor symptoms or cognitive symptoms, as these outcomes were reported by less than 10 studies, the minimum number for reporting reliable outcomes (Borenstein and Higgins, 2013). This led to assessment of the following four outcome domains in this meta-analysis: 1) positive symptoms; 2) negative symptoms; 3) disorganization symptoms; and 4) depressive symptoms. A detailed overview of which outcomes are categorized in which outcome domain has been described in Supplementary Material 2.

2.4. Assessment of duration of illness and follow-up subgroups

Included studies investigated patients with different durations of illness (DOI) at baseline, and assessed outcomes over different follow-up periods. Therefore, we categorized included studies according to the patients' average DOI at baseline into five subgroups: 1) DOI<2 years; 2) DOI between 2 and 5 years; 3) DOI between 5 and 10 years; 4) DOI >10 years; 5) DOI unknown. Within each baseline DOI subgroup we also divided the outcomes of the included studies into separate subgroups based on their follow-up length: 1) follow-up <2 years; 2) follow-up between 2 and 5 years; 3) follow-up between 5 and 8 years; 4) follow-up >8 years. This method of categorization, as shown in Text Box 1, was based on categorizations described in previous studies (Breitborde et al., 2009; Frascarelli et al., 2015; Preston, 2000), and, in case of follow-up length, the potential availability of study data for that subgroup. This procedure was comparable to the method we used in a previous meta-analysis on social functioning in psychosis (De Winter et al., 2021).

This overview shows that combinations of illness duration at baseline and length of follow-up (i.e., the duration of illness at follow-up) can overlap between subgroups after the follow-up assessment. Despite this overlap, we considered clustering studies in these separate DOI and follow-up subgroups as the most optimal classification for current study.

2.5. Selection and assessment of moderators of outcome

We selected potential moderators at baseline for the symptomatic outcomes through a two-step approach. First, we extracted 72 variables

Textbox 1
Assessment of subgroups based on duration of illness at baseline and follow-up length.

Duration of illness at baseline	Length of follow-up	Duration of illness at follow-up
1. Duration of illness <2 years	1. Follow-up < 2 years	1.1 Duration of illness 1–4 years
	2. Follow-up 2–5 years	1.2 Duration of illness 2–7 years
	3. Follow-up 5–8 years	1.3 Duration of illness 5–10 years
	4. Follow-up >8 years	1.4 Duration of illness >8 years
2. Duration of illness 2–5 years	1. Follow-up < 2 years	2.1 Duration of illness 3–7 years
	2. Follow-up 2–5 years	2.2 Duration of illness 4–10 years
	3. Follow-up 5–8 years	2.3 Duration of illness 7–13 years
	4. Follow-up >8 years	2.4 Duration of illness >10 years
3. Duration of illness 5–10 years	1. Follow-up < 2 years	3.1 Duration of illness 6–12 years
	2. Follow-up 2–5 years	3.2 Duration of illness 7–15 years
	3. Follow-up 5–8 years	3.3 Duration of illness 10–18 years
	4. Follow-up >8 years	3.4 Duration of illness >13 years
4. Duration of illness >10 years	1. Follow-up < 2 years	4.1 Duration of illness >11 years
	2. Follow-up 2–5 years	4.2 Duration of illness >12 years
	3. Follow-up 5–8 years	4.3 Duration of illness >15 years
	4. Follow-up >8 years	4.4 Duration of illness >18 years

that have been found to influence outcomes in at least one of the included studies or in comparable reviews (Lally et al., 2017; Salazar de Pablo et al., 2021; McGinty et al., 2018). Second, we extracted baseline data from our included studies of each of the 72 primarily included potential moderator. If baseline data of the moderator could be extracted from at least 10 studies (Borenstein and Higgins, 2013), we analyzed the influence of the specific moderator at baseline on the study outcomes. Based on these criteria, we were able to select 26 potential moderators at baseline (see Supplementary Materials 4). For the outcome domains of disorganization and depressive symptoms not all 26 moderators met inclusion criteria.

For moderators that were evaluated by different assessment instruments (e.g. assessment of symptoms or cognition) we calculated percentile scores based on normative data to ensure that each assessment was assessed in the same scale range.

2.6. Quality assessment

Quality assessment was conducted using the Quality in Prognostic Studies (QUIPS) tool (Hayden et al., 2013). It was based on six criteria: participation, attrition, prognostic factor measurement, handling confounders, outcome measurement, and analysis and reporting. For each criterion we assigned a high, moderate or low risk of bias score for each study.

The first author (LdW) assessed all studies and another assessor (MO) independently conducted quality assessment of 10% of the studies. The level of agreement was fair to good ($\kappa = 0.61$). Disagreements were resolved by consensus. We investigated the influence of study quality on outcomes by sensitivity analysis.

2.7. Statistical analysis

2.7.1. Meta-analytic procedure

Meta-analyses were conducted using RevMan 5.3 (The Nordic Cochrane Centre, 2014). We calculated effect sizes of change in study outcomes between baseline and follow-up assessment. For studies with multiple follow-up assessments we calculated separate effect sizes of change between baseline and each follow-up assessment. As a result, outcomes from the same study could be presented in multiple subgroups based on different follow-up length. For clinical trials we analyzed the longitudinal outcomes of the total study sample of both treatment and control group together. Overall effect sizes of categorical outcomes were converted into Cohen's d (Chinn, 2000) to analyze homogeneous and consistent patterns for both continuous and categorical outcomes. Magnitude of effect was considered marginal and clinically not relevant when $d < 0.2$, small when $d \geq 0.2$ and < 0.5, medium when $d \geq 0.5$ and < 0.8, and large when $d \geq 0.8$ (Chinn, 2000). All outcomes were reported with 95% confidence intervals (CIs). We used random effects models, weighted by the method of inverse variance (Higgins, 2008). Statistical heterogeneity was assessed by calculating the I^2 statistic (including 95% CI), describing the percentage of observed heterogeneity not expected by chance (Higgins and Thompson, 2002).

2.7.2. Subgroup analyses and calculation of moderators

We investigated the influence of potential moderators on the five outcome domains through a meta-regression analysis using R (R Core Team, 2016). We analyzed differences in effect sizes of change between subgroups of studies with different baseline DOI and length of follow-up as well as differences between studies with high levels or presence versus studies with low levels or absence of any significant moderator from the meta-regression analysis, using an analysis of subgroup differences (Borenstein and Higgins, 2013). We controlled for multiple testing effects in all analyses through a Benjamini-Hochberg correction, with the false discovery rate set on 0.3 (Benjamini and Hochberg, 1995).

2.7.3. Handling outliers and publication bias

We defined outliers as individual study outcomes which confidence interval (CI) of the effect size exceeded the upper or lower bound of the CI of the effect size of overall study outcomes. A correction on potential influence of outliers on the overall study outcomes was executed by comparing subgroups of study outcomes in which outliers are still included in the analysis with subgroups in which outliers are excluded through an analysis of subgroup differences (Borenstein and Higgins, 2013). Potential publication bias was detected by visual inspection of funnel plots.

3. Results

3.1. Study flow

Of the 8,483 records retrieved through database search and reference tracking, we excluded 7,763 records after title and abstract screening. Of

the remaining 720 records, we excluded 616 records after full-text screening (see Fig. 1 for study flow and reasons of exclusion). The remaining 104 articles reported results of 82 studies.

3.2. Study characteristics

We selected 82 studies describing the course of symptomatic outcomes of 14,936 participants with a schizophrenia spectrum disorder. Regarding patient population, the mean age of participants was 37.6 years (*SD* = 14.5), and 36.2% were female. Forty-two studies (51.2%) exclusively included patients diagnosed with schizophrenia. For study design, fifteen studies (18.3%) were clinical trials and 67 studies (81.7%) were cohort studies. Regarding treatment, in 20 studies (24.4%; 13 cohort studies and 7 clinical trials) a subgroup of participants received integrated community treatment, in 14 studies (17.1%; 10 cohort studies and 4 clinical trials) a subgroup of participants received psychotherapy and in 54 studies (65.9%; 47 cohort studies and 7 clinical trials) at least 80% of the study sample received antipsychotics. In 28 studies (34.1%) the drop-out rate was low (i.e. <20%), in 31 studies (37.8%) the drop-out rate was moderate (i.e. ≥ 20% - ≤ 40%), and in 21 studies (25.6%) the drop-out rate was high (i.e. > 40%) (see Table 1). Two studies reported no drop-out rates.

We observed that studies investigating a shorter baseline duration of illness reported patients with a higher severity of positive symptoms, a shorter duration of untreated psychosis (DUP), a younger age and more years of education completed compared to subgroups with a longer baseline duration of illness.

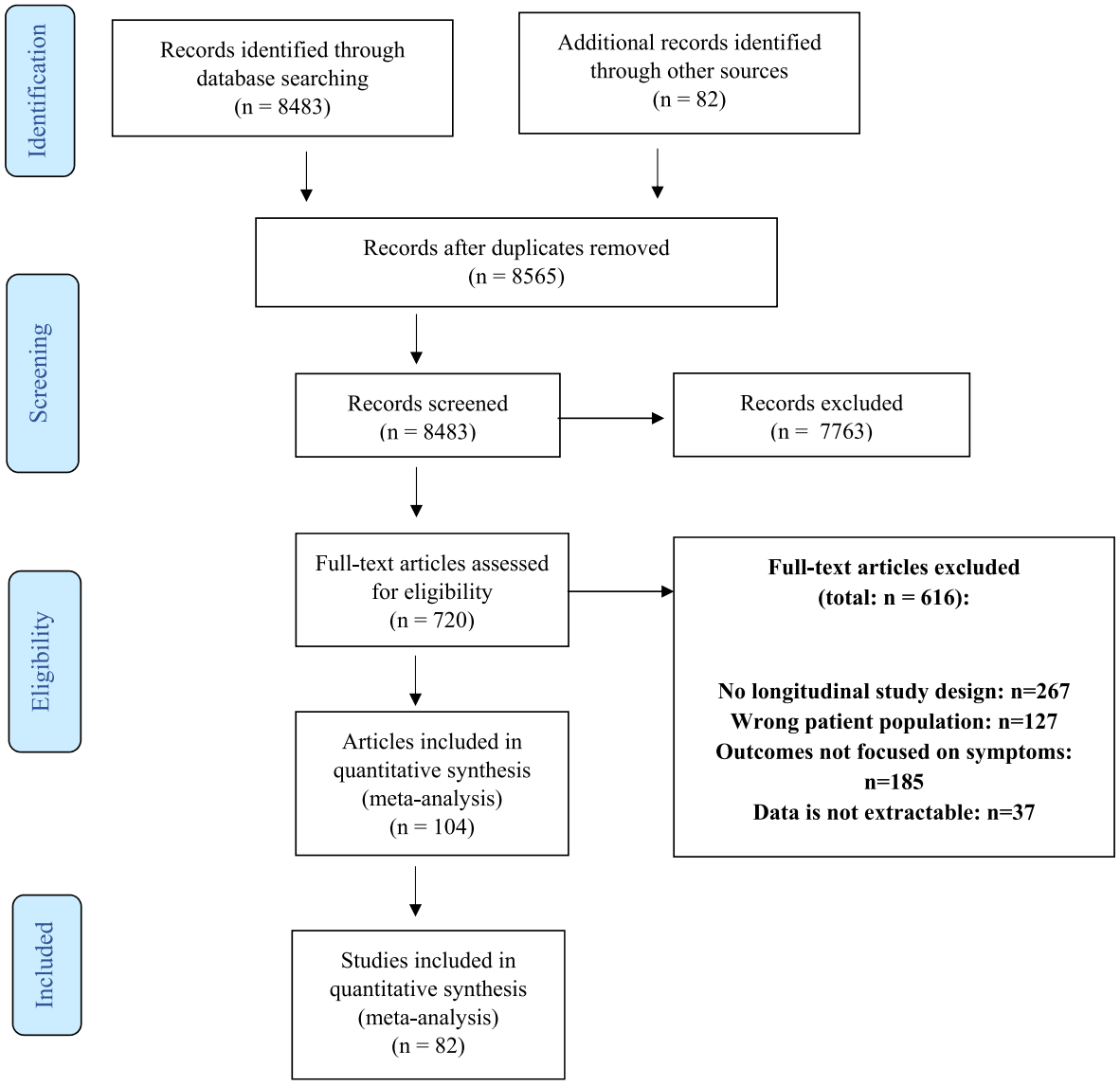

Fig. 1. Flow Chart selection studies conform Prisma Guidelines.

Table 1
Descriptive statistics of included studies.

Study name[a]	N(baseline-FU)	Age (SD)	% female	Primary diagnosis	Comorbidity	Treatment	Baseline DOI (y)	FU duration (y)	Attrition rate	Outcome categories reported
Addington 2000[1]	80–65	33.2 (8.9)	21.10%	Schizophrenia (100%)	NR	Antipsychotics (100%); Routine care (100%)	11.2	2.5	18.8%	Negative symptoms; Positive symptoms
Balanzá -Martínez 2005[2,3]	47–47	33.4 (8.2)	21.30%	Schizophrenia (100%)	NR	Antipsychotics (100%); Antidepressants (12.8%); Benzodiazepines (31.9%); Psychosocial rehabilitation (19.2%)	8.7	1; 3	9.6%	Depressive symptoms; Negative symptoms; Positive symptoms
Bhullar 2017[4]	65–65	28.8 (NR)	24.60%	Schizophrenia spectrum disorder (84.6%); Other psychotic disorder (15.4%)	NR	Prevention and Early Intervention Program for Psychoses (100%)	4.9	10	48.5%	Negative symptoms; Positive symptoms
Breier 2018[6]	60–60	23.6 (4.89)	21.70%	Schizophrenia (68.3%); Schizophreniform disorder (13.3%); Schizoaffective disorder (8.3%); Psychotic disorder NOS (10.0%)	NR	Antipsychotics (100%)	1.4	1	46.7%	Negative symptoms; Positive symptoms
Cechnicki 2017[7]	67–67	26.6 (5.83)	56.70%	Schizophrenia (100%)	NR	Community treatment program (50%); Individual treatment program (50%)	0.8	3; 12	16.3%	Negative symptoms; Positive symptoms
Češková 2007[8,9]	93–93	23.0 (NR)	0.00%	schizophrenia (100%)	NR	Antipsychotics (100%)	0.8	1; 4; 7	52.7%	Negative symptoms; Positive symptoms
Chan 2018[11]	148–107	20.9 (3.07)	49.30%	Schizophrenia spectrum disorder (100%)	NR	Early Assessment Service for Young People with Psychosis (EASY) program (100%)	0	3	27.7%	Negative symptoms; Positive symptoms
Chen 2000[13]	50–43	48.9 (8.9)	30.20%	Schizophrenia (100%)	NR	NA	23.5	3	14.0%	Negative symptoms; Positive symptoms
Chen 2005[14–20]	138–88	31.7 (9.2)	54.80%	Schizophrenia (80.6%); Schizophreniform disorder (14.0%); Schizoaffective disorder (5.4%)	NR	Antipsychotics (48.4%); Antidepressants (12.9%); Benzodiazepines (12.9%)	1.5	1/2/3	39.2%	Depressive symptoms; Negative symptoms; Positive symptoms
Ciudad 2009[21]	1005–375	37.7 (10.5)	35.60%	Schizophrenia (100%)	Substance/Alcohol abuse: 34.3%	NA	13.7	1	16.8%	Depressive symptoms; Negative symptoms; Positive symptoms
Conley 2007[22]	2228–1164	41.8 (11.2)	38.50%	Schizophrenia (57.2%); Schizoaffective disorder (33.6%); Other psychotic disorder (9.2%)	Substance use disorder: 28.0%; Personality disorder: 14.5%; Depressive disorder: 39.4%	Antidepressants (38.8%); Anti-anxiety agents (11.3%); Mood stabilizers (31.2%); Hypnotics (1.7%); Antiparkinsonian agents (44.8%); Atypical antipsychotics (59.8%); Typical antipsychotics (58.2%)	21.6	3	4.3%	Negative symptoms
Cullberg 2002[23,24]	170–170	28.2 (7.07)	45.00%	Schizophrenia syndromes (schizophrenia, schizophreniform psychosis and schizoaffective psychosis;	NR	Need adapted treatment (100%); Antipsychotics (41.8%); Benzodiazepines (70.6%); Antidepressants or lithium (44.7%)	0	1/3/5	30.8%	Depressive symptoms; Negative symptoms

(*continued on next page*)

Table 1 (*continued*)

Study name[a]	N(baseline-FU)	Age (SD)	% female	Primary diagnosis	Comorbidity	Treatment	Baseline DOI (y)	FU duration (y)	Attrition rate	Outcome categories reported
				40.8%); Non-schizophrenia syndromes (delusional disorder, brief psychosis and psychotic disorder not otherwise specified (NOS); 59.2%)						
Dal Santo 2020[25]	17–17	45.4 (8.06)	0.00%	Schizophrenia (100%)	NR	Clozapine (100%)	NR	2.9	0.0%	Positive symptoms
De Haan 2013[26]	176–176	21.1 (3.0)	14.20%	Schizophrenia (57.1%); Schizoaffective disorder (22.2%); Schizophreniform disorder (19.1%)	Obsessive Compulsive Disorder (9.1%)	NA	2.2	0.1/3/5	19.6%	Depressive symptoms
Dixon 2015[27]	65–65	22.2 (4.2)	36.90%	Schizophrenia (66.2%); Schizoaffective disorder (13.9%); Schizophreniform disorder (6.2%); Psychosis NOS (4.6%); Brief psychotic disorder (1.5%); No diagnosis (3.1%); Unknown (4.6%)	Bipolar disorder NOS (3.1%); Depressive disorder NOS (23.1%); Panic disorder (4.6%); Social phobia (3.1%); obsessive compulsive disorder (1.5%); Post-Traumatic Stress Disorder (7.7%); Anxiety disorder NOS (4.6%); Alcohol use disorder (18.5%); Sedative-Hypnotic-Anxiolytic use disorder (1.5%); Cannabis use disorder (33.9%); Stimulant use disorder (1.5%); Opioid use disorder (3.1%); Cocaine use disorder (4.6%); Hallucinogen use disorder (4.6%)	Treatment connection program (100%)	≤2 y	0.5/1/1.5/2	69.2%	Negative symptoms; Positive symptoms
Domen 2017[28]	55–55	28.1 (7.05)	23.50%	Schizophrenia (100%)	NR	NR	5.7	3/4.6	4.4%	Disorganization symptoms; Negative symptoms; Positive symptoms
Evensen 2016[29]	148–148	32.9 (7.94)	30.60%	Schizophrenia (88.5%); Schizoaffective disorder (7.5%); Psychosis NOS (2.0%); Delusional disorder (2.1%)	NR	Cognitive Behavioral Therapy (56.8%); Cognitive remediation (43.2%)	7.2	2	12.2%	Depressive symptoms
Friedman 2002[30]	124–124	72.4 (6.3)	54.80%	Schizophrenia (100%)	NR	Antipsychotics (74.0%); Anticholinergics (13.0%)	NR	1.2/4	59.7%	Negative symptoms; Positive symptoms
Galderisi 2020[31]	921–618	40.2 (10.7)	30.40%	Schizophrenia (100%)	Substance abuse (5.0%); Alcohol abuse (4.9%)	Antipsychotics (76.8%); Integrated treatment (26.8%)	16.2	4	32.9%	Depressive symptoms; Disorganization symptoms; Negative

(*continued on next page*)

Table 1 (*continued*)

Study name[a]	N(baseline-FU)	Age (SD)	% female	Primary diagnosis	Comorbidity	Treatment	Baseline DOI (y)	FU duration (y)	Attrition rate	Outcome categories reported
										symptoms; Positive symptoms
Ganella 2018[32]	29–14	21.3 (2.0)	24.10%	First Episode Psychotic Disorder (100%)	NR	NR	1.4	1	51.7%	Depressive symptoms; Negative symptoms; Positive symptoms
Gaughran 2017[33]	403–259	44.2 (10.1)	42.40%	Psychotic disorder (100%)	NR	Health promotion intervention (52.5%)	NR	1/1.3	25.9%	Depressive symptoms
Godin 2019[34]	770–325	32.7 (9.9)	26.00%	Schizophrenia (100%)	Anxiety disorder (37.4%); Tobacco smoking (51.6%); Cannabis use disorder (28.4%); Alcohol use disorder (20.1%)	Antipsychotics (21.2%); Antidepressants (25.7%)	10.7	1	57.8%	Depressive symptoms
Gorwood 2019[35]	303–228	29.3 (4.9)	26.40%	Schizophrenia (100%)	NR	Antipsychotics (100%); Benzodiazepines (38.3%); Other psychotropic treatment (59.4%); Psychotherapy (57.3%); Psychosocial care (37.8%)	7.6	0.5/1	26.1%	Negative symptoms; Positive symptoms
Granholm 2010[36]	107–101	56.1 (8.4)	38.30%	Schizophrenia (100%)	NR	Antipsychotics (94.0%); Anticholinergics (43.1%); Antidepressants (58.8%)	NR	1/2/3	52.3%	Negative symptoms; Positive symptoms
Häfner 1999[37]	115–89	27.9 (NR)	50.40%	Schizophrenia (100%)	Depression (82.6%); Alcohol abuse (28.7%); Drug abuse (21.7%)	Antipsychotics (12.0%)	0	0.5/1/2/3/5	50.4%	Depressive symptoms
Haro 2018[38]	1344–1272	42 (11.43)	29.10%	Schizophrenia (100%)	Mood disorder (5.3%); Anxiety disorder (4.8%); Drug abuse (24.7%)	Antipsychotics (99.4%)	12.2	1	11.0%	Negative symptoms; Positive symptoms
Harvey 1996[39]	174–174	76.3 (10.8)	64.20%	Schizophrenia (100%)	NR	NA	46.9	1/6	34.4%	Negative symptoms; Positive symptoms
Harvey 1999[40]	57–57	77.8 (8.2)	56.10%	Schizophrenia (100%)	NR	Antipsychotics (100%); Anticholinergics (8.8%); Benzodiazepines (14.0%); Anticonvulsants (5.3%)	47.1	2.6	3.5%	Negative symptoms; Positive symptoms
Hauser 1990[41]	67–67	36.7 (9.4)	46.30%	Schizophrenia (100%)	NR	Antipsychotics (76.1%)	10.3	2.5	31.5%	Positive symptoms
Hill 2012[42]	123–123	29.1 (12.0)	42.10%	Schizophrenia/ schizophreniform disorder (59.1%); Other psychosis (40.9%)	Substance abuse (25.5%)	NA	1.9	12	28.1%	Disorganization symptoms; Negative symptoms; Positive symptoms
Hoff 2005[43]	21–21	26.3 (7.4)	26.20%	Schizophrenia (74.3%); Schizoaffective disorder (25.7%)	NR	Typical antipsychotics (92.9%)	1.5	10	50.0%	Negative symptoms; Positive symptoms
Horan 2006[44]	89–34	23.8 (4.9)	15.70%	Schizophrenia (100%)	NR	Antipsychotics (100%); Behaviorally oriented individual case management (100%); Group psychosocial therapy (100%)	1.2	1.3	61.8%	Negative symptoms; Positive symptoms

(*continued on next page*)

Table 1 (*continued*)

Study name[a]	N(baseline-FU)	Age (SD)	% female	Primary diagnosis	Comorbidity	Treatment	Baseline DOI (y)	FU duration (y)	Attrition rate	Outcome categories reported
Hwu 2002[46]	163–163	30.4 (7.3)	45.40%	Schizophrenia (100%)	NR	Antipsychotics (100%); Psychosocial management (100%)	7.9	1	41.9%	Disorganization symptoms; Negative symptoms; Positive symptoms
Hyza 2016[47]	58–52	23.4 (5.1)	0.00%	Schizophrenia (100%)	NR	Atypical antipsychotics (100%)	0.6	1/4	34.5%	Negative symptoms; Positive symptoms
Ito 2015[48]	76–55	30.6 (10.1)	53.20%	Schizophrenia spectrum disorder (100%)	NR	Antipsychotics (100%)	2	0.5/1/1.5	53.9%	Negative symptoms; Positive symptoms
Jørgensen 2015[49]	101–94	37.5 (12.6)	53.50%	Schizophrenia (92.1%); Schizoaffective disorder (7.9%)	NR	Antipsychotics (100%); Guided self-determination intervention (100%)	9.8	0.25/0.5/1	7.9%	Negative symptoms; Positive symptoms
Kane 2016[51,52]	404–404	23.1 (5.1)	27.50%	Schizophrenia (53.0%); Schizoaffective disorder, bipolar (6.0%); Schizoaffective disorder, depressive (14.1%); Schizophreniform disorder (16.6%); Brief psychotic disorder (0.5%); Psychotic disorder NOS (9.9%)	Alcohol abuse/ dependence (36.4%); Cannabis abuse/ dependence (35.6%)	Antipsychotics (83.4%); NAVIGATE treatment (55.2%); Community care (44.8%)	3.7	0.5/1/1.5/2	43.8%	Depressive symptoms; Disorganization symptoms; Positive symptoms
Kelly 2009[53]	43–43	44.1 (8.3)	27.90%	Schizophrenia (100%)	NR	Haloperidol (58.1%); Olanzapine (41.9%)	22.1	1	23.2%	Negative symptoms; Positive symptoms
Klærke 2019[54]	70–70	26.5 (6.2)	28.60%	Schizophrenia (95.7%); Schizoaffective disorder (4.3%)	Substance abuse (11.0%)	NR	2.1	9.6	51.1%	Negative symptoms; Positive symptoms
Koshiyama 2017[55]	14–14	22.6 (5.4)	21.40%	First Episode Psychosis (100%)	NR	Antipsychotics (100%)	0.7	1.9	NR	Negative symptoms; Positive symptoms
Landolt 2012[58]	341–340	26.0 (5.6)	40.20%	Schizophrenia (53.2%); Schizophreniform disorder (39.8%); Schizoaffective disorder (7.0%)	Major Depressive Disorder (36.1%)	Haloperidol (20.7%); Amisulpride (20.9%); Olanzapine (21.1%); Quetiapine (20.9%); Ziprasidone (16.5%)	≤2 y	1	31.3%	Depressive symptoms
Lee 1991[60]	153–97	32.0 (7.5)	47.10%	Schizophrenia (100%)	NR	NA	7.4	1	36.6%	Negative symptoms; Positive symptoms
Lindenmayer 1987[61]	37–19	24.0 (4.2)	40.50%	Schizophrenia (100%)	NR	Antipsychotics (73.7%)	1.4	2.2	48.7%	Depressive symptoms
Liu 2011[62]	31–31	27.6 (10.0)	32.30%	Schizophrenia (58.1%); Schizoaffective disorder (6.4%); Schizophreniform disorder (35.5%)	NR	Antipsychotics (100%); Antidepressants (12.9%); Anticholinergics (35.5%); Benzodiazepines (9.7%)	NR	1/3	0.0%	Negative symptoms; Positive symptoms
Luckhoff 2018[63]	106–64	24.2 (NR)	27.40%	Schizophrenia (74.5%); Schizophreniform disorder (24.5%); Schizoaffective disorder (0.9%)	Substance abuse (47.2%)	Antipsychotics (100%)	0.7	1	39.6%	Disorganization symptoms; Negative symptoms; Positive symptoms
Mäkinen 2010[65]	38–38	23.2 (4.1)	59.00%	Schizophrenia (100%)	Substance abuse (40.0%); Personality disorder (38.5%)	NA	0.5	10	24.6%	Negative symptoms

(*continued on next page*)

Table 1 (*continued*)

Study name[a]	N(baseline-FU)	Age (SD)	% female	Primary diagnosis	Comorbidity	Treatment	Baseline DOI (y)	FU duration (y)	Attrition rate	Outcome categories reported
Malla 2007[66–67]	153–153	24.3 (6.6)	22.00%	Schizophrenia (48.0%); Schizophreniform disorder (17.2%); Schizoaffective disorder (10.1%); Delusional disorder (3.5%); Brief psychotic disorder (2.5%); Psychosis NOS (18.7%)	Depressive disorder (30.6%); Anxiety disorder (22.7%); Substance related disorder (18.3%)	Antipsychotics (100%)	1.4	1/2/5/10	23.5%	Negative symptoms; Positive symptoms
McGurk 2000[68]	168–168	74.2 (6.6)	51.80%	Schizophrenia (100%)	NR	NA	NR	1.3	0.0%	Negative symptoms; Positive symptoms
McGurk 2003[69]	30–27	39.7 (6.9)	23.30%	Schizophrenia (53.3%); Schizaffective disorder (46.7%)	NR	Supported employment (100%); Antipsychotics (100%); Clozapine (23.3%); Risperidone (26.7%); Olanzapine (13.3%)	15.7	2	10.0%	Depressive symptoms; Negative symptoms; Positive symptoms
Meagher 2004[70]	82–82	68.7 (10.1)	41.90%	Schizophrenia (100%)	NR	Antipsychotics (100%)	44.7	2.9	36.4%	Negative symptoms; Positive symptoms
Melle 2008[71,72]	281–231	30.0 (10.0)	44.30%	Schizophrenia spectrum disorder (72.1%)	Alcohol use problems (7.5%); Drug use problems (10.0%)	First-episode treatment programs consisting of antipsychotic psychopharmacology, assertively oriented individual outpatient treatment, and psychoeducational family work (100%).	≤2 y	0.25/1/2/10	33.2%	Negative symptoms; Positive symptoms
Na 2016[73]	25–25	28.1 (6.4)	48.00%	Schizophrenia (60.0%); Schizoaffective disorder (12.0%); Psychotic disorder NOS (28.0%)	NR	Antipsychotics (100%); Mind flower program (100%)	NR	0.5/1	4.0%	Negative symptoms; Positive symptoms
Nakamura 2019[74]	37–37	61.0 (7.8)	43.20%	Schizophrenia (100%)	NR	Illness Management and Recovery (100%)	34.7	1	14.0%	Negative symptoms; Positive symptoms
Nordentoft 2006[75,76]	83–62	26.6 (24.4)	54.20%	Schizotypal disorder (100%)	Alcohol or substance abuse (20.5%)	Antipsychotics (67.5%)	9.8	1/2/3.5	28.9%	Disorganization symptoms; Negative symptoms; Positive symptoms
Oh 2017[77]	22–22	22.9 (5.6)	52.50%	Psychotic disorder (100%)	NR	NR	≤2 y	1	45.0%	Negative symptoms; Positive symptoms
Oribe 2015[79,80]	18–18	21.7 (4.6)	27.80%	Schizophrenia (100%)	NR	Atypical antipsychotics (72.2%); Mood stabilizers (5.6%); Antidepressants (33.3%); Anxiolytics (16.7%)	1.2	1	0.0%	Negative symptoms; Positive symptoms
Ozawa 2019[81]	35–35	63.9 (7.9)	34.30%	Schizophrenia (100%)	NR	Antipsychotics (100%)	35.7	1	25.5%	Depressive symptoms; Negative symptoms; Positive symptoms

(*continued on next page*)

Table 1 (*continued*)

Study name[a]	N(baseline-FU)	Age (SD)	% female	Primary diagnosis	Comorbidity	Treatment	Baseline DOI (y)	FU duration (y)	Attrition rate	Outcome categories reported
Pillman 2002[82]	52–47	34.6 (11.1)	80.80%	Schizophrenia (50%); Brief psychotic disorder (50%)	NR	NA	5.8	2.5	9.6%	Positive symptoms
Putnam 1996[83,84]	233–233	67.4 (15.8)	49.90%	Schizophrenia (100%)	NR	Antipsychotics (81.0%)	41.5	1	24.4%	Negative symptoms; Positive symptoms
Rais 2008[85]	51–51	22.7 (4.7)	11.80%	Schizophrenia (76.5%); Schizophreniform disorder (17.7%); Schizoaffective disorder (2.0%); Psychosis NOS (3.9%)	Cannabis abuse (37.3%)	Antipsychotics (100%)	1	5	0.0%	Negative symptoms; Positive symptoms
Rodríguéz-Sánchez 2008[86–88]	307–307	29.1 (8.9)	42.00%	Schizophrenia (60.0%); Schizophreniform disorder (46.7%); Psychosis NOS (3.3%); Brief psychotic disorder (6.0%)	NR	Antipsychotics (100%); Anticholinergics (6.5%); Hypnotics (15.5%); Benzodiazepines (52.9%); PAFIP (100%)	2.3 (Rodriguez-Sanchez et al., 2008); 1.2 (Pelayo-Terran, 2018)	0.1/1/3	27.5%	Depressive symptoms; Disorganization symptoms; Negative symptoms; Positive symptoms
Rossi 2009[89]	326–326	44.2 (11.4)	38.00%	Schizophrenia (74.9%); Schizoaffective disorder (25.1%)	NR	Risperidone (100%); Benzodiazepines (45.6%)	17.3	1	30.0%	Negative symptoms; Positive symptoms
Ryu 2006[90–92]	78–78	54.6 (7.2)	34.60%	Schizophrenia (100%)	NR	Optimal Treatment Project strategies (100%)	31.5	1/2/3/4/5/6/12/15	28.2%	Negative symptoms; Positive symptoms
Salyers 2014[93–94]	118–118	47.7 (8.9)	20.70%	Schizophrenia (46.6%); Schizoaffective disorder (55.2%)	NR	Illness Management and Recovery (50.8%); Intensive problem-solving (49.2%); Usual treatment (100%)	NR	0.75/1.5	40.7%	Negative symptoms; Positive symptoms
Scottish Schizophrenia Research group 1989[95]	48–34	30.6 (NR)	53.10%	Schizophrenia (100%)	NR	antipsychotics (100%)	0.2	1/2/5	16.3%	Negative symptoms
She 2017[96]	170–169	32.4 (8.3)	37.10%	Schizophrenia (100%)	NR	Integrated group treatment (50.6%); Antipsychotics (100%)	7.2	0.25/0.5/1	36.5%	Negative symptoms; Positive symptoms
Siegel 2006[97]	98–92	28.6 (7.4)	40.80%	Schizophrenia (100%)	NR	Antipsychotics (85.9%)	6.1	3	52.9%	Depressive symptoms; Negative symptoms; Positive symptoms
Smith 2002[98]	46–45	37.0 (9.0)	41.30%	Schizophrenia (60.9%); Schizoaffective disorder (39.1%)	NR	Outpatient treatment program (100%); Antipsychotics (100%)	19	0.25/0.5/0.75/1	37.5%	Disorganization symptoms; Negative symptoms; Positive symptoms
Sweeney 1991[99]	39–39	28.6 (8.6)	38.50%	Schizophrenia (74.4%); Schizophreniform disorder (10.3%); Schizoaffective disorder (15.4%)	NR	Antipsychotics (100%)	6.6	0.25/1/1.25/1.5	0.0%	Negative symptoms; Positive symptoms
Tabo 2017[100]	120–120	40.9 (10.9)	28.30%	Schizophrenia (100%)	NR	NR	16.3	1	NR	Negative symptoms; Positive symptoms

(*continued on next page*)

Table 1 (*continued*)

Study name[a]	N(baseline-FU)	Age (SD)	% female	Primary diagnosis	Comorbidity	Treatment	Baseline DOI (y)	FU duration (y)	Attrition rate	Outcome categories reported
Torgalsbøen 2015[101,102]	25–25	21.0 (2.6)	39.30%	Schizophrenia (75.0%); Schizoaffective disorder (21.4%); Psychotic disorder NOS (3.6%)	Substance abuse (3.6%)	Psychotherapy (71.4%); Group therapy (7.1%); Psychoeducation (64.3%)	<0.4 y	2	10.7%	Negative symptoms; Positive symptoms
Üçok 2011[104]	96–96	21.1 (4.8)	43.30%	Schizophrenia (100%)	Alcohol and/or substance use (16.5%)	Antipsychotics (63.0%)	1.2	0.25/1/2/3/4	25.6%	Negative symptoms; Positive symptoms
Van Haren 2007[105]	96–96	32.2 (11.1)	27.10%	Schizophrenia (100%)	NR	Antipsychotics (92.7%)	11	4.8	39.6%	Negative symptoms; Positive symptoms
Veerman 2016[106]	25–25	42.0 (10.4)	24.00%	Schizophrenia (100%)	Alcohol use (20%); Nicotine use (56%); Cocaine use (12%)	Clozapine (100%); Psychotherapy (8%)	19.6	1	19.4%	Depressive symptoms; Negative symptoms; Positive symptoms
Waddington 1995[108]	49–49	62.6 (13.4)	43.20%	Schizophrenia (100%)	NR	Antipsychotics (100%)	34.6	5/10	51.5%	Negative symptoms
Wang 2010[109]	374–374	32.6 (10.8)	57.40%	Schizophrenia (100%)	Tobacco use (12.9%)	Risperidone (100%)	6.7	1	7.4%	Depressive symptoms; Disorganization symptoms; Negative symptoms; Positive symptoms
Whitehorn 2002[110]	103–56	21.9 (5.7)	33.10%	Schizophrenia spectrum disorder (100%)	NR	Second generation antipsychotic use (100%); Multidisciplinary treatment (100%); Psychoeducation (100%)	≤2 y	0.5/1	52.4%	Disorganization symptoms; Negative symptoms; Positive symptoms
Wilson-d'Almeida 2013[112]	306–306	41.1 (10.1)	30.10%	Schizophrenia (100%)	NR	Antipsychotics (100%)	NR	0.5/1	12.3%	Negative symptoms; Positive symptoms
Wittorf 2004[113]	11–11	31.9 (10.9)	66.70%	Schizophrenia (93.3%); Schizoaffective disorder (6.7%)	NR	Antipsychotics (100%)	6.1	1.1	26.7%	Disorganization symptoms; Negative symptoms; Positive symptoms
Wunderink 2009[114]	125–125	26.4 (6.4)	31.20%	schizophrenia (45.6%); Other non-affective psychosis (54.4%)	Cannabis dependence (24.0%)	Antipsychotics (100%)	0.7	0.5/1.25/2	14.4%	Negative symptoms; Positive symptoms
Xie 2005[115,116]	152–152	32.4 (7.2)	22.40%	Schizophrenia (70.4%); Schizoaffective disorder (29.6%)	Substance use disorder (100%); Alcohol use disorder (81.6%); Cannabis use disorder (44.7%); Cocaine use disorder (15.1%); Bipolar disorder (100%)	Dual disorder treatment (100%)	12	0.5/1/1.5/2/2.5/3/4/5/6/7/8/9/10	23.1%	Disorganization symptoms; Negative symptoms; Positive symptoms
Xu 2014[117]	60–60	25.3 (10.4)	45.00%	Schizophrenia (51.7%); Schizophreniform disorder (20.0%); Psychosis NOS (21.7%); Schizoaffective disorder (6.7%)	NR	Antipsychotics (95.0%); Anticholinergics (18.3%)	0	1	23.1%	Disorganization symptoms; Negative symptoms; Positive symptoms

Abbreviations: NA = Not Applicable; NR = Not Reported; y = years.

[a] The reference list of the included studies are presented in Supplementary materials 8.

3.3. Meta-analysis of study outcomes with different durations of illness

We presented a general overview of the outcomes and differences between duration of illness (DOI) and follow-up subgroups in Fig. 2 and Table 2. In the text below *d* stands for the effect size of change, I^2 stands for the heterogeneity of the outcomes and *k* stands for the number of studies reporting on the outcomes.

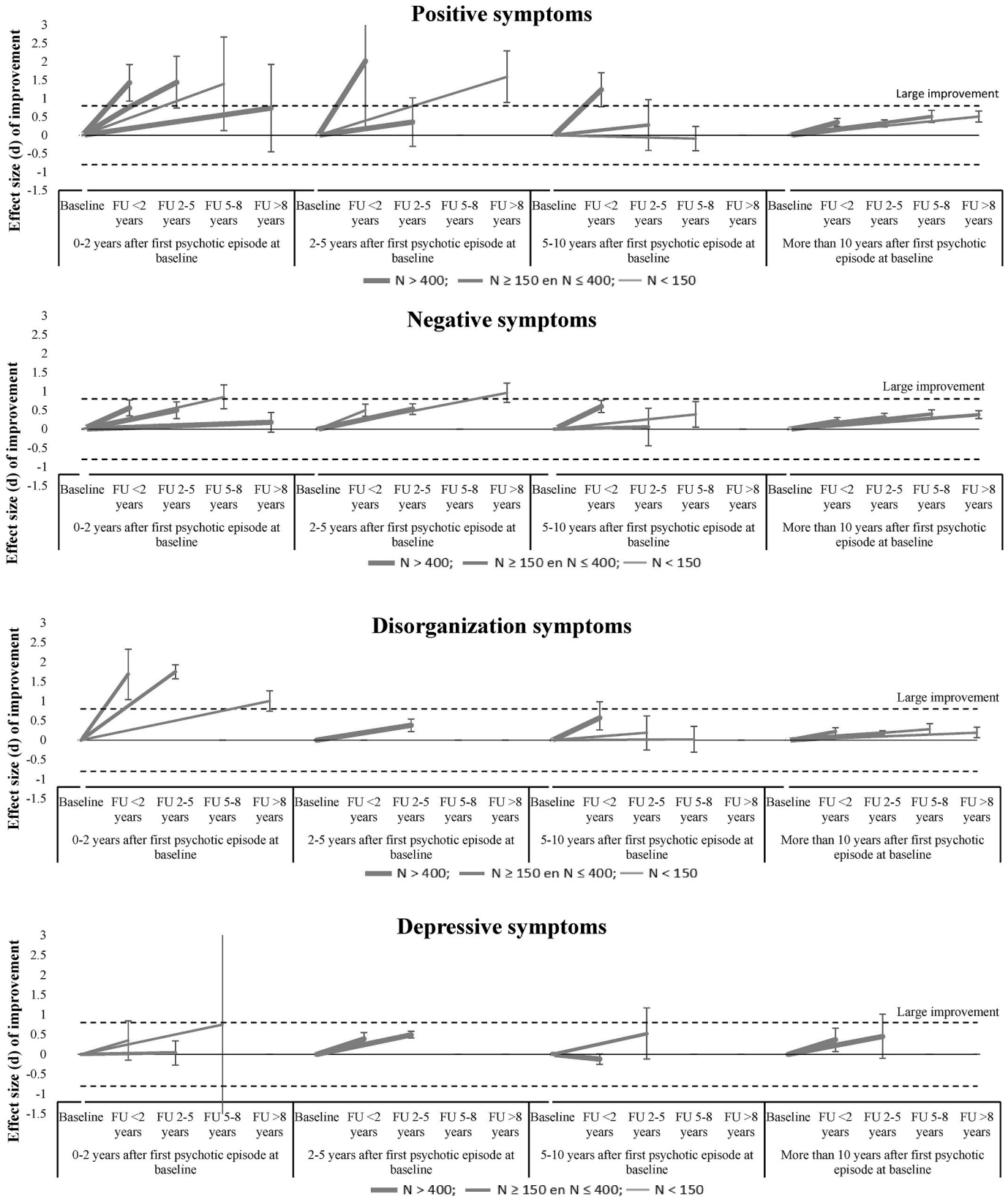

Fig. 2. Effect sizes of improvement and/or deterioration of the five symptomatic outcome categories within the four baseline DOI subgroups.

Table 2
Meta-analysis of symptomatic outcomes.

Positive symptoms						
(Sub)analysis		K (studies (outcomes))	N (baseline-FU)	Effect size (95% CI)[a] and magnitude of effect[b]	K (%) large effect[b] [+/−][c]	Heterogeneity (I^e (95%CI))[a]
All studies and outcomes		74 (145)	9882–9025	d = **0.84** [L] (0.71–0.96)	+ = 49 (33.8%)/- = 0 (0.0%)	I^e = **97%** (96–97%)
Subgroups						
Baseline subgroup	*Follow-up cohort*					
Duration of illness <2 years	<2 years	15 (20)	1060–921	d = **1.43** [L] (0.93–1.92)[g]	+ = 12 (60.0%)/- = 0 (0.0%)	I^e = **97%** (96–98%)
	≥2 - < 5 years	10 (13)	1026–965	d = **1.44** [L] (0.74–2.15)[234]	+ = 8 (61.5%)/- = 0 (0.0%)	I^e = **99%** (98–99%)
	≥5 - < 8 years	2 (2)	95–95	d = **1.40** [L] (0.13–2.67)[f]	+ = 1 (50.0%)/- = 0 (0.0%)	I^e = **93%** (NA)
	≥8 years	4 (4)	492–397	d = 0.74 [M] (−0.45 – 1.93)	+ = 2 (50.0%)/- = 0 (0.0%)	I^e = **98%** (97–99%)
	Subgroup differences between follow-up cohorts			χ^e = 1.16; df = 3; p = 0.76		
Duration of illness 2–5 years	<2 years	2 (3)	516–516	d = **2.02** [L] (0.21–3.82)	+ = 2 (66.7%)/- = 0 (0.0%)	I^e = **99%** (99–100%)
	≥2 - < 5 years	2 (2)	536–363	d = 0.36 [S] (−0.30 – 1.02)[d]	+ = 0 (0.0%)/- = 0 (0.0%)	I^e = **95%** (NA)
	≥8 years	2 (2)	135–135	d = **1.59** [L] (0.89–2.30)[g]	+ = 2 (100.0%)/- = 0 (0.0%)	I^e = **84%** (NA)
	Subgroup differences between follow-up cohorts			χ^e = 7.57; df = 2; p < 0.05		
Duration of illness 5–10 years	<2 years	9 (15)	1339–1190	d = **1.24** [L] (0.78–1.70)[g]	+ = 10 (66.7%)/- = 0 (0.0%)	I^e = **98%** (97–98%)
	≥2 - < 5 years	5 (5)	299–266	d = 0.28 [S] (−0.41 – 0.97)[d]	+ = 1 (20.0%)/- = 0 (0.0%)	I^e = **93%** (87–97%)
	≥5 - < 8 years	1 (1)	83–59	d = −0.09 [O] (−0.42 – 0.24)[d]	+ = 0 (0.0%)/- = 0 (0.0%)	Not Applicable
	Subgroup differences between follow-up cohorts			χ^e = 20.77; df = 2; p < 0.01		
Duration of illness >10 years	<2 years	15 (32)	3169–3068	d = **0.35** [S] (0.24–0.46)[13]	+ = 3 (9.4%)/- = 0 (0.0%)	I^e = **83%** (79–87%)
	≥2 - < 5 years	10 (16)	1859–1462	d = **0.32** [S] (0.23–0.42)[d]	+ = 0 (0.0%)/- = 0 (0.0%)	I^e = **54%** (38–66%)
	≥5 - < 8 years	2 (4)	208–203	d = **0.51** [M] (0.35–0.68)	+ = 0 (0.0%)/- = 0 (0.0%)	I^e = 37% (0–68%)
	≥8 years	1 (3)	130–120	d = **0.51** [M] (0.36–0.66)[e]	+ = 0 (0.0%)/- = 0 (0.0%)	I^e = 0% (0–60%)
	Subgroup differences between follow-up cohorts			χ^e = 7.40; df = 3; p = 0.06		
Duration of illness unclear	<2 years	9 (14)	1211–1126	d = **0.91** [L] (0.60–1.22)	+ = 6 (42.9%)/- = 0 (0.0%)	I^e = **97%** (96–98%)
	≥2 - < 5 years	5 (6)	510–486	d = 0.78 [M] (−0.11 – 1.67)	+ = 2 (33.3%)/- = 0 (0.0%)	I^e = **98%** (96–99%)
	Subgroup differences between follow-up cohorts			χ^e = 0.08; df = 1; p = 0.78		
Negative symptoms						
(Sub)analysis		K (studies (outcomes))	N (baseline-FU)	Effect size (95% CI)[a] and magnitude of effect[b]	K (%) large effect[b] [+/−][c]	Heterogeneity (I^e (95%CI))[a]
All studies and outcomes		74 (147)	12159–10018	d = **0.39** [S] (0.34–0.45)	+ = 26 (17.7%)/- = 0 (0.0%)	I^e = **89%** (87–90%)
Subgroups						
Baseline subgroup	*Follow-up cohort*					
Duration of illness <2 years	<2 years	17 (22)	1281–1128	d = **0.56** [M] (0.35–0.77)[g]	+ = 6 (27.3%)/- = 0 (0.0%)	I^e = **89%** (85–92%)
	≥2 - < 5 years	12 (15)	1196–1105	d = **0.50** [M] (0.28–0.72)	+ = 4 (26.7%)/- = 0 (0.0%)	I^e = **85%** (77–89%)
	≥5 - < 8 years	2 (2)	95–95	d = **0.85** [L] (0.54–1.17)[g]	+ = 1 (50.0%)/- = 0 (0.0%)	I^e = 9% (NA)
	≥8 years	6 (6)	668–531	d = 0.18 [O] (−0.08 – 0.44)[e]	+ = 0 (0.0%)/- = 0 (0.0%)	I^e = **75%** (49–88%)
	Subgroup differences between follow-up cohorts			χ^e = 10.92; df = 3; p < 0.05		
Duration of illness 2–5 years	<2 years	1 (2)	112–112	d = 0.50 [M] (0.34–0.66)[g]	+ = 0 (0.0%)/- = 0 (0.0%)	I^e = 0% (NA)
	≥2 - < 5 years	3 (3)	629–456	d = **0.53** [M] (0.39–0.67)[g]	+ = 0 (0.0%)/- = 0 (0.0%)	I^e = 17% (0–50%)
	≥8 years	2 (2)	135–135	d = **0.96** [L] (0.71–1.22)[14]	+ = 2 (100.0%)/- = 0 (0.0%)	I^e = 0% (NA)
	Subgroup differences between follow-up cohorts			χ^e = 10.17; df = 2; p < 0.01		
Duration of illness 5–10 years	<2 years	8 (14)	1260–1128	d = **0.60** [M] (0.44–0.76)[g]	+ = 6 (42.9%)/- = 0 (0.0%)	I^e = **87%** (81–91%)
	≥2 - < 5 years	4 (4)	247–219	d = 0.06 [O] (−0.44 – 0.55)	+ = 0 (0.0%)/- = 0 (0.0%)	I^e = **84%** (58–94%)
	≥5 - < 8 years	1 (1)	83–59	d = 0.39 [S] (0.05–0.73)	+ = 0 (0.0%)/- = 0 (0.0%)	Not Applicable
	Subgroup differences between follow-up cohorts			χ^e = 4.84; df = 2; p = 0.09		
Duration of illness >10 years	<2 years	14 (31)	3189–3109	d = **0.21** [S] (0.12–0.31)[123]	+ = 2 (6.5%)/- = 0 (0.0%)	I^e = **86%** (83–89%)
	≥2 - < 5 years	11 (17)	4100–2624	d = **0.30** [S] (0.18–0.42)[e]	+ = 0 (0.0%)/- = 0 (0.0%)	I^e = **85%** (78–89%)
	≥5 - < 8 years	3 (5)	257–250	d = **0.40** [S] (0.29–0.51)[d]	+ = 0 (0.0%)/- = 0 (0.0%)	I^e = 0% (0–31%)
	≥8 years	2 (4)	179–169	d = **0.38** [S] (0.28–0.49)[e]	+ = 0 (0.0%)/- = 0 (0.0%)	I^e = 0% (0–36%)
	Subgroup differences between follow-up cohorts			χ^e = 8.58; df = 3; p < 0.05		
Duration of illness unclear	<2 years	9 (14)	1213–1160	d = **0.49** [S] (0.28–0.70)	+ = 5 (35.7%)/- = 0 (0.0%)	I^e = **91%** (87–94%)
	≥2 - < 5 years	4 (5)	493–469	d = 0.12 [O] (−0.23 – 0.47)	+ = 0 (0.0%)/- = 0 (0.0%)	I^e = **87%** (71–94%)
	Subgroup differences between follow-up cohorts			χ^e = 3.12; df = 1; p = 0.08		
Disorganization symptoms						
(Sub)analysis		K (studies (outcomes))	N (baseline-FU)	Effect size (95% CI)[a] and magnitude of effect[b]	K (%) large effect[b] [+/−][c]	Heterogeneity (I^e (95%CI))[a]
All studies and outcomes		14 (37)	2908–2321	d = **0.47** [S] (0.32–0.62)	+ = 7 (18.9%)/- = 0 (0.0%)	I^e = **93%** (91–94%)
Subgroups						

(*continued on next page*)

Table 2 (*continued*)

Positive symptoms						
(Sub)analysis		K (studies (outcomes))	N (baseline-FU)	Effect size (95% CI)[a] and magnitude of effect[b]	K (%) large effect[b] [+/−][c]	Heterogeneity (I^e (95%CI))[a]
Baseline subgroup	*Follow-up cohort*					
Duration of illness <2 years	<2 years	2 (2)	166–124	d = **1.69** [L] (1.04–2.33)[34]	+ = 2 (100.0%)/- = 0 (0.0%)	I^e = **83%** (NA)
	≥2 - < 5 years	1 (1)	307–307	d = **1.75** [L] (1.57–1.93)[234]	+ = 1 (100.0%)/- = 0 (0.0%)	Not Applicable
	≥8 years	1 (1)	123–123	d = **1.00** [L] (0.74–1.26)[g]	+ = 1 (100.0%)/- = 0 (0.0%)	Not Applicable
	Subgroup differences between follow-up cohorts			$\chi^e = 20.98$; $df = 2$; $p < 0.01$		
Duration of illness 2–5 years	≥2 - < 5 years	1 (1)	404–231	d = **0.38** [S] (0.22–0.54)[14]	+ = 0 (0.0%)/- = 0 (0.0%)	Not Applicable
	Subgroup differences between follow-up cohorts			Not Applicable		
Duration of illness 5–10 years	<2 years	4 (7)	627–610	d = 0.57 [M] (0.26–0.98)[14]	+ = 1 (14.3%)/- = 0 (0.0%)	I^e = **91%** (85–95%)
	≥2 - < 5 years	2 (2)	143–112	d = 0.19 [O] (−0.25 – 0.62)[d]	+ = 0 (0.0%)/- = 0 (0.0%)	I^e = 66% (NA)
	≥5 - < 8 years	1 (1)	83–59	d = 0.02 [O] (−0.31 – 0.35)[d]	+ = 0 (0.0%)/- = 0 (0.0%)	Not Applicable
	Subgroup differences between follow-up cohorts			$\chi^e = 5.80$; $df = 2$; $p = 0.06$		
Duration of illness >10 years	<2 years	2 (8)	198–197	d = 0.22 [S] (0.11–0.32)[13]	+ = 0 (0.0%)/- = 0 (0.0%)	I^e = 0% (0–33%)
	≥2 - < 5 years	2 (6)	1073–770	d = 0.16 [O] (0.09–0.24)[12]	+ = 0 (0.0%)/- = 0 (0.0%)	I^e = 0% (0–56%)
	≥5 - < 8 years	1 (3)	130–125	d = 0.28 [S] (0.14–0.42)	+ = 0 (0.0%)/- = 0 (0.0%)	I^e = 0% (0–33%)
	≥8 years	1 (3)	130–120	d = 0.19 [O] (0.06–0.33)[d]	+ = 0 (0.0%)/- = 0 (0.0%)	I^e = 0% (0–33%)
	Subgroup differences between follow-up cohorts			$\chi^e = 2.36$; $df = 3$; $p = 0.50$		
Duration of illness unclear	<2 years	1 (2)	103–56	d = **1.05** [L] (0.81–1.30)	+ = 2 (100.0%)/- = 0 (0.0%)	I^e = 0% (NA)
	Subgroup differences between follow-up cohorts			Not Applicable		
Depressive symptoms						
(Sub)analysis		K (studies (outcomes))	N (baseline-FU)	Effect size (95% CI)[a] and magnitude of effect[b]	K (%) large effect[b] [+/−][c]	Heterogeneity (I^e (95%CI))[a]
All studies and outcomes		20 (30)	5316–3736	d = **0.33** [S] (0.21–0.45)	+ = 3 (10.0%)/- = 0 (0.0%)	I^e = **84%** (79–87%)
Subgroups						
Baseline subgroup	*Follow-up cohort*					
Duration of illness <2 years	<2 years	2 (3)	144–103	d = 0.35 [S] (−0.15 – 0.84)	+ = 0 (0.0%)/- = 0 (0.0%)	I^e = 0% (0–48%)
	≥2 - < 5 years	3 (4)	267–203	d = 0.04 [O] (−0.27 – 0.34)[e]	+ = 0 (0.0%)/- = 0 (0.0%)	I^e = 0% (0–79%)
	≥5 - < 8 years	1 (1)	115–86	d = 0.75 [M] (−1.57 – 3.07)	+ = 0 (0.0%)/- = 0 (0.0%)	Not Applicable
	Subgroup differences between follow-up cohorts			$\chi^e = 1.37$; $df = 2$; $p = 0.50$		
Duration of illness 2–5 years	<2 years	3 (4)	667–666	d = **0.39** [S] (0.24–0.55)[f]	+ = 1 (25.0%)/- = 0 (0.0%)	I^e = 0% (0–56%)
	≥2 - < 5 years	3 (4)	730–730	d = **0.49** [S] (0.41–0.58)	+ = 0 (0.0%)/- = 0 (0.0%)	I^e = 16% (0–40%)
	Subgroup differences between follow-up cohorts			$\chi^e = 1.28$; $df = 1$; $p = 0.26$		
Duration of illness 5–10 years	<2 years	2 (2)	421–421	d = −0.12 [O] (−0.25 – 0.01)[g]	+ = 0 (0.0%)/- = 0 (0.0%)	I^e = 0% (NA)
	≥2 - < 5 years	2 (2)	246–240	d = 0.52 [M] (−0.12 – 1.17)	+ = 1 (50.0%)/- = 0 (0.0%)	I^e = **92%** (NA)
	Subgroup differences between follow-up cohorts			$\chi^e = 3.65$; $df = 1$; $p = 0.06$		
Duration of illness >10 years	<2 years	4 (5)	1835–760	d = 0.37 [S] (0.07–0.66)[f]	+ = 0 (0.0%)/- = 0 (0.0%)	I^e = **86%** (69–94%)
	≥2 - < 5 years	3 (3)	1044–738	d = 0.45 [S] (−0.10 – 1.01)	+ = 1 (33.3%)/- = 0 (0.0%)	I^e = **91%** (71–97%)
	Subgroup differences between follow-up cohorts			$\chi^e = 0.07$; $df = 1$; $p = 0.78$		
Duration of illness unclear	<2 years	1 (2)	403–259	d = 0.03 [O] (−0.09 – 0.15)	+ = 0 (0.0%)/- = 0 (0.0%)	I^e = 0% (NA)
	Subgroup differences between follow-up cohorts			Not Applicable		

Abbreviations: CI = Confidence Interval; FU = Follow-up; L = Large effect; M = Medium effect; O=No effect; S = Small effect; N = number of patients; K = number of studies.

Notes.

[a] Outcomes in **bold** are significant ($p < 0.05$) after Benjamini-Hochberg correction; Outcomes underlined are no longer significant after Benjamini-Hochberg correction for multiple testing.

[b] N=No effect (d > -0.20 - <0.20); S = Small effect (d ≤ -0.20 and >-0.50 - ≥0.20 and < 0.50); M = Medium effect (d ≤ -0.50 and >-0.80 - ≥0.50 and < 0.80); L = Large effect (d < -0.80 - >0.80).

[c] + = improvement of outcome at follow-up; - = deterioration of outcome at follow-up.

[d] Significant subgroup differences with the duration of illness <2 years subgroup outcome within the same follow-up cohort.

[e] Significant subgroup differences with the duration of illness 2–5 years subgroup outcome within the same follow-up cohort.

[f] Significant subgroup differences with the duration of illness 5–10 years subgroup outcome within the same follow-up cohort.

[g] Significant subgroup differences with the duration of illness >10 years subgroup outcome within the same follow-up cohort.

3.3.1. Positive symptoms

We found a large overall improvement of positive symptoms ($d = 0.84$; $I^2 = 97\%$; $k = 74$). For subgroups with a baseline DOI <2 years we found a large improvement of positive symptoms up to 8 years of follow-up. For both subgroups with a baseline DOI 2–5 years and 5–10 years, there was a larger improvement of positive symptoms after a follow-up of less than 2 years compared with outcomes with longer follow-up lengths ($\chi^2 = 7.57$; $df = 2$; $p < 0.05$ and $\chi^2 = 20.77$; $df = 1$; $p < 0.01$). For the subgroup with a baseline DOI >10 years small to medium improvement of positive symptoms was found regardless of the follow-up length. Overall, subgroups with a baseline DOI <2 years showed larger improvement of positive symptoms after a short follow-up (i.e. <5 years) than subgroups with a larger baseline DOI.

3.3.2. Negative symptoms

We found a small overall improvement of negative symptoms ($d = 0.39$; $I^2 = 89\%$; $k = 74$). For subgroups with a baseline DOI <2 years and 2–5 years there was a medium improvement of negative symptoms after a follow-up shorter than 5 years and a large improvement of negative symptoms after a follow-up longer than 5 years. For both DOI subgroups we found a significantly larger improvement of negative symptoms for study outcomes with longer follow-up ($\chi^2 = 10.92$; $df = 3$; $p < 0.05$ and $\chi^2 = 10.17$; $df = 2$; $p < 0.01$). For the subgroup with a baseline DOI 5–10 years a medium improvement of negative symptoms was found after a follow-up shorter than 2 years. For the subgroup with a baseline DOI >10 years we found a small improvement of negative symptoms regardless of the follow-up length. Overall, both subgroups with a baseline DOI <2 years and 2–5 years showed larger improvement of negative symptoms than the subgroup with a baseline DOI >10 years.

3.3.3. Disorganization symptoms

We found a small overall improvement of disorganization symptoms ($d = 0.47$; $I^2 = 93\%$; $k = 14$). For subgroups with a baseline DOI <2 years there was a larger improvement of disorganization symptoms, regardless of the follow-up length, compared with the subgroups with a baseline DOI of more than 2 years. For subgroups with a baseline DOI 2–5 years we found a small improvement of disorganization symptoms and for the subgroups with baseline DOI 5–10 years and >10 years no improvement of disorganization symptoms was found.

3.3.4. Depressive symptoms

We found a small overall improvement of depressive symptoms ($d = 0.33$; $I^2 = 84\%$; $k = 20$). For the subgroups with a baseline DOI 2–5 years, a small improvement of depressive symptoms regardless of the follow-up length was found. There was no improvement of depressive symptoms in the remaining baseline DOI subgroups (i.e. <2 years, 5–10 years and >10 years). Also, we found no differences between baseline DOI subgroups or study outcomes with different follow-up lengths.

3.4. Outliers and publication bias

We found 21 positive and 19 negative outliers for positive symptoms, 25 positive and 18 negative outliers for negative symptoms, 0 positive and 1 negative outlier for disorganization symptoms and no outliers for depressive symptoms. The positive outliers exclusively influenced results of negative symptoms within the subgroup with a DOI <2 years.

We also found a positively skewed funnel plot for negative symptoms (see Supplementary Materials 7), which indicates publication bias for the negative symptoms outcome data. All other outcome domains showed no indications of publication bias.

The influence of positive outliers and the positively skewed funnel plot indicated a potential overestimation of the negative symptom outcomes, especially within the subgroup with a DOI <2 years.

3.5. Analysis of potential moderators of change in outcomes at baseline

Meta-regression outcomes and sensitivity analyses are presented in Supplementary Material 4 and Table 3.

3.5.1. Moderators of change in positive symptoms

Meta-regression showed that age at baseline, study samples with a diagnosis of schizophrenia, baseline overall symptoms, and baseline positive symptoms were significant moderators for changes in positive symptoms. Sensitivity analyses indicated that subgroups with a lower age at baseline, in which only a part of the study sample was diagnosed with schizophrenia, or with a high severity of overall and positive symptoms at baseline were associated with larger improvement in positive symptoms ($\chi^2 = 38.61$; $df = 1$; $p < 0.01$; $\chi^2 = 9.80$; $df = 1$; $p < 0.01$; $\chi^2 = 42.32$; $df = 1$; $p < 0.01$; $\chi^2 = 36.51$; $df = 1$; $p < 0.01$). The influence of age at baseline, or baseline severity of overall and positive symptoms on changes in positive symptoms also applied to outcomes in the subgroup with a baseline DOI <2 years. The influence of baseline severity of positive symptoms on changes in positive symptoms also applied to outcomes in the subgroup with a baseline DOI 5–10 years.

3.5.2. Moderators of change in negative symptoms

Meta-regression showed that age at baseline, baseline severity of depressive symptoms, overall symptoms and positive symptoms and baseline executive functioning were significant moderators for changes in negative symptoms. Overall, sensitivity analyses indicated that subgroups with a lower age at baseline, or a high baseline severity of depressive symptoms, overall symptoms, or positive symptoms were associated with greater improvement in negative symptoms ($\chi^2 = 22.39$; $df = 1$; $p < 0.01$; $\chi^2 = 10.30$; $df = 1$; $p < 0.01$; $\chi^2 = 7.22$; $df = 1$; $p < 0.01$; $\chi^2 = 10.42$; $df = 1$; $p < 0.01$). The influence of age at baseline and severity of both overall and positive symptoms on changes in negative symptoms also applied to outcomes in the subgroup with a baseline DOI <2 years. The influence of a high baseline severity of depressive symptoms and positive symptoms on changes in negative symptoms also applied to outcomes in the subgroup with a baseline DOI 5–10 years.

3.5.3. Moderators of change in symptoms of disorganization symptoms

Meta-regression showed that study samples with a schizophrenia diagnosis, or a high baseline severity of positive symptoms were significant moderators for changes in disorganization symptoms. Sensitivity analyses indicated that a high baseline severity of positive symptoms was associated with larger improvement in disorganization symptoms ($\chi^2 = 23.04$; $df = 1$; $p < 0.01$). This moderating effect also applied to outcomes in the subgroup with a baseline DOI 5–10 years. We found no moderating effects of a schizophrenia diagnosis on changes in disorganization symptoms in the sensitivity analysis.

3.5.4. Moderators of change in depressive symptoms

Meta-regression showed that baseline severity of positive symptoms was a significant moderator for changes in depressive symptoms. Sensitivity analyses indicated no overall moderating effect of baseline positive symptoms on changes in depressive symptoms. However, within the subgroup with a baseline DOI 5–10 years we found that a high baseline severity of positive symptoms was associated with greater improvement in depressive symptoms.

3.6. Quality assessment

We presented the quality assessment and the sensitivity analysis, in which we analyzed the influence of study quality on the outcomes, in Supplementary Materials 5 and 6. Quality assessments indicated that the quality of the included studies was considered good concerning patient recruitment and outcome assessment and analysis. A relatively larger

Table 3
Sensitivity analysis of significant moderators.

Positive symptoms						
(Sub)analysis		*All studies and outcomes*				
		K (studies (outcomes))	N (baseline-FU)	Effect size (95% CI)[a] and magnitude of effect[b]	K (%) large effect[b] [+/−][c]	Heterogeneity (I^2 (95% CI))[a]
Confounder	*Rating*					
Age at baseline	Young (below median)	35 (64)	3393–3061	d = **1.30** [L] (1.07–1.52)	+ = 36 (56.3%); - = 0 (0.0%)	$\mathbf{I^2 = 98\%}$ (97–98%)
	Old (above median)	35 (78)	6211–5696	d = **0.47** [S] (0.35–0.60)	+ = 11 (14.1%); - = 0 (0.0%)	$\mathbf{I^2 = 95\%}$ (94–95%)
	Subgroup differences between follow-up cohorts			$\chi^2 = 38.61$; $df = 1$; $p < 0.01$		
Diagnosis of schizophrenia	100% diagnosed with schizophrenia	37 (65)	6352–5740	d = **0.62** [M] (0.46–0.78)	+ = 18 (27.7%); - = 0 (0.0%)	$\mathbf{I^2 = 96\%}$ (96–97%)
	<100% diagnosed with schizophrenia	34 (77)	3342–3104	d = **1.02** [L] (0.83–1.21)	+ = 29 (37.7%); - = 0 (0.0%)	$\mathbf{I^2 = 97\%}$ (97–98%)
	Subgroup differences between follow-up cohorts			$\chi^2 = 9.80$; $df = 1$; $p < 0.01$		
Baseline overall symptom severity	High (above median)	23 (58)	2558–2403	d = **1.20** [L] (0.99–1.42)	+ = 28 (48.3%); - = 0 (0.0%)	$\mathbf{I^2 = 98\%}$ (97–98%)
	Low (below median)	21 (34)	2488–2070	d = **0.38** [S] (0.27–0.50)	+ = 4 (11.8%); - = 0 (0.0%)	$\mathbf{I^2 = 84\%}$ (79–87%)
	Subgroup differences between follow-up cohorts			$\chi^2 = 42.32$; $df = 1$; $p < 0.01$		
Baseline positive symptom severity	High (above median)	29 (55)	4360–4124	d = **1.35** [L] (1.11–1.58)	+ = 31 (56.4%); - = 0 5(0.0%)	$\mathbf{I^2 = 98\%}$ (97–98%)
	Low (below median)	38 (83)	5032–4388	d = **0.50** [M] (0.36–0.64)	+ = 14 (16.9%); - = 0 (0.0%)	$\mathbf{I^2 = 96\%}$ (96–97%)
	Subgroup differences between follow-up cohorts			$\chi^2 = 36.51$; $df = 1$; $p < 0.01$		
(Sub)analysis		***Duration of illness at baseline <2 years study outcomes***				
		K (studies (outcomes))	N (baseline-FU)	Effect size (95% CI)[a] and magnitude of effect[b]	K (%) large effect[b] [+/−][c]	Heterogeneity (I^2 (95% CI))[a]
Confounder	*Rating*					
Age at baseline	Young (below median)	23 (36)	2056–1781	d = **1.28** [L] (0.90–1.65)	+ = 19 (52.8%); - = 0 (0.0%)	$\mathbf{I^2 = 98\%}$ (97–98%)
	Old (above median)	1 (3)	93–93	d = 2.33 [L] (2.16–2.51)	+ = 3 (100.0%); - = 0 (0.0%)	I^2 = 0% (0–80%)
	Subgroup differences between follow-up cohorts			$\chi^2 = 25.0$; $df = 1$; $p < 0.01$		
Diagnosis of schizophrenia	100% diagnosed with schizophrenia	6 (14)	419–358	d = 0.95 [L] (0.43–1.47)	+ = 5 (35.7%); - = 0 (0.0%)	$\mathbf{I^2 = 97\%}$ (96–98%)
	<100% diagnosed with schizophrenia	18 (25)	1730–1516	d = **1.59** [L] (1.15–2.03)	+ = 17 (68.0%); - = 0 (0.0%)	$\mathbf{I^2 = 98\%}$ (97–98%)
	Subgroup differences between follow-up cohorts			$\chi^2 = 3.43$; $df = 1$; $p = 0.06$		
Baseline overall symptom severity	High (above median)	12 (19)	1085–1016	d = 1.81 [L] (1.32–2.30)	+ = 15 (78.9%); - = 0 (0.0%)	$\mathbf{I^2 = 97\%}$ (96–98%)
	Low (below median)	9 (11)	560–490	d = 0.59 [M] (0.18–1.01)	+ = 3 (27.3%); - = 0 (0.0%)	$\mathbf{I^2 = 93\%}$ (89–95%)
	Subgroup differences between follow-up cohorts			$\chi^2 = 13.76$; $df = 1$; $p < 0.05$		
Baseline positive symptom severity	High (above median)	12 (19)	968–858	d = **2.09** [L] (1.67–2.51)	+ = 16 (84.2%); - = 0 (0.0%)	$\mathbf{I^2 = 95\%}$ (94–96%)
	Low (below median)	12 (19)	1132–967	d = 0.58 [M] (0.10–1.05)	+ = 5 (26.3%); - = 0 (0.0%)	$\mathbf{I^2 = 98\%}$ (97–98%)
	Subgroup differences between follow-up cohorts			$\chi^2 = 21.59$; $df = 1$; $p < 0.05$		
(Sub)analysis		***Duration of illness at baseline 2–5 years study outcomes***				
		K (studies (outcomes))	N (baseline-FU)	Effect size (95% CI)[a] and magnitude of effect[b]	K (%) large effect[b] [+/−][c]	Heterogeneity (I^2 (95% CI))[a]
Confounder	*Rating*					
Age at baseline	Young (below median)	5 (7)	783–783	d = 1.42 [L] (0.67–2.16)	+ = 4 (57.1%); - = 0 (0.0%)	$\mathbf{I^2 = 98\%}$ (98–99%)
	Old (above median)	X	X	X	X	X
	Subgroup differences between follow-up cohorts			Not Applicable		
Diagnosis of schizophrenia	100% diagnosed with schizophrenia	X	X	X	X	X
	<100% diagnosed with schizophrenia	5 (7)	783–783	d = 1.42 [L] (0.67–2.16)	+ = 4 (57.1%); - = 0 (0.0%)	$\mathbf{I^2 = 98\%}$ (98–99%)
	Subgroup differences between follow-up cohorts			Not Applicable		
Baseline overall symptom severity	High (above median)	4 (5)	379–379	d = 1.74 [L] (0.47–3.01)	+ = 4 (80.0%); - = 0 (0.0%)	$\mathbf{I^2 = 99\%}$ (98–99%)
	Low (below median)	1 (2)	404–404	d = 0.62 [M] (0.52–0.73)	+ = 0 (0.0%); - = 0 (0.0%)	I^2 = 0% (NA)
	Subgroup differences between follow-up cohorts			$\chi^2 = 2.94$; $df = 1$; $p = 0.09$		
Baseline positive symptom severity	High (above median)	2 (2)	202–202	d = 0.62 [M] (−0.57 – 1.82)	+ = 1 (50.0%); - = 0 (0.0%)	$\mathbf{I^2 = 97\%}$ (NA)
	Low (below median)	3 (5)	581–581	d = 1.73 [L] (0.78–2.68)	+ = 3 (60.0%); - = 0 (0.0%)	$\mathbf{I^2 = 99\%}$ (98–99%)
	Subgroup differences between follow-up cohorts			$\chi^2 = 2.04$; $df = 1$; $p = 0.15$		
(Sub)analysis		***Duration of illness at baseline 5–10 years study outcomes***				
		K (studies (outcomes))	N (baseline-FU)	Effect size (95% CI)[a] and magnitude of effect[b]	K (%) large effect[b] [+/−][c]	Heterogeneity (I^2 (95% CI))[a]

(*continued on next page*)

Table 3 (*continued*)

Positive symptoms						
(Sub)analysis		*All studies and outcomes*				
		K (studies (outcomes))	N (baseline-FU)	Effect size (95% CI)[a] and magnitude of effect[b]	K (%) large effect[b] [+/−][c]	Heterogeneity (I^2 (95% CI))[a]
Confounder	*Rating*					
Age at baseline	Young (below median)	6 (12)	741–636	$d = \underline{0.92}$ [L] (0.52–0.73)	+ = 6 (50.0%); - = 0 (0.0%)	**I^2 = 95%** (94–97%)
	Old (above median)	6 (9)	808–745	$d = \underline{0.95}$ [L] (0.02–1.88)	+ = 4 (44.4%); - = 0 (0.0%)	**I^2 = 98%** (98–99%)
	Subgroup differences between follow-up cohorts			$\chi^2 = 0.00$; $df = 1$; $p = 0.96$		
Diagnosis of schizophrenia	100% diagnosed with schizophrenia	8 (15)	1364–1225	$d = \underline{1.07}$ [L] (0.61–1.54)	+ = 9 (60.0%); - = 0 (0.0%)	**I^2 = 98%** (97–98%)
	<100% diagnosed with schizophrenia	4 (6)	185–159	$d = \underline{0.60}$ [M] (0.02–1.18)	+ = 1 (16.7%); - = 0 (0.0%)	**I^2 = 91%** (83–95%)
	Subgroup differences between follow-up cohorts			$\chi^2 = 1.56$; $df = 1$; $p = 0.21$		
Baseline overall symptom severity	High (above median)	3 (7)	292–270	$d =$ **1.38** [L] (0.53–2.22)	+ = 4 (57.1%); - = 0 (0.0%)	**I^2 = 98%** (97–99%)
	Low (below median)	1 (2)	47–47	$d = -0.03$ [O] (−0.38 – 0.32)	+ = 0 (0.0%); - = 0 (0.0%)	I^2 = 0% (NA)
	Subgroup differences between follow-up cohorts			$\chi^2 = 9.03$; $df = 1$; $p = 0.07$		
Baseline positive symptom severity	High (above median)	3 (8)	431–424	$d = \underline{1.75}$ [L] (1.44–2.06)	+ = 8 (100.0%); - = 0 (0.0%)	**I^2 = 91%** (84–94%)
	Low (below median)	7 (11)	912–816	$d = \underline{0.37}$ [S] (0.02–0.72)	+ = 1 (9.1%); - = 0 (0.0%)	**I^2 = 94%** (91–96%)
	Subgroup differences between follow-up cohorts			$\chi^2 = 33.94$; $df = 1$; $p < 0.05$		
(Sub)analysis		***Duration of illness at baseline >10 years study outcomes***				
		K (studies (outcomes))	N (baseline-FU)	Effect size (95% CI)[a] and magnitude of effect[b]	K (%) large effect[b] [+/−][c]	Heterogeneity (I^2 (95% CI))[a]
Confounder	*Rating*					
Age at baseline	Young (below median)	X	X	X	X	X
	Old (above median)	23 (55)	4472–4058	$d =$ **0.36** [S] (0.29–0.43)	+ = 3 (5.5%); - = 0 (0.0%)	**I^2 = 77%** (73–81%)
	Subgroup differences between follow-up cohorts			Not Applicable		
Diagnosis of schizophrenia	100% diagnosed with schizophrenia	18 (27)	3847–3441	$d =$ **0.39** [S] (0.26–0.52)	+ = 3 (11.1%); - = 0 (0.0%)	**I^2 = 84%** (80–88%)
	<100% diagnosed with schizophrenia	5 (28)	655–644	$d =$ **0.34** [S] (0.27–0.41)	+ = 0 (0.0%); - = 0 (0.0%)	**I^2 = 57%** (47–65%)
	Subgroup differences between follow-up cohorts			$\chi^2 = 0.48$; $df = 1$; $p = 0.49$		
Baseline overall symptom severity	High (above median)	3 (22)	598–598	$d = \underline{0.41}$ [S] (0.32–0.50)	+ = 1 (4.5%); - = 0 (0.0%)	**I^2 = 76%** (67–82%)
	Low (below median)	8 (15)	1336–1020	$d = \underline{0.29}$ [S] (0.20–0.39)	+ = 1 (6.7%); - = 0 (0.0%)	I^2 = 34% (19–47%)
	Subgroup differences between follow-up cohorts			$\chi^2 = 3.19$; $df = 1$; $p = 0.07$		
Baseline positive symptom severity	High (above median)	9 (19)	2457–2385	$d = \underline{0.43}$ [S] (0.31–0.55)	+ = 2 (10.5%); - = 0 (0.0%)	**I^2 = 81%** (74–87%)
	Low (below median)	12 (35)	1948–1606	$d =$ **0.31** [S] (0.22–0.40)	+ = 1 (2.9%); - = 0 (0.0%)	**I^2 = 74%** (68–79%)
	Subgroup differences between follow-up cohorts			$\chi^2 = 2.25$; $df = 1$; $p = 0.13$		
Negative symptoms						
(Sub)analysis		***All studies and outcomes***				
		K (studies (outcomes))	N (baseline-FU)	Effect size (95% CI)[a] and magnitude of effect[b]	K (%) large effect[b] [+/−][c]	Heterogeneity (I^2 (95% CI))[a]
Confounder	*Rating*					
Age at baseline	Young (below median)	38 (68)	3694–3123	$d =$ **0.54** [M] (0.46–0.63)	+ = 17 (25.0%); - = 0 (0.0%)	**I^2 = 83%** (80–85%)
	Old (above median)	34 (79)	8398–6813	$d =$ **0.28** [S] (0.21–0.35)	+ = 9 (11.4%); - = 0 (0.0%)	**I^2 = 89%** (87–90%)
	Subgroup differences between follow-up cohorts			$\chi^2 = 22.39$; $df = 1$; $p < 0.01$		
Baseline severity of depressive symptoms	High (above median)	13 (20)	3801–2496	$d =$ **0.54** [M] (0.34–0.75)	+ = 6 (30.0%); - = 0 (0.0%)	**I^2 = 94%** (92–95%)
	Low (below median)	13 (24)	3695–3305	$d =$ **0.19** [O] (0.11–0.26)	+ = 1 (4.2%); - = 0 (0.0%)	**I^2 = 71%** (62–78%)
	Subgroup differences between follow-up cohorts			$\chi^2 = 10.30$; $df = 1$; $p < 0.05$		
Baseline executive functioning	High (above median)	7 (9)	264–261	$d = \underline{0.60}$ [M] (0.20–1.01)	+ = 4 (44.4%); - = 0 (0.0%)	**I^2 = 83%** (70–90%)
	Low (below median)	7 (12)	1553–1185	$d =$ **0.37** [S] (0.21–0.53)	+ = 2 (16.7%); - = 0 (0.0%)	**I^2 = 69%** (52–79%)
	Subgroup differences between follow-up cohorts			$\chi^2 = 1.09$; $df = 1$; $p = 0.30$		
Baseline overall symptoms severity	High (above median)	18 (51)	2044–1888	$d =$ **0.54** [M] (0.45–0.64)	+ = 14 (27.5%); - = 0 (0.0%)	**I^2 = 87%** (84–89%)
	Low (below median)	25 (42)	2742–2182	$d =$ **0.38** [S] (0.29–0.46)	+ = 5 (11.9%); - = 0 (0.0%)	**I^2 = 69%** (62–74%)
	Subgroup differences between follow-up cohorts			$\chi^2 = 7.22$; $df = 1$; $p < 0.05$		
Baseline positive symptoms severity	High (above median)	31 (61)	4602–4300	$d =$ **0.52** [M] (0.41–0.62)	+ = 17 (27.9%); - = 0 (0.0%)	**I^2 = 93%** (91–94%)
	Low (below median)	36 (79)	4737–4053	$d =$ **0.32** [S] (0.26–0.38)	+ = 7 (8.9%); - = 0 (0.0%)	**I^2 = 80%** (77–83%)
	Subgroup differences between follow-up cohorts			$\chi^2 = 10.42$; $df = 1$; $p < 0.01$		

(*continued on next page*)

Table 3 (*continued*)

Positive symptoms

(Sub)analysis		*All studies and outcomes*				
		K (studies (outcomes))	N (baseline-FU)	Effect size (95% CI)[a] and magnitude of effect[b]	K (%) large effect[b] [+/−][c]	Heterogeneity (I^2 (95% CI))[a]
(Sub)analysis		***Duration of illness at baseline <2 years study outcomes***				
		K (studies (outcomes))	N (baseline-FU)	Effect size (95% CI)[a] and magnitude of effect[b]	K (%) large effect[b] [+/−][c]	Heterogeneity (I^2 (95% CI))[a]
Confounder	*Rating*					
Age at baseline	Young (below median)	26 (41)	2297–1956	d = **0.54** [M] (0.39–0.69)	+ = 11 (26.8%); - = 0 (0.0%)	**I^2 = 87%** (84–89%)
	Old (above median)	1 (4)	93–93	d = 0.21 [S] (0.05–0.38)	+ = 0 (0.0%); - = 0 (0.0%)	I^2 = 60% (0–84%)
	Subgroup differences between follow-up cohorts			χ^2 = 8.04; df = 1; p < 0.05		
Baseline severity of depressive symptoms	High (above median)	2 (5)	238–186	d = 0.78 [M] (0.33–1.22)	+ = 3 (60.0%); - = 0 (0.0%)	**I^2 = 84%** (63–93%)
	Low (below median)	4 (7)	494–479	d = 0.22 [S] (0.08–0.36)	+ = 0 (0.0%); - = 0 (0.0%)	I^2 = 47% (14–67%)
	Subgroup differences between follow-up cohorts			χ^2 = 5.49; df = 1; p = 0.13		
Baseline executive functioning	High (above median)	2 (2)	85–85	d = 1.30 [L] (0.72–1.88)	+ = 2 (100.0%); - = 0 (0.0%)	I^2 = 59% (NA)
	Low (below median)	3 (4)	421–421	d = 0.32 [S] (0.07–0.56)	+ = 1 (25.0%); - = 0 (0.0%)	I^2 = 39% (0–70%)
	Subgroup differences between follow-up cohorts			χ^2 = 9.44; df = 1; p = 0.15		
Baseline overall symptoms severity	High (above median)	10 (16)	876–791	d = 0.85 [L] (0.62–1.09)	+ = 8 (50.0%); - = 0 (0.0%)	**I^2 = 81%** (73–87%)
	Low (below median)	10 (17)	627–557	d = 0.33 [S] (0.20–0.46)	+ = 2 (11.8%); - = 0 (0.0%)	**I^2 = 68%** (56–77%)
	Subgroup differences between follow-up cohorts			χ^2 = 14.20; df = 1; p < 0.01		
Baseline positive symptoms severity	High (above median)	14 (25)	1210–1034	d = **0.69** [M] (0.49–0.89)	+ = 9 (36.0%); - = 0 (0.0%)	**I^2 = 89%** (86–92%)
	Low (below median)	11 (18)	827–662	d = 0.26 [S] (0.14–0.39)	+ = 1 (5.6%); - = 0 (0.0%)	**I^2 = 69%** (58–78%)
	Subgroup differences between follow-up cohorts			χ^2 = 12.26; df = 1; p < 0.05		
(Sub)analysis		***Duration of illness at baseline 2–5 years study outcomes***				
		K (studies (outcomes))	N (baseline-FU)	Effect size (95% CI)[a] and magnitude of effect[b]	K (%) large effect[b] [+/−][c]	Heterogeneity (I^2 (95% CI))[a]
Confounder	*Rating*					
Age at baseline	Young (below median)	5 (7)	783–610	d = **0.61** [M] (0.47–0.76)	+ = 2 (28.6%); - = 0 (0.0%)	I^2 = 55% (22–74%)
	Old (above median)	X	X	X	X	X
	Subgroup differences between follow-up cohorts			Not Applicable		
Baseline severity of depressive symptoms	High (above median)	4 (4)	671–498	d = 0.68 [M] (0.42–0.94)	+ = 2 (50.0%); - = 0 (0.0%)	**I^2 = 74%** (27–90%)
	Low (below median)	1 (3)	112–112	d = 0.55 [M] (0.41–0.69)	+ = 0 (0.0%); - = 0 (0.0%)	I^2 = 0% (0–96%)
	Subgroup differences between follow-up cohorts			χ^2 = 0.74; df = 1; p = 0.39		
Baseline executive functioning	High (above median)	X	X	X	X	X
	Low (below median)	1 (3)	112–112	d = 0.55 [M] (0.41–0.69)	+ = 0 (0.0%); - = 0 (0.0%)	I^2 = 0% (0–96%)
	Subgroup differences between follow-up cohorts			Not Applicable		
Baseline overall symptoms severity	High (above median)	2 (4)	244–244	d = 0.51 [M] (0.39–0.63)	+ = 0 (0.0%); - = 0 (0.0%)	I^2 = 0% (0–86%)
	Low (below median)	3 (3)	539–366	d = 0.79 [M] (0.44–1.13)	+ = 2 (66.7%); - = 0 (0.0%)	**I^2 = 76%** (8–94%)
	Subgroup differences between follow-up cohorts			χ^2 = 2.15; df = 1; p = 0.14		
Baseline positive symptoms severity	High (above median)	2 (2)	202–202	d = 0.63 [M] (0.16–1.10)	+ = 1 (50.0%); - = 0 (0.0%)	**I^2 = 80%** (NA)
	Low (below median)	3 (5)	581–408	d = 0.61 [M] (0.45–0.77)	+ = 1 (20.0%); - = 0 (0.0%)	I^2 = 53% (4–76%)
	Subgroup differences between follow-up cohorts			χ^2 = 0.01; df = 1; p = 0.94		
(Sub)analysis		***Duration of illness at baseline 5–10 years study outcomes***				
		K (studies (outcomes))	N (baseline-FU)	Effect size (95% CI)[a] and magnitude of effect[b]	K (%) large effect[b] [+/−][c]	Heterogeneity (I^2 (95% CI))[a]
Confounder	*Rating*					
Age at baseline	Young (below median)	6 (11)	741–636	d = **0.45** [S] (0.32–0.58)	+ = 1 (9.09%); - = 0 (0.0%)	**I^2 = 76%** (61–85%)
	Old (above median)	5 (8)	755–698	d = 0.56 [M] (0.14–0.98)	+ = 5 (62.50%); - = 0 (0.0%)	**I^2 = 93%** (88–96%)
	Subgroup differences between follow-up cohorts			χ^2 = 0.23; df = 1; p = 0.63		
Baseline severity of depressive symptoms	High (above median)	2 (2)	109–103	d = 0.67 [M] (0.40–0.95)	+ = 1 (50.0%); - = 0 (0.0%)	I^2 = 0% (NA)
	Low (below median)	2 (3)	421–421	d = 0.09 [O] (−0.05 – 0.22)	+ = 0 (0.0%); - = 0 (0.0%)	I^2 = 0% (0–96%)
	Subgroup differences between follow-up cohorts			χ^2 = 13.86; df = 1; p < 0.01		
Baseline executive functioning	High (above median)	1 (1)	11–11	d = 0.97 [L] (0.08–1.86)	+ = 1 (100.0%); - = 0 (0.0%)	Not Applicable
	Low (below median)	2 (3)	86–86	d = 0.20 [S] (−0.50 – 0.89)	+ = 1 (33.3%); - = 0 (0.0%)	**I^2 = 82%** (31–95%)
	Subgroup differences between follow-up cohorts			χ^2 = 1.80; df = 1; p = 0.18		
Baseline overall symptoms severity	High (above median)	2 (3)	122–98	d = 0.31 [S] (−0.32 – 0.93)	+ = 1 (33.3%); - = 0 (0.0%)	**I^2 = 88%** (58–96%)

(*continued on next page*)

Table 3 (*continued*)

Positive symptoms

(Sub)analysis		*All studies and outcomes*				
		K (studies (outcomes))	N (baseline-FU)	Effect size (95% CI)[a] and magnitude of effect[b]	K (%) large effect[b] [+/−][c]	Heterogeneity (I^2 (95% CI))[a]
	Low (below median)	1 (2)	47–47	$d = -0.10$ [O] (−0.45 – 0.26)	+ = 0 (0.0%); - = 0 (0.0%)	I^2 = 0% (NA)
	Subgroup differences between follow-up cohorts			$\chi^2 = 1.23$; $df = 1$; $p = 0.27$		
Baseline positive symptoms severity	High (above median)	3 (8)	431–424	$d = \underline{0.70}$ [M] (0.50–0.89)	+ = 3 (37.5%); - = 0 (0.0%)	**I^2 = 81%** (67–89%)
	Low (below median)	7 (10)	912–813	$d = \underline{0.29}$ [S] (0.07–0.50)	+ = 2 (20.0%); - = 0 (0.0%)	**I^2 = 88%** (81–93%)
	Subgroup differences between follow-up cohorts			$\chi^2 = 7.64$; $df = 1$; $p < 0.01$		
(Sub)analysis		***Duration of illness at baseline >10 years study outcomes***				
		K (studies (outcomes))	N (baseline-FU)	Effect size (95% CI)[a] and magnitude of effect[b]	K (%) large effect[b] [+/−][c]	Heterogeneity (I^2 (95% CI))[a]
Confounder	*Rating*					
Age at baseline	Young (below median)	X	X	X	X	X
	Old (above median)	23 (57)	6682–5210	$d =$ **0.26** [S] (0.20–0.33)	+ = 2 (3.5%); - = 0 (0.0%)	**I^2 = 84%** (82–87%)
	Subgroup differences between follow-up cohorts			Not Applicable		
Baseline severity of depressive symptoms	High (above median)	3 (5)	2359–1285	$d = 0.14$ [O] (−0.14 – 0.41)	+ = 0 (0.0%); - = 0 (0.0%)	**I^2 = 84%** (63–93%)
	Low (below median)	7 (11)	2780–2405	$d = \underline{0.08}$ [O] (0.02–0.13)	+ = 0 (0.0%); - = 0 (0.0%)	I^2 = 38% (18–53%)
	Subgroup differences between follow-up cohorts			$\chi^2 = 0.18$; $df = 1$; $p = 0.67$		
Baseline executive functioning	High (above median)	3 (4)	137–134	$d = 0.08$ [O] (−0.33 – 0.50)	+ = 0 (0.0%); - = 0 (0.0%)	**I^2 = 66%** (9–87%)
	Low (below median)	2 (2)	1001–683	$d = \underline{0.22}$ [S] (0.01–0.43)	+ = 0 (0.0%); - = 0 (0.0%)	I^2 = 42% (NA)
	Subgroup differences between follow-up cohorts			$\chi^2 = 0.32$; $df = 1$; $p = 0.57$		
Baseline overall symptoms severity	High (above median)	3 (22)	598–598	$d =$ **0.34** [S] (0.27–0.41)	+ = 1 (4.5%); - = 0 (0.0%)	**I^2 = 74%** (66–81%)
	Low (below median)	9 (16)	1386–1069	$d = \underline{0.38}$ [S] (0.27–0.50)	+ = 0 (0.0%); - = 0 (0.0%)	**I^2 = 56%** (35–61%)
	Subgroup differences between follow-up cohorts			$\chi^2 = 0.37$; $df = 1$; $p = 0.54$		
Baseline positive symptoms severity	High (above median)	9 (19)	2457–2385	$d = \underline{0.29}$ [S] (0.15–0.42)	+ = 2 (10.5%); - = 0 (0.0%)	**I^2 = 92%** (90–94%)
	Low (below median)	11 (34)	1911–1575	$d =$ **0.27** [S] (0.22–0.33)	+ = 0 (0.0%); - = 0 (0.0%)	**I^2 = 52%** (43–59%)
	Subgroup differences between follow-up cohorts			$\chi^2 = 0.03$; $df = 1$; $p = 0.87$		
Disorganization symptoms						
(Sub)analysis		***All studies and outcomes***				
		K (studies (outcomes))	N (baseline-FU)	Effect size (95% CI)[a] and magnitude of effect[b]	K (%) large effect[b] [+/−][c]	Heterogeneity (I^2 (95% CI))[a]
Confounder	*Rating*					
Diagnosis of schizophrenia	100% diagnosed with schizophrenia	4 (7)	1513–1210	$d = \underline{0.41}$ [S] (0.14–0.68)	+ = 0 (0.0%); - = 0 (0.0%)	**I^2 = 93%** (88–96%)
	<100% diagnosed with schizophrenia	10 (30)	1395–1108	$d =$ **0.48** [S] (0.30–0.67)	+ = 7 (23.3%); - = 0 (0.0%)	**I^2 = 93%** (91–94%)
	Subgroup differences between follow-up cohorts			$\chi^2 = 0.19$; $df = 1$; $p = 0.66$		
Baseline positive symptoms severity	High (above median)	5 (9)	555–466	$d =$ **1.01** [L] (0.76–1.27)	+ = 5 (55.6%); - = 0 (0.0%)	**I^2 = 87%** (78–92%)
	Low (below median)	9 (28)	2353–1852	$d =$ **0.29** [S] (0.14–0.44)	+ = 2 (7.1%); - = 0 (0.0%)	**I^2 = 91%** (88–93%)
	Subgroup differences between follow-up cohorts			$\chi^2 = 23.04$; $df = 1$; $p < 0.01$		
(Sub)analysis		***Duration of illness at baseline <2 years study outcomes***				
		K (studies (outcomes))	N (baseline-FU)	Effect size (95% CI)[a] and magnitude of effect[b]	K (%) large effect[b] [+/−][c]	Heterogeneity (I^2 (95% CI))[a]
Confounder	*Rating*					
Diagnosis of schizophrenia	100% diagnosed with schizophrenia	X	X	X	X	X
	<100% diagnosed with schizophrenia	4 (4)	596–554	$d = \underline{1.53}$ [L] (1.09–1.97)	+ = 4 (100.0%); - = 0 (0.0%)	**I^2 = 90%** (74–96%)
	Subgroup differences between follow-up cohorts			Not Applicable		
Baseline positive symptoms severity	High (above median)	3 (3)	289–247	$d = \underline{1.45}$ [L] (0.83–2.06)	+ = 3 (100.0%); - = 0 (0.0%)	**I^2 = 90%** (68–97%)
	Low (below median)	1 (1)	307–307	$d = \underline{1.75}$ [L] (1.57–1.93)	+ = 1 (100.0%); - = 0 (0.0%)	Not Applicable
	Subgroup differences between follow-up cohorts			$\chi^2 = 0.86$; $df = 1$; $p = 0.35$		
(Sub)analysis		***Duration of illness at baseline 2–5 years study outcomes***				
		K (studies (outcomes))	N (baseline-FU)	Effect size (95% CI)[a] and magnitude of effect[b]	K (%) large effect[b] [+/−][c]	Heterogeneity (I^2 (95% CI))[a]
Confounder	*Rating*					
Diagnosis of schizophrenia	100% diagnosed with schizophrenia	X	X	X	X	X

(*continued on next page*)

Table 3 (*continued*)

Positive symptoms						
(Sub)analysis		*All studies and outcomes*				
		K (studies (outcomes))	N (baseline-FU)	Effect size (95% CI)[a] and magnitude of effect[b]	K (%) large effect[b] [+/−][c]	Heterogeneity (I^2 (95% CI))[a]
	<100% diagnosed with schizophrenia	1 (1)	404–231	*d* = 0.38 [S] (0.22–0.54)	+ = 0 (0.0%); - = 0 (0.0%)	Not Applicable
	Subgroup differences between follow-up cohorts			Not Applicable		
Baseline positive symptoms severity	High (above median)	X	X	X	X	X
	Low (below median)	1 (1)	404–231	*d* = 0.38 [S] (0.22–0.54)	+ = 0 (0.0%); - = 0 (0.0%)	Not Applicable
	Subgroup differences between follow-up cohorts			Not Applicable		
(Sub)analysis		***Duration of illness at baseline 5–10 years study outcomes***				
		K (studies (outcomes))	N (baseline-FU)	Effect size (95% CI)[a] and magnitude of effect[b]	K (%) large effect[b] [+/−][c]	Heterogeneity (I^2 (95% CI))[a]
Confounder	*Rating*					
Diagnosis of schizophrenia	100% diagnosed with schizophrenia	3 (6)	592–592	*d* = 0.46 [S] (0.12–0.80)	+ = 0 (0.0%); - = 0 (0.0%)	**I^2 = 93%** (87–96%)
	<100% diagnosed with schizophrenia	2 (4)	94–73	*d* = 0.36 [S] (0.05–0.67)	+ = 1 (25.0%); - = 0 (0.0%)	I^2 = 56% (0–82%)
	Subgroup differences between follow-up cohorts			$\chi^2 = 0.17$; $df = 1$; $p = 0.68$		
Baseline positive symptoms severity	High (above median)	1 (4)	163–163	*d* = 0.70 [M] (0.58–0.81)	+ = 0 (0.0%); - = 0 (0.0%)	I^2 = 4% (0–11%)
	Low (below median)	4 (6)	523–499	*d* = 0.20 [S] (−0.04 – 0.44)	+ = 1 (16.7%); - = 0 (0.0%)	**I^2 = 69%** (38–85%)
	Subgroup differences between follow-up cohorts			$\chi^2 = 13.49$; $df = 1$; $p < 0.01$		
(Sub)analysis		***Duration of illness at baseline >10 years study outcomes***				
		K (studies (outcomes))	N (baseline-FU)	Effect size (95% CI)[a] and magnitude of effect[b]	K (%) large effect[b] [+/−][c]	Heterogeneity (I^2 (95% CI))[a]
Confounder	*Rating*					
Diagnosis of schizophrenia	100% diagnosed with schizophrenia	1 (1)	921–618	*d* = **0.14** [O] (0.04–0.24)	+ = 0 (0.0%); - = 0 (0.0%)	Not Applicable
	<100% diagnosed with schizophrenia	2 (19)	198–197	*d* = **0.22** [S] (0.16–0.27)	+ = 0 (0.0%); - = 0 (0.0%)	I^2 = 0% (0–5%)
	Subgroup differences between follow-up cohorts			$\chi^2 = 1.61$; $df = 1$; $p = 0.20$		
Baseline positive symptoms severity	High (above median)	X	X	X	X	X
	Low (below median)	3 (20)	1119–815	*d* = **0.20** [S] (0.15–0.25)	+ = 0 (0.0%); - = 0 (0.0%)	I^2 = 0% (0–5%)
	Subgroup differences between follow-up cohorts			Not Applicable		
Depressive symptoms						
(Sub)analysis		***All studies and outcomes***				
		K (studies (outcomes))	N (baseline-FU)	Effect size (95% CI)[a] and magnitude of effect[b]	K (%) large effect[b] [+/−][c]	Heterogeneity (I^2 (95% CI))[a]
Confounder	*Rating*					
Baseline positive symptoms severity	High (above median)	7 (13)	907–847	*d* = 0.42 [S] (0.23–0.61)	+ = 3 (23.1%); - = 0 (0.0%)	I^2 = 41% (23–55%)
	Low (below median)	11 (14)	3858–2462	*d* = **0.35** [S] (0.19–0.51)	+ = 1 (7.1%); - = 0 (0.0%)	**I^2 = 90%** (85–93%)
	Subgroup differences between follow-up cohorts			$\chi^2 = 0.28$; $df = 1$; $p = 0.60$		
(Sub)analysis		***Duration of illness at baseline <2 years study outcomes***				
		K (studies (outcomes))	N (baseline-FU)	Effect size (95% CI)[a] and magnitude of effect[b]	K (%) large effect[b] [+/−][c]	Heterogeneity (I^2 (95% CI))[a]
Confounder	*Rating*					
Baseline positive symptoms severity	High (above median)	3 (7)	267–214	*d* = 0.10 [O] (−0.19 – 0.38)	+ = 1 (14.3%); - = 0 (0.0%)	I^2 = 0% (0–43%)
	Low (below median)	1 (1)	29–14	*d* = 0.31 [S] (−0.33 – 0.95)	+ = 0 (0.0%); - = 0 (0.0%)	Not Applicable
	Subgroup differences between follow-up cohorts			$\chi^2 = 0.35$; $df = 1$; $p = 0.55$		
(Sub)analysis		***Duration of illness at baseline 2–5 years study outcomes***				
		K (studies (outcomes))	N (baseline-FU)	Effect size (95% CI)[a] and magnitude of effect[b]	K (%) large effect[b] [+/−][c]	Heterogeneity (I^2 (95% CI))[a]
Confounder	*Rating*					
Baseline positive symptoms severity	High (above median)	2 (3)	517–516	*d* = 0.47 [S] (0.26–0.68)	+ = 1 (33.3%); - = 0 (0.0%)	I^2 = 32% (0–98%)
	Low (below median)	2 (5)	554–554	*d* = **0.47** [S] (0.39–0.54)	+ = 0 (0.0%); - = 0 (0.0%)	I^2 = 0% (0–71%)
	Subgroup differences between follow-up cohorts			$\chi^2 = 0.0$; $df = 1$; $p = 0.97$		

(*continued on next page*)

Table 3 (*continued*)

Positive symptoms						
(Sub)analysis	*All studies and outcomes*					
	K (studies (outcomes))	N (baseline-FU)	Effect size (95% CI)[a] and magnitude of effect[b]	K (%) large effect[b] [+/−][c]	Heterogeneity (I^2 (95% CI))[a]
(Sub)analysis	***Duration of illness at baseline 5–10 years study outcomes***				
	K (studies (outcomes))	N (baseline-FU)	Effect size (95% CI)[a] and magnitude of effect[b]	K (%) large effect[b] [+/−][c]	Heterogeneity (I^2 (95% CI))[a]
Confounder	*Rating*				
Baseline positive symptoms severity	High (above median)	1 (1)	98–92	d = <u>0.86</u> [L] (0.57–1.15)	+ = 1 (100.0%); - = 0 (0.0%)	Not Applicable
	Low (below median)	2 (2)	421–421	d = -0.12 [O] (−0.25 – 0.01)	+ = 0 (0.0%); - = 0 (0.0%)	I^2 = 0% (NA)
	Subgroup differences between follow-up cohorts		χ^2 = 35.69; df = 1; p < 0.01		
(Sub)analysis	***Duration of illness at baseline >10 years study outcomes***				
	K (studies (outcomes))	N (baseline-FU)	Effect size (95% CI)[a] and magnitude of effect[b]	K (%) large effect[b] [+/−][c]	Heterogeneity (I^2 (95% CI))[a]
Confounder	*Rating*				
Baseline positive symptoms severity	High (above median)	1 (2)	25–25	d = 0.38 [S] (−0.02 – 0.77)	+ = 0 (0.0%); - = 0 (0.0%)	I^2 = 0% (NA)
	Low (below median)	6 (6)	2854–1473	d = <u>0.40</u> [S] (0.14–0.67)	+ = 1 (16.7%); - = 0 (0.0%)	I^2 = 92% (85–96%)
	Subgroup differences between follow-up cohorts		χ^2 = 0.01; df = 1; p = 0.92		

Abbreviations: CI = Confidence Interval; FU = Follow-up; L = Large effect; M = Medium effect; O═No effect; S = Small effect; N = number of patients; K = number of studies.

Notes.

[a] Outcomes in **bold** are significant (p < 0.05) after Benjamini-Hochberg correction; Outcomes <u>underlined</u> are no longer significant after Benjamini-Hochberg correction for multiple testing.

[b] N═No effect (d > -0.20 - <0.20); S = Small effect (d ≤ -0.20 and >-0.50 - ≥0.20 and < 0.50); M = Medium effect (d ≤ -0.50 and >-0.80 - ≥0.50 and < 0.80); L = Large effect (d < -0.80 - >0.80).

[c] + = improvement of outcome at follow-up; - = deterioration of outcome at follow-up.

number of studies reported high risk of bias regarding the other QUIPS items study attrition (i.e. handling missing data), study confounding and prognostic factor measurement (i.e. handling confounders and predictors in the analysis).

The sensitivity analyses showed that a low risk of bias on study participation was associated with larger improvement of both positive and negative symptoms. Furthermore, a high risk of bias on study attrition was associated with larger improvement of positive, negative, and disorganization symptoms.

4. Discussion

The current meta-analysis assessed changes in symptom dimensions in patients with schizophrenia spectrum disorders, with different durations of illness and length of follow-up. Overall, we found a large improvement of positive symptoms and a small improvement of negative, disorganization, and depressive symptoms. Subgroups with a baseline duration of illness (DOI) up until 10 years had the most substantial improvement of positive symptoms, specifically after short follow-up. In addition, the largest improvement of disorganization symptoms was found for patients with a DOI of less than 2 years, regardless of the follow-up length. There were no differences between different DOI subgroups in changes of both negative symptoms and depressive symptoms. For negative symptoms, we found that the most substantial improvement of outcomes takes place after a long follow-up, which is in contrast with the results for positive symptoms.

4.1. The influence of duration of illness on changes in symptoms

Taken together, outcomes of positive and disorganization symptoms showed relatively similar patterns of improvement over time, with the largest improvement for patients earlier in the course of illness. Of note, improvement of positive symptoms is larger and is also observed after a longer DOI (i.e. until 10 years after onset of schizophrenia spectrum disorder) compared with improvement patterns for disorganization symptoms, which show small or no improvement of outcomes in subgroups with a baseline DOI longer than 2 years. These findings could be explained by the fact that the baseline severity of positive symptoms was higher, and the duration of untreated psychosis was shorter earlier in the course of illness (see Supplementary Materials 3). This might indicate that participants are in a more acute phase during the first years of psychosis. As a consequence, treatment is started more quickly for subgroups with a shorter DOI and the focus might be prioritized on exacerbation of positive and disorganized symptoms. These findings are also in line with the idea that the first 5 years after onset of a schizophrenia spectrum disorder could be labeled as a ''critical period of recovery'' (Birchwood et al., 1998), and that this period may even be extended up to 10 years for positive symptoms.

In contrast to our findings for positive and disorganization symptoms, we did not find any differences between DOI subgroups for both depressive and negative symptoms. This is in line with previous studies that indicated that positive symptoms improve, but also relapse, faster than negative symptoms, and that both depressive and negative symptoms often remain more stable, but more persistent with less improvement over time (Salazar de Pablo et al., 2021; Arndt et al., 1995; Häfner, 2015; Haro et al., 2018; Ventura et al., 2015). The positive publication bias that was found for negative symptom outcomes in our included studies indicates that the improvement in negative symptoms is probably overestimated and thus may be even less favorable. This suggests that the focus on improvement of negative symptoms in included studies might explain this positive publication bias. However, we did not find any moderating effects of specific types of treatment or antipsychotic use that could explain this finding. The relatively small improvement of depressive symptoms that we found throughout the course of schizophrenia spectrum disorder is in contrast with previous findings (Salazar de Pablo et al., 2021). Both the small improvement of negative and

depressive symptoms and the absence of differences between DOI subgroups could not be contributed by a selective study sample at baseline. Supplementary materials 3 showed that the standardized scores of severity of both negative and depressive symptoms, based on normative data, were around average at baseline for people with schizophrenia spectrum disorders, and baseline symptom severity did not differ between the DOI subgroups. Our meta-analysis confirms that nowadays we still did not find sufficient manners to improve negative and depressive symptoms for this target group. Previous findings indicated that both negative and depressive symptoms are an important barrier for the improvement of other outcome domains as well, such as personal recovery (Best et al., 2020; Van Eck et al., 2018), and social functioning (De Winter et al., 2021). This underscores the clinical importance to focus on ways to address these symptom dimensions, and its interrelationships with other recovery domains, in future research and clinical practice.

It is important to address that we could not provide results on motor or cognitive symptoms over time, because these outcomes were reported by less than 10 studies. Previous research already indicated that both motor and cognitive symptoms are important symptom dimensions in the recovery process (Tandon et al., 2009; Walther and Strik, 2012). However, given the fact that studies investigating changes in these symptom dimensions are underreported, we recommend more focus on longitudinal assessments of change in motor and cognitive symptoms in future research.

4.2. The influence of moderators of changes in symptoms

Results of moderating effects on study outcomes showed that a high severity of both positive and overall symptoms at baseline was associated with more improvement of all symptom dimensions. A plausible explanation for this might be that patients with higher baseline symptom severity can decrease in symptom severity to a larger extent than those who already had a lower symptom severity at baseline. These findings support the optimistic point of view that people with severe psychotic symptoms have a high potential in showing symptomatic improvement. On the other hand, previous research indicated that initial symptom severity is associated with more residual symptoms at follow-up (Madhoo and Levine, 2015). However, because we assessed the course of psychotic symptom dimensions over time on a study level, we could not indicate the level of residual symptoms at follow-up for individual patients. Therefore, despite the fact that we found a positive association between baseline severity of symptoms and improvement in the different symptom dimensions over time, we are not able to indicate if symptoms have been remitted at follow-up.

We also found that a diagnosis of schizophrenia is negatively associated with improvement in positive symptoms. This is in line with the results of several previous studies, and might be explained by the clinical definition of schizophrenia, indicating less substantial improvement of positive symptoms throughout the course of illness compared with other schizophrenia spectrum disorders (Goghari et al., 2013; Harrow et al., 1997; Kotov et al., 2017). In contrast with our findings, previous studies also found a negative influence of schizophrenia diagnosis on improvement of negative and disorganization symptoms (Goghari et al., 2013; Kotov et al., 2017), while we observed small improvement of disorganization and negative symptoms over time, regardless of the diagnosis of the patient population. These findings indicate that the schizophrenia diagnosis is especially associated with a limited improvement of positive symptoms, whereas the changes in other symptom domains are less sensitive for the diagnosis of the patient population.

Furthermore, a lower age at baseline is associated with more substantial improvement in positive and negative symptoms. However, as shown in Supplementary Materials 3, patients within longer DOI subgroups also have an older age at baseline compared to patients in the subgroups with a shorter baseline DOI. Therefore, the influence of age at baseline could be explained by the influence of the baseline DOI of the patients on the improvement of both positive and negative symptoms.

We also found that a high baseline severity of depressive symptoms and a high baseline level of executive functioning are associated with larger improvement in negative symptoms. This is in line with previous studies that found a positive association between both executive functioning and depressive symptoms with the course of negative symptoms (Calderon-Mediavilla et al., 2020; Clark et al., 2010; Edwards et al., 2019; Rodriguez-Sanchez et al., 2008), especially on the executive components of inhibition of behavior and response initiation (Clark et al., 2010; Greenwood et al., 2008). This suggests a conceptual overlap between negative symptoms, executive functioning and depressive symptoms, that manifests itself in behavioral inhibition, apathy and reduced activity.

In contrast with previous findings, we did not find any indication for commonly found moderating effects, such as duration of untreated psychosis (DUP), education level, premorbid adjustment, ethnic background or medication adherence (Díaz et al., 2013; Lally et al., 2017; Lambert et al., 2008, 2010). One possible explanation is that these moderators are only reported by a limited number of studies, leading to a lack of consistency and statistical power for the moderating effects. Another explanation might be that there is less diversity between these moderators when analyzing and assessing them on a study level, compared with moderating effects assessed on an individual patient level.

Differences between DOI subgroups might also be explained by differences in sample characteristics at baseline between these DOI subgroups. We found differences between the DOI subgroups in baseline severity of positive symptoms, age of the study sample, DUP and years of education (see Supplementary Materials 3). Both DUP and years of education were no significant moderators of outcomes in our study. However, baseline severity of positive symptoms and age of the study sample were significant moderators of study outcomes and may have driven differences in outcomes between DOI subgroups. Especially the fact that baseline severity of positive symptoms was lower for patients with a longer DOI, combined with the minor improvement of positive symptoms we found after 10 years of illness, might indicate less room for improvement for the subgroup with a longer DOI for positive symptoms. Moreover, since DOI was associated with DUP, the effect of DUP on outcome was more relevant for short- and median-outcomes of people with first episode.

Finally, we also found that a low risk of bias on study participation and a high risk of bias on study attrition was associated with larger improvement of outcomes. The low risk of bias on study participation indicates that a more representative study sample with a good representation of the population of interest, was associated with a larger improvement of symptoms (Hayden et al., 2013). However, the positive association between high risk of attrition bias and improvement of symptoms might indicate that in some studies a selected group of participants with larger improvement of symptoms might be selectively reached at follow-up, whereas participants with reduced symptomatic improvement might drop out (Hayden et al., 2013). This might have overestimated the observed improvement of symptoms over time.

4.3. Contribution and limitations of the meta-analysis

In this meta-analysis we used a unique approach of evaluating improvement of symptoms throughout the course of schizophrenia spectrum disorders by comparing subgroups of studies that evaluated patients with different durations of illness at baseline and different lengths of follow-up. By using this approach, we were able to show whether, and how strong, patients with schizophrenia spectrum disorders improve on different symptom dimensions, and which factors contribute to symptomatic improvement. Despite the fact that this approach does not directly reflect on the episodic character of the course of illness of an individual, and potential fluctuations in between time-

points are not measured, it still adds knowledge to other clinical staging models about the overall course of symptoms in patients with schizophrenia spectrum disorders (Lieberman et al., 2001; McGorry et al., 2010). Nevertheless, several limitations should be addressed. First, the subgroup and sensitivity analyses for the outcomes of disorganization and depressive symptoms were based on a limited number of studies evaluating a largely variety of different outcome assessment, making the outcomes less reliable (Böhning et al., 2017). Furthermore, study outcomes were heterogeneous, which makes it complex to properly interpret study outcomes. For the overall analysis, large clinical heterogeneity was inevitable given the variety of studies executed in different settings and contexts. We made an effort to capture which variables explain heterogeneity by executing meta-regression analyses on potential moderators of outcomes and we took these into consideration in the interpretation of these outcomes. We found that generally patient characteristics, such as schizophrenia diagnosis, severity of symptoms and age, explain the heterogeneity of the outcomes. We did not find evidence that contextual factors, such as the study design or the type of treatment that patients received, influenced the outcomes and caused heterogeneity. In our analysis of study characteristics we also did not find an overrepresentation of specific types of treatment in clinical trials or cohort studies. This supports our decision of including both clinical trials and cohort studies in our meta-analysis. Another limitation is that studies with multiple follow-up assessments are represented in multiple subgroups based on the follow-up length in the analysis. As a consequence, these studies are overrepresented in the overall analyses of each outcome, which might have biased the results. Another limitation is that in most studies it is not clear how DOI of the study sample was assessed. Although all included studies reported DOI similarly as the duration illness after onset of schizophrenia spectrum disorder, it is not clear at what time-point each study started registering the DOI of its sample. This could have led to modest differences between studies. In addition, we categorized studies into different DOI subgroups based on the mean DOI of the study sample. This inevitably means that in some studies a part of the study sample has a shorter or longer DOI than the upper or lower limit of the DOI subgroup. Therefore, the DOI subgroup of a study is not necessarily accounting for the individual DOI of all the included participants in that study. In the meta-regression analysis we compared studies in which all participants in the study had a baseline DOI that matches the range of the DOI subgroup in which the study was categorized ($k = 32$), with studies in which a part of the study sample had a baseline DOI that was different from the DOI subgroup the study was categorized in based on its mean DOI at baseline ($k = 37$). However, we found no significant differences between both groups. Therefore, despite the fact that the DOI subgroup of a study is not necessarily representative for the whole study sample, it did not affect the reported outcomes. Furthermore, our inclusion criteria were relatively strict in order to keep our study population and study design as homogeneous as possible. Especially our inclusion criterion to exclusively select studies of which 100% of the study sample had a diagnosis of schizophrenia spectrum disorder, and our inclusion criterion to select studies with a follow-up length of at least one year, resulted in exclusion of many reports. Furthermore, we found that the baseline severity of positive symptoms is higher in subgroups with shorter DOI compared with subgroups with longer DOI. This indicates that the potential of improvement in positive symptoms is larger in studies investigating patients earlier in the course of illness, which might overestimate differences in improvement of positive symptoms over time between short and long baseline DOI subgroups. Finally, we also need to address that outcomes are evaluated on a study level, and as a consequence we were not able to adequately grasp the clinical diversity of psychotic patients on an individual level within each study.

4.4. Conclusions

In conclusion, we found large improvements of positive symptoms and modest improvements in disorganization, negative and depressive symptoms over time for patients with schizophrenia spectrum disorders. Both positive and disorganization symptom dimensions show comparable trends of improvement over time with the highest potential of improvement for patients earlier in the course of illness. In contrast, we observed only modest improvement of negative and depressive symptoms throughout the course of illness. Despite the differences between DOI subgroups, results did indicate a rather optimistic insight that improvement in symptoms is still possible for patients with longer durations of illness. Future research should focus on exploring ways to improve negative symptoms and depressive symptoms, preferably in relation with relevant cognitive and functional outcomes of psychosis, to increase the potential for symptomatic, functional and personal recovery in patients with schizophrenia spectrum disorder.

Declaration of competing interest

All authors declare no financial relationships with commercial interest

Acknowledgements

We would like to thank and acknowledge Kete Klaver (KK) and Ravi Meijer (RM) for their important contribution to the study selection and Matthijs Oud (MO) for his advice and support in the quality assessment procedure.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.jpsychires.2023.06.031.

References

American Psychiatric Association, 2013. Diagnostic and Statistical Manual of Mental Disorders, fifth ed. American Psychiatric Association, Washington, DC.
Andreasen, N.C., 1995. Symptoms, signs, and diagnosis of schizophrenia. Lancet 346 (8973), 477–481.
Arndt, S., Andreasen, N.C., Flaum, M., Miller, D., Nopoulos, P.A., 1995. Longitudinal study of symptom dimensions in schizophrenia: prediction and patterns of change. Arch. Gen. Psychiatr. 52 (5), 352–360.
Benjamini, Y., Hochberg, Y., 1995. Controlling the false discovery rate: a practical and powerful approach to multiple testing. J. Roy. Stat. Soc. B 57 (1), 289–300.
Best, M.W., Law, H., Pyle, M., Morrison, A.P., 2020. Relationships between psychiatric symptoms, functioning and personal recovery in psychosis. Schizophr. Res. 223, 112–118.
Birchwood, M., Todd, P., Jackson, C., 1998. Early intervention in psychosis: the critical period hypothesis. Br. J. Psychiatr. 172 (S33), 53–59.
Böhning, D., Lerdsuwansri, R., Holling, H., 2017. Some general points on the I^2-measure of heterogeneity in meta-analysis. Metrika 80 (6–8), 685–695.
Boonstra, N., Klaassen, R., Sytema, S., Marshall, M., De Haan, L., Wunderink, L., Wiersma, D., 2012. Duration of untreated psychosis and negative symptoms—a systematic review and meta-analysis of individual patient data. Schizophr. Res. 142 (1–3), 12–19.
Borenstein, M., Higgins, J.P., 2013. Meta-analysis and subgroups. Prev. Sci. 14 (2), 134–143.
Breitborde, N.J., Srihari, V.H., Woods, S.W., 2009. Review of the operational definition for first-episode psychosis. Early Intervent. Psychiatr. 3 (4), 259–265.
Calderon-Mediavilla, M., Vila-Badia, R., Dolz, M., Butjosa, A., Barajas, A., Del Cacho, N., Sánchez, B., Pardo, M., Baños, I., Usall, J., Ochoa, S., GENIPE Group, PROFEP Group, 2020. Depressive symptoms and their relationship with negative and other psychotic symptoms in early onset psychosis. Eur. Child Adolesc. Psychiatr. 30, 1383–1390.
Chinn, S., 2000. A simple method for converting an odds ratio to effect size for use in meta-analysis. Stat. Med. 19, 3127–3131.
Clark, L.K., Warman, D., Lysaker, P.H., 2010. The relationships between schizophrenia symptom dimensions and executive functioning components. Schizophr. Res. 124 (1–3), 169–175.
De Winter, L., Couwenbergh, C., Van Weeghel, J., Hasson-Ohayon, I., Vermeulen, J.M., Mulder, C.L., Boonstra, N., Klaver, K.M., Oud, M., De Haan, L., Veling, W., 2021. Changes in social functioning over the course of psychotic disorders – a meta-analysis. Schizophr. Res. 239, 55–82.
Díaz, I., Pelayo-Terán, J.M., Pérez-Iglesias, R., Mata, I., Tabarés-Seidesos, R., Suárez-Pinilla, P., Vázquez-Barquero, J.L., Crespo-Facorro, B., 2013. Predictors of clinical remission following a first episode of non-affective psychosis: sociodemographics, premorbid and clinical variables. Psychiatr. Res. 206 (2–3), 181–187.

Edwards, C.J., Garety, P., Hardy, A., 2019. The relationship between depressive symptoms and negative symptoms in people with non-affective psychosis: a meta-analysis. Psychol. Med. 49 (15), 2486–2498.
Emsley, R., Oosthuizen, P.P., Kidd, M., Koen, L., Niehaus, D.J., Turner, H.J., 2006. Remission in first-episode psychosis: predictor variables and symptom improvement patterns. J. Clin. Psychiatr. 67 (11), 1707–1712.
Frascarelli, M., Tognin, S., Mirigliani, A., Parente, F., Buzzanca, A., Torti, M.C., Tinelli, E., Caramia, F., De Fabio, F., Biondi, M., Fusar-Poli, P., 2015. Medial frontal gyrus alterations in schizophrenia: relationship with duration of illness and executive dysfunction. Psychiatr. Res. Neuroimaging 231 (2), 103–110.
Goghari, V.M., Harrow, M., Grossman, L.S., Rosen, C., 2013. A 20-year multi-follow-up of hallucinations in schizophrenia, other psychotic, and mood disorders. Psychol. Med. 43 (6), 1151–1160.
Greenwood, K.E., Morris, R., Sigmundsson, T., Landau, S., Wykes, T., 2008. Executive functioning in schizophrenia and the relationship with symptom profile and chronicity. J. Int. Neuropsychol. Soc. 14 (5), 782–792.
Häfner, H., 2019. From onset and prodromal stage to a life-long course of schizophrenia and its symptom dimensions: how sex, age, and other risk factors influence incidence and course of illness. Psychiatr. J., 9804836
Häfner, H., 2015. What is schizophrenia? 25 years of research into schizophrenia-the Age Beginning Course Study. World J. Psychiatr. 5 (2), 167–169.
Haro, J.M., Altamura, C., Corral, R., Elkis, H., Evans, J., Krebs, M.O., Zink, M., Malla, A., Méndez, J.I., Bernasconi, C., Lalonde, J., Nordstroem, A.L., 2018. Understanding the course of persistent symptoms in schizophrenia: longitudinal findings from the pattern study. Psychiatr. Res. 267, 56–62.
Harrow, M., Sands, J.R., Silverstein, M.L., Goldberg, J.F., 1997. Course and outcome for schizophrenia versus other psychotic patients: a longitudinal study. Schizophr. Bull. 23 (2), 287–303.
Hayden, J.A., Van der Windt, D.A., Cartwright, J.L., Côté, P., Bombardier, C., 2013. Assessing bias in studies of prognostic factors. Ann. Intern. Med. 158 (4), 280–286.
Higgins, J.P., 2008. Cochrane Handbook for Systematic Reviews of Interventions. The Cochrane Collaboration, version 5.0.1. http://www.cochrane-handbook.org.
Higgins, J.P., Thompson, S.G., 2002. Quantifying heterogeneity in a meta-analysis. Stat. Med. 21 (11), 1539–1558.
Kotov, R., Fochtmann, L., Li, K., Tanenberg-Karant, M., Constantino, E.A., Rubinstein, J., Perlman, G., Velthorst, E., Fett, A.J., Carlson, G., Bromet, E.J., 2017. Declining clinical course of psychotic disorders over the two decades following first hospitalization: evidence from the Suffolk County Mental Health Project. Am. J. Psychiatr. 174 (11), 1064–1074.
Lally, J., Ajnakina, O., Stubbs, B., Culinane, M., Murphy, K.C., Gaughran, F., Murray, R. M., 2017. Remission and recovery from first-episode psychosis in adults: systematic review and meta-analysis of long-term outcome studies. Br. J. Psychiatr. 211 (6), 350–358.
Lambert, M., Karow, A., Leucht, S., Schimmelmann, B.G., Naber, D., 2010. Remission in schizophrenia: validity, frequency, predictors, and patients' perspective 5 years later. Dialogues Clin. Neurosci. 12 (3), 393–407.
Lambert, M., Naber, D., Schacht, A., Wagner, T., Hundemer, H.P., Karow, A., Huber, C. G., Suarez, D., Haro, J.M., Novick, D., Dittmann, R.W., Schimmelmann, B.G., 2008. Rates and predictors of remission and recovery during 3 years in 392 never-treated patients with schizophrenia. Acta Psychiatr. Scand. 118 (3), 220–229.
Lefebvre, S., Pavlidou, A., Walther, S., 2020. What is the potential of neurostimulation in the treatment of motor symptoms in schizophrenia? Expert Rev. Neurother. 20 (7), 697–706.
Lieberman, J.A., Perkins, D., Belger, A., Chakos, M., Jarskog, F., Boteva, K., Gilmore, J., 2001. The early stages of schizophrenia: speculations on pathogenesis, pathophysiology, and therapeutic approaches. Soc. Biol. Psychiatr. 50 (11), 884–897.
Madhoo, M., Levine, S.Z., 2015. Initial severity effects on residual symptoms in response and remission: a STAR* D study during and after failed citalopram treatment. J. Clin. Psychopharmacol. 35 (4), 450–453.
McGinty, J., Haque, M.S., Upthegrove, R., 2018. Depression during first episode psychosis and subsequent suicide risk: a systematic review and meta-analysis of longitudinal studies. Schizophr. Res. 195, 58–66.
McGorry, P.D., Nelson, B., Goldstone, S., Yung, A.R., 2010. Clinical staging: a heuristic and practical strategy for new research and better health and social outcomes for psychotic and related mood disorders. Can. J. Psychiatr. 55 (8), 486–497.
Moher, D., Liberati, A., Tetzlaff, J., Altman, D.G., Prisma Group, 2009. Reprint—preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. Phys. Ther. 89 (9), 873–880.
Preston, N.J., 2000. Predicting community survival in early psychosis and schizophrenia populations after receiving intensive case management. Aust. N. Z. J. Psychiatr. 34 (1), 122–128.
R Core Team, 2016. R: A Language and Environment for Statistical Computing. Vienna, Austria. https://www.R-project.org/.
Rodriguez-Sanchez, J.M., Crespo-Facorro, B., Gonzalez-Blanch, C., Pérez-Iglesias, R., Álvarez-Jiménez, M., Martínez, O., Vázquez-Barquero, J.L., 2008. Cognitive functioning and negative symptoms in first episode schizophrenia: different patterns of correlates. Neurotox. Res. 14 (2), 227–235.
Rowe, A.R., Mercer, L., Casetti, V., Sendt, K.V., Giaroli, G., Shergill, S.S., Tracy, D.K., 2015. Dementia praecox redux: a systematic review of the nicotinic receptor as a target for cognitive symptoms of schizophrenia. J. Psychopharmacol. 29 (2), 197–211.
Salazar de Pablo, G.S., Besana, F., Arienti, V., Catalan, A., Vaquerizo-Serrano, J., Cabras, A., Pereira, J., Soardo, L., Coronelli, F., Kaur, S., Da Silva, J., Oliver, D., Petros, N., Moreno, C., González-Pinto, A., Díaz-Caneja, C.M., Shin, J.I., Politi, P., Solmi, M., Norgatti, R., Mensi, M.M., Arango, C., Correll, C.U., McGuire, P., Fusar-Poli, P., 2021. Longitudinal outcome of attenuated positive symptoms, negative symptoms, functioning and remission in people at clinical high risk for psychosis: a meta-analysis. E Clin.Med. 36, 100909.
Schnack, H.G., 2019. Improving individual predictions: machine learning approaches for detecting and attacking heterogeneity in schizophrenia (and other psychiatric diseases). Schizophr. Res. 214, 34–42.
Sullivan, S.A., Kounali, D., Cannon, M., Davis, A.S., Fletcher, P.C., Holmans, P., Jones, H., Jones, P.B., Linden, D.E.J., Lewis, G., Owen, M.J., O'Donovan, M.O., Rammos, A., Thompson, A., Wolke, D., Heron, J., Zammit, S., 2020. A population-based cohort study examining the incidence and impact of psychotic experiences from childhood to adulthood, and prediction of psychotic disorder. Am. J. Psychiatr. 177 (4), 308–317.
Tandon, R., Nasrallah, H.A., Keshavan, M.S., 2009. Schizophrenia, "just the facts" 4. Clinical features and conceptualization. Schizophr. Res. 110 (1–3), 1–23.
The Nordic Cochrane Centre, 2014. Review Manager (RevMan) [Computer Program]. Version 5.3. The Nordic Cochrane Centre, The Cochrane Collaboration, Copenhagen.
Van Eck, R.M., Burger, T.J., Vellinga, A., Schirmbeck, F., De Haan, L., 2018. The relationship between clinical and personal recovery in patients with schizophrenia spectrum disorders: a systematic review and meta-analysis. Schizophr. Bull. 44 (3), 631–642.
Ventura, J., Subotnik, K.L., Gitlin, M.J., Gretchen-Deoorly, D., Ered, A., Villa, K.F., Hellemann, G.S., Nuechterlein, K.H., 2015. Negative symptoms and functioning during the first year after a recent onset of schizophrenia and 8 years later. Schizophr. Res. 161 (2–3), 407–413.
Walther, S., Strik, W., 2012. Motor symptoms and schizophrenia. Neuropsychobiology 66 (2), 77–92.