Schizophrenia symptoms are inherently heterogeneous: a systematic review of cluster and group-based studies

Tesfa Dejenie Habtewold[1,2*], Lyan H. Rodijk[3], Edith J. Liemburg[2,4], Grigory Sidorenkov[1], Richard Bruggeman[2,4], Behrooz Z. Alizadeh[1,2*]

[1]University of Groningen, University Medical Center Groningen, Department of Epidemiology, Groningen, The Netherlands
[2]University of Groningen, University Medical Center Groningen, University Center for Psychiatry, Rob Giel Research Center, Department of Psychiatry, Groningen, The Netherlands
[3]University of Groningen, University Medical Center Groningen, Department of Pediatric Surgery, Groningen, The Netherlands
[4]University of Groningen, University Medical Center Groningen, Department of Neuroscience, Groningen, The Netherlands

*Correspondence to Tesfa Dejenie Habtewold (t.d.habtewold@umcg.nl) and Behrooz Z. Alizadeh (b.z.alizadeh@umcg.nl)

Abstract

Introduction: Schizophrenia (SCZ) is a complex, heterogeneous, psychotic disorder with variable phenotypic expressions, variable course patterns and a complex etiology. It affects individuals, families, and the society at large, with most of the affected population having severe symptomatic and functional outcome.

Aims: The objectives of this systematic review were to summarize the number of clusters and trajectories of schizophrenia symptoms course patterns discovered until now, to identify predictors of clusters and trajectories, to highlight knowledge gaps and to point out a way forward to optimize future cluster- and trajectory-based studies.

Methods: PsycINFO, PubMed, PsycTESTS, PsycARTICLES, SCOPUS, EMBASE, and Web of Science databases were searched using a comprehensive search strategy adapted to each database. Cross-sectional and longitudinal studies published from 2008 to 2018, that reported at least two clusters or trajectory groups of patients, siblings and controls using a statistical method across positive, negative, and cognitive impairment symptom dimensions or combination of symptom dimensions were included. Two reviewers independently screened and extracted the data from included studies. A narrative synthesis was performed, and data were summarized using tables.

Results: Of 2,282 studies identified, 47 fulfilled the inclusion criteria and were included in the qualitative synthesis. These studies were conducted globally in more than 17 countries (15 studies in the USA) and published from 2009 to 2019. Sixteen of the included studies had a longitudinal design, involving 11,475 patients with schizophrenia-spectrum disorders, 1,059 siblings and 653 controls, whereas 31 studies had a cross-sectional design involving 5,271 patients with schizophrenia-spectrum disorders, 7,423 siblings, and 2,346 controls. The longitudinal studies discovered two to five patient trajectory groups based on positive and negative symptoms, and four to five patient and sibling trajectory groups identified based on cognitive deficits. Regarding cross-sectional studies, three clusters of patients were found based on positive and negative symptoms while four clusters of siblings were identified based on positive and negative schizotypy. Regarding cognitive deficits, three to five clusters were reported in patients and their unaffected siblings. Age, gender, ethnicity, educational status, age of illness onset, diagnosis of schizophrenia, depressive symptoms, general psychopathology, severity of positive and negative symptoms, cognitive performance, premorbid functioning, quality of life and global functioning were important predictors among patients and their unaffected siblings.

Conclusions: The evidence from cluster- and trajectory-based studies in the past decade showed that clinical symptoms of schizophrenia are clearly heterogeneous across patients, siblings and controls. Despite this fact, the extent of heterogeneity is yet to be investigated. To fully understand the heterogeneity, further work is expected from psychiatric researchers targeting longitudinal study design, unaffected siblings and utilizing genetic markers as a predictor.

Introduction

Schizophrenia is a heterogeneous complex psychotic disorder with variable phenotypic expression, variable patterns of course and complex etiology that affects individuals, families and the society at large, with most of the affected population having a severe course of symptomatic and functional outcome.[1] The prevalence of schizophrenia is 4.6 per 1.000 individuals with a lifetime morbidity risk of 0.7%.[2] The incidence in men and women is 15 and 10 per 100,000 individuals, respectively. The first episode of schizophrenia usually occurs in late adolescence or early adulthood.[2] Schizophrenia has been associated with various environmental and genetic factors.[2] It is also known that siblings of patients with schizophrenia may develop schizophrenia over time due to shared genetic and environmental factors.[3,4] The concordance rate of schizophrenia is 33% in monozygotic twins and 7% in dizygotic twins.[4]

Schizophrenia has three groups of clinical symptoms including positive symptoms, negative symptoms and cognitive deficits, which are important outcomes quite often used in psychiatric research. These symptoms assessed by standard psychometric assessment tools with or without validation to the local population. Positive symptom includes hallucinations, delusions, and disorganized thinking.[5] Negative symptoms include emotional expressive deficit, social amotivation, social withdrawal and difficulty in experiencing pleasure and the prevalence is 50-90% in FEP and persists in 20-40% of patients with SCZ.[6-8] Negative and positive symptoms assessed by the positive and negative syndrome scales.[9-12] Cognitive deficit affects 75-80% of patients with schizophrenia.[13] The most common deficit occurred in executive function, processing speed, memory (e.g. episodic, verbal and working), attention, verbal fluency, problem-solving and social cognition.[14-22] Cognitive dysfunction can be evaluated by various standard neuropsychological battery tests.[23,24]

From the beginning of conception of the term schizophrenia until know, various attempts have been made to identify sociodemographic, clinical, neurocognitive and other factors that can influence the heterogeneity of clinical and functional outcomes. Existing evidence shows that the course of schizophrenia has four different trajectories over time - progressive deterioration, relapsing, progressive amelioration, and stability.[25] These divergent views of the course of schizophrenia have recently been investigated by subtyping using imaging, biological and symptom data.[26] The other methods of subtyping are using statistical approaches, such as cluster analysis, latent class analysis, and growth mixture models, based on clinical symptom (sub)scale scores examining baseline and subsequent assessments.[25-27] These models identify groups of individuals who have a similar profile or course of symptoms over time and estimate the effects of predictors on trajectory shape and group membership.[28] A trajectory or cluster is a group of people that has a homogenous symptom profile within that group and a significantly dissimilar (i.e., heterogeneous) profile from other groups.[29] In this review, we used 'trajectory' for groups identified by longitudinal studies and 'cluster' for groups identified by cross-sectional studies. Subgrouping approaches are useful for categorization of patients, understanding of etiologies and pathophysiology, and to predict treatment response.[27]

Despite a century of efforts, understanding the heterogeneity and course of schizophrenia has been unsuccessful due to nature of its clinical symptoms, variation in

response to treatment, and the lack of valid, stable, and meaningful subtyping methods.[27,30,31] Heterogeneity in clinical outcomes can be manifested between groups or subjects, within subjects over time, within and between diagnostic groupings, and caused by several intrinsic and extrinsic factors.[30,32] The classification of clinical outcomes into dichotomous categories, such as recovered or not, and symptom remission or not, is a common practice within schizophrenia research. However, this can also be problematic if dichotomization is undertaken without evidence of the distribution of a population. Cut-off scores used for categorization are often arbitrary in nature. All these pitfalls may lead to the loss of information, inefficient analysis of continuous data and difficulties in the translation of results into clinically meaning information.[29]

Cluster- and trajectory-based studies of clinical symptoms of schizophrenia show inconsistent findings due to high symptomatic variability between patients and within patients over time and also have several limitations.[6,15] Previous studies are often hampered by the heterogeneity of age, sex, and diagnosis of patients, severity of symptoms, use of various assessment tools, use of different clustering algorithms, use of different scoring and standardization techniques, small sample size and shorter duration of follow-up.[33] Furthermore, not all prior studies included sibling and control participants.[33] All these factors blur our understanding of the current state-of-art regarding heterogeneity of schizophrenia symptoms. Therefore, there is a pressing need to synthesize the contemporary evidence, evaluate the extent and origin of heterogeneity, and develop a consensus outline for clinical practice.

To our knowledge, there is no comprehensive review based on cluster- and trajectory-based studies of positive symptoms, negative symptoms and cognitive deficits in patients with schizophrenia-spectrum disorders, their unaffected siblings and healthy controls. To date, reviews have been conducted on various aspects of cognitive dysfunction[13,34-43], negative symptoms[8,44,45], and positive symptoms.[46] The focus of these past reviews has largely been based on the traditional approach of determining average change in course of symptoms over time, and variation between subjects (patient vs relatives, relatives vs controls, patients vs controls) and diagnosis. They are also based on correlation analysis which is believed not the strong measure of association. In addition, none of these reviews fully addressed symptomatic clusters or trajectories in patients with SCZ, their unaffected siblings and healthy controls. In this review, we summarized the number of clusters and trajectories in patients with schizophrenia-spectrum disorders, their unaffected siblings and healthy controls that discovered by longitudinal and cross-sectional studies until now. The predictors of clusters or trajectories were also identified and discussed. We further highlight gaps in current knowledge and point out a way forward to optimize evidence from the future cluster- and trajectory-based studies.

Methods

Registration and reporting

This systematic review was conducted and reported based on a registered protocol (http://www.crd.york.ac.uk/PROSPERO/display_record.php?ID=CRD42018093566) and the Preferred Reporting Items for Systematic Review and Meta-Analysis (PRISMA) statement guideline respectively.[47,48] The screening and selection process of the reviewed articles are further illustrated using a PRISMA flow diagram.

Databases and search terms

For all available publications, a systematic search of PubMed, PsycINFO, PsycTESTS, PsycARTICLES, SCOPUS, EMBASE and Web of Science electronic databases was performed. A comprehensive search strategy was developed for PubMed and adapted to each database in consultation with a medical information specialist (Supplementary file 1). The following search terms were used in their singular or plural form in their title, abstract, keywords and text: "schizophrenia", "psychosis", "non-affective psychosis", "cognitive deficit", "cognitive dysfunction", "cognitive alteration", "negative symptoms", "deficit syndrome", "positive symptoms", "psychopathology", "cognit*", "neuropsycholog*", "neurocognition", "longitudinal", "follow-up", "course", "heterogeneity", "endophenotype", "profile", "cluster analysis", "siblings", "healthy controls", "latent class analyses", "Symptom trajectories", "traject*", "group modeling" and "trajectory". Cross-references of included articles and grey literature were also hand-searched. Furthermore, we searched the table of contents of the journals of Schizophrenia Research, Schizophrenia Bulletin, Acta Psychiatrica Scandinavica and British Journal of Psychiatry to explore relevant studies. The final search was conducted in March 2019.

Inclusion and exclusion criteria

Studies meeting the following criteria were included: (1) cross-sectional and longitudinal studies; (2) studies that reported at least two clusters or trajectory groups of individuals using a statistical method based on distinct positive symptom, negative symptom, and neurocognitive or social cognitive) impairment dimensions or a combination of these symptom dimensions; (3) studies conducted in patients with schizophrenia-spectrum disorders, their unaffected siblings, and healthy controls irrespective of any clinical (e.g. medication status, severity of illness) and sociodemographic characteristics; and (4) English published studies from 2009 to 2019. The publication year was limited to the last decade to capture the latest available evidence. In addition, the number of large sample cohorts has been increasing in the last decade which we believe studies can provide statistically powerful precise estimates and successful subtyping of schizophrenia symptoms. In order to maximize the number of searched articles, the follow-up period in longitudinal studies was not restricted. Trajectory-based studies based on mean score change over time were excluded because they did not cross-sectionally or longitudinally reveal the actual heterogeneity of groups.[13,34-43] In addition, studies based on the non-statistical methods of clustering (e.g. family-based clustering) were excluded. Review papers, commentaries, duplicate studies, editorials, and qualitative studies were excluded as well.

Furthermore, we excluded studies where the trajectory groups or clusters generated based on scores constructed using a combination of schizophrenia symptoms and other unspecified psychotic symptoms.

Data retrieval and synthesis

Studies retrieved from all databases were exported to RefWorks version 2.0 for Windows web-based citation manager. Close and exact duplicates were deleted. All independent studies were exported to Microsoft Excel spreadsheet to screen for further inclusion criteria. TD and LR independently screened the titles and abstracts followed by a test of agreement using Cohen's Kappa coefficient. The two reviewers had substantial agreement, which Kappa coefficient was 0.62. Inconsistent decisions on title and abstract inclusion were managed by discussion. Finally, full-text review was performed and the following data were independently extracted by TD and LR: first author name, publication year, country, cohort/research center, study population, sample size, symptom dimension(s), assessment tool, study design, duration of follow-up (only for longitudinal studies), frequency of assessment, method of calculating tests composite score, method of clustering/trajectory analysis, number of identified clusters/trajectory groups and significant predictors of clusters or trajectories.[49] The corresponding author(s) were contacted by email when full-text of an included article was not accessible. Whenever the cohort or research center was not clearly reported, we extracted the institutional affiliation of the first/corresponding author. The disagreement was resolved by consensus and in consultation with BZ and RB. Due to substantial heterogeneity of studies, a narrative synthesis was done and summarized using tables.

Results

Search results

In total, 2,262 studies were identified through database searching and an additional 20 studies through manual searching of cross-references and table of contents of relevant journals. After removing duplicate articles and applying the inclusion and exclusion criteria, title and abstract of 1,291 articles were screened which resulted in the exclusion of 1,236 articles. As a result, 55 articles were selected for full-text review and seven articles[50-57] were excluded due to unclear outcome, mixed diagnosis of the study population, use of non-statistical method of clustering and clustering based on different phenotypes of schizophrenia. Finally, data were extracted from 47 cluster- and trajectory-based studies. The PRISMA flow diagram of screening and selection process is shown in Figure 1.

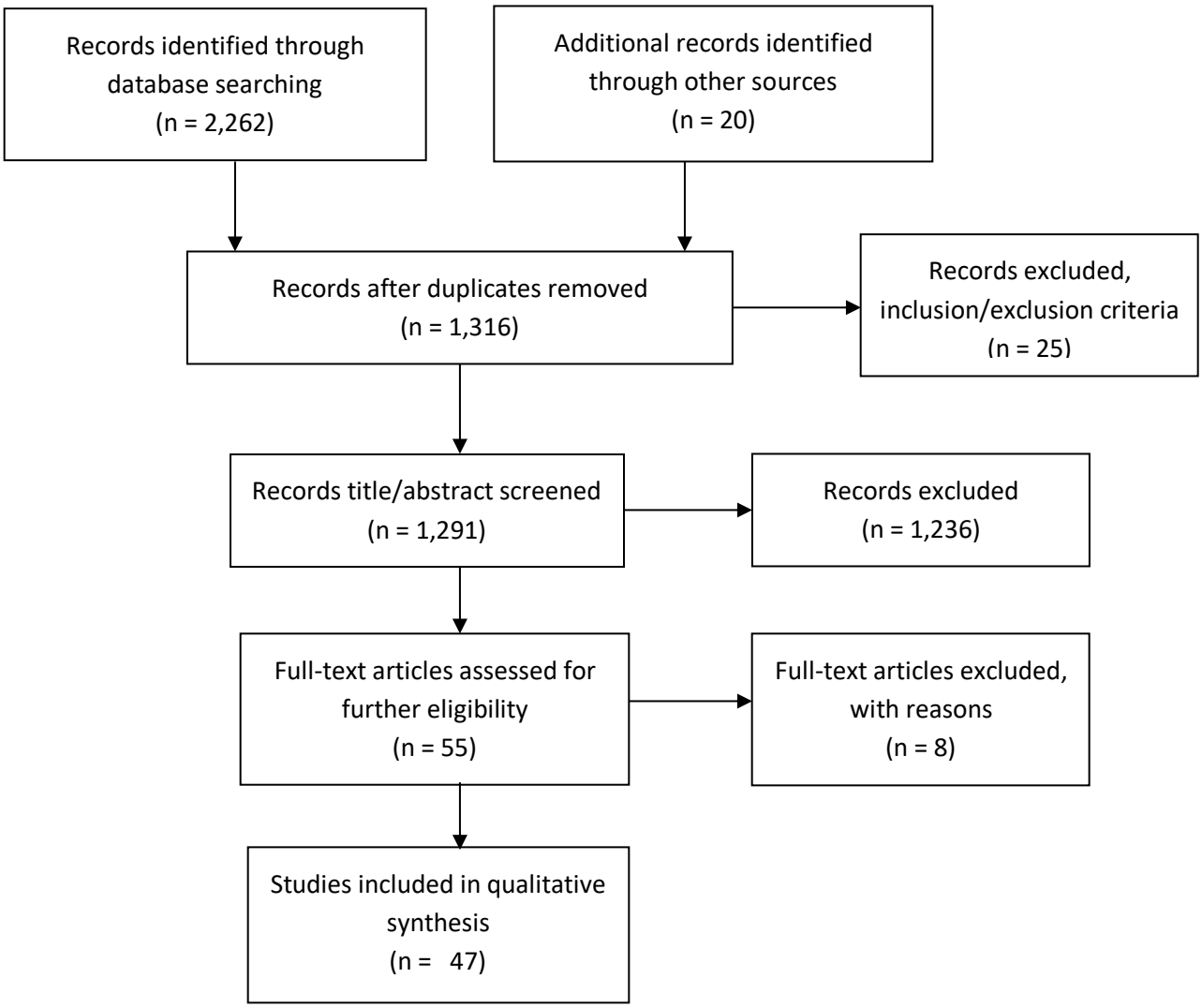

Figure 1: PRISMA flow diagram showing schematic presentation of literature screening and selection process.

Overview of included studies

The included 47 studies were conducted globally in more than 17 countries (15 studies in USA) and published from 2009 to 2019. Of these, 16 studies were longitudinal studies involving 11,475 patients, 1,059 siblings and 653 controls, whereas 31 studies were cross-sectional studies involving 5,271 patients, 7,423 siblings, and 2,346 controls. Only one longitudinal study[58] and six cross-sectional studies[59-63] included siblings. Almost all longitudinal studies reported trajectories of positive and negative symptoms whereas most cross-sectional studies reported clusters based on cognitive function.

Evidence from longitudinal studies

From the total of 16 longitudinal studies, conducted in more than eight countries, 14 studies[25,29,31,64-74] investigated the trajectory of positive and negative positive symptoms in patients, and only two studies[30,58] examined the trajectory of cognitive impairment in patients and siblings. The duration of follow-up ranges from six weeks to 10 years and included all population age groups. The total sample size in each study ranges from 138 to 1990 subjects though it varied in symptom dimensions. One study[58] investigated the association between patients' and siblings' cognitive trajectories whereas another study[73] examined the association between positive and negative symptom trajectories. Moreover, five studies reported the influence of trajectories on long-term social, occupational and global functioning, and health-related or general quality of life.[68,69,71,72,74] Even though all studies had similar aims, they used different name for the trajectory analysis methods, such as growth mixture modelling (GMM)[31,65,73], latent class growth analysis (LCGA)[29,30,66,69,71,72,74], mixed mode latent class regression modelling[25,64,70] and group-based trajectory modelling (GBTM).[58,67,68] Akaike's Information Criterion (AIC), Bayesian information criterion (BIC), logged Bayes factor, sample-size-adjusted BIC [aBIC], bootstrap likelihood ratio test [BLRT], Lo–Mendell–Rubin Likelihood Ratio Test (LMR-LRT) and entropy were reported model selection indices. Of these indices, Bayesian information criterion (BIC) reported by all studies except one study[30] that reported deviance information criterion (DIC).

Among studies with less than two years of follow-up (Table 1), three studies[25,65,67] discovered five trajectories, and the other three studies[31,64,73] identified three trajectories of positive symptoms. These trajectories were predicted by age, gender, ethnicity, cannabis use, age of illness onset, diagnosis, duration of untreated psychosis, extrapyramidal symptoms, depressive symptoms, quality of life, general psychopathology, types of antipsychotic medication, cognitive performance, premorbid functioning, severity of positive and negative symptoms.[25,31,64,65,67,73] Similarly for the negative symptom dimension, three studies[25,65,67] discovered five trajectories and the other three studies[31,72,73] reported four trajectories. The identified predictors of were age, gender, ethnicity, family history of non-affective psychosis, age of onset of illness, extrapyramidal symptoms, quality of life, general psychopathology, diagnosis of schizophrenia, cognitive performance, premorbid functioning, premorbid adjustment, depressive symptoms, types of antipsychotic medication, and severity of positive and negative symptoms.[25,31,64,65,67,72,73]

Combining both positive and negative symptom dimensions, two studies[25,66] discovered five trajectories, one study[31] found out four trajectories and one study[73] identified three trajectories. The predictors of these trajectories were ethnicity, age of illness onset, duration of illness, previous hospitalizations, extrapyramidal and depressive symptoms, quality of life, severity of positive and negative symptoms, general psychopathology, diagnosis of schizophrenia, cognitive performance and premorbid functioning.[25,31,66,73]

Table 1: Detailed characteristics of included longitudinal studies with a duration of follow-up of less than two years (n = 8).

Authors' and publication year	Country	Research centre/Cohort	Participants	Sample size	Symptom dimension	Tool	Frequency of assessment	Duration of follow-up	Method of calculating test score	Method of trajectory analysis	Number of trajectories identified	Predictors of trajectories
Chen 2013[73]	USA	Mulitcenter trial study, mental health outpatient clinics	Patients with Schizophrenia spectrum disorder and treated with first- and second-generation antipsychotics	400	Positive and negative symptoms	PANSS	Seven times	1 year	Sum score	Growth mixture modelling (GMM)	Three for positive symptom: Class 1, Class 2, Class 3 Four for negative symptom: Class 1, Class 2, Class 3, Class 4 Positive and negative symptom combined: dramatic and sustained early improvement, mild and sustained improvement, no improvement	Positive and negative symptoms positively associated. No other factors reported
Gee 2016[72]	UK	National EDEN study	Patients with first episode psychosis and receiving treatment for 12 months	1006	Negative symptoms	PANSS	Three times	1 year	Mean score	Latent class growth analysis (LCGA)	Four trajectories: minimal decreasing, mild stable, high decreasing, high stable.	Male gender, family history of non-affective psychosis, poor premorbid adolescent adjustment and baseline depression
Levine 2010[25]	12 countries	International cohort/ Johnson & Johnson Pharmaceutical Research and Development	Patients with early episode psychosis and receiving treatment for more than three months	491	Positive and negative symptoms	PANSS	Six times	6 months	Sum score	Mixed mode latent class regression modelling	Five: stable (3 groups), improved and stable, marked improvement)	Not having a diagnosis of schizophrenia, age of onset, cognitive functioning, premorbid functioning
Case et al 2011[31]	3 countries	64 research centres	Patients with psychosis and treated with antipsychotics	628	Negative and positive symptoms	PANSS	Eight times	3 months	Sum score	Growth-mixture modelling (GMM)	Four: moderate-gradual, rapid, high-gradual, unsustained improvement	Extrapyramidal and depression symptoms, quality of life, age at onset of illness, ethnicity, positive and negative symptoms, general psychopathology
Pelayo-Terán et al 2014[67]	Spain	University Hospital Marqués de Valdecilla/Clinical Programme on First-Episode Psychosis of Cantabria (PAFIP)	Patients with a first episode of non-affective psychosis and no prior treatment	161	Negative and positive symptoms	SANS and SAPS	Six times	6 weeks	Sum score	Group-based trajectory modelling (GBTM)	Five for positive symptom: responders, dramatic responders, partial responders, slow partial responders, non-responders	Positive symptom: Duration of untreated psychosis and cannabis use Negative symptom:

											Five for negative symptom: responders, mild non-responders, moderate non-responders, partial responders, poor responders	SCZ diagnosis
Schennach et al 2012[66]	German	Multi-centre study/ German Research Network on Schizophrenia (GRNS)	Patients with schizophrenia spectrum disorder	399	Negative and positive symptoms	PANSS	More than 10 times	> 5 months	Sum score	Latent class growth analysis (LCGA)	Five: early and considerable response, rapid and dramatic response, early and satisfying response, gradual response	Depressive symptoms at admission, functioning, duration of illness, previous hospitalizations
Stauffer et al 2011[65]	USA and other countries	Multicentre study	Patients with chronic schizophrenia and receiving treatment	1,990	Negative and positive symptoms	PANSS	11 times	≤6 months	Sum score	Growth mixture modelling (GMM)	Five: dramatic responders, partial responders, partial responders-unsustained (late), partial responders-unsustained (early), Delayed Responders	Age, gender, ethnicity
Levine et al 2012[64]	USA	57 clinical sites	Patients with chronic schizophrenia and receiving treatment	1,124	Negative and positive symptoms	PANSS	Eight times	1.5 years	Sum score adjusted for the baseline score	Mixed-mode latent regression modelling	Three: low deteriorators, responders, high deteriorators	Type of antipsychotics

PANSS = Positive and Negative Syndrome Scale; SAPS = Scale for the Assessment of Positive Symptoms; SANS=Scale for the Assessment of Negative Symptoms; HEN = High Royds Evaluation of Negativity Scale

In studies with two years or longer duration of follow-up (Table 2), two studies[29,70] discovered five trajectories and two studies[71,74] found two trajectories of positive symptoms. The predictors were gender, educational status, duration of untreated psychosis, global functioning, living situation, involuntary admission, premorbid functioning, cognitive performance, severity of positive and negative symptoms, substance abuse and diagnosis.[29,70,71,74] Regarding the negative symptom dimension, one study[70] identified five trajectories, two studies[29,74] discovered four trajectories, one study[69] depicted three trajectories and one study[71] found two trajectories. In addition, a study[68] from our research group identified four trajectories of negative symptom subdomains of social amotivation and expressive deficits. The predictors were age, gender, employment status, ethnicity, marital status, educational status, duration of untreated psychosis, diagnosis, global functioning, living situation, involuntary admission, quality of life, premorbid functioning, cognitive performance, severity of positive and negative symptoms, premorbid adjustment, social functioning, and disorganized and depressive symptoms.[29,68-71,74] Combining both positive and negative symptom dimensions, one study[70] identified five trajectories. The predictors were diagnosis, premorbid functioning, cognitive performance, severity of positive and negative symptoms.[70]

Furthermore, the six years longitudinal study[58] from our research group discovered five trajectories of cognitive impairment in patients and four trajectories in healthy siblings. Another study[30] reported three trajectories of global cognitive function combining patients and controls together. The predictors identified by both studies were educational status, IQ, premorbid functioning, age, gender, ethnicity, severity of positive and negative symptoms, frequency of psychotic experiences, cognitive performance, and living situation.[30,58]

Table 2: Detailed characteristics of included longitudinal studies with two years or longer duration of follow-up (n = 8).

Authors' and publication year	Country	Research centre/Cohort	Participants	Sample size	Symptom dimension	Tool	Frequency of assessment	Duration of follow-up	Method of calculating test score	Method of trajectory analysis	Number of trajectories identified	Predictors of trajectories
Abdin 2017[74]	Singapore	Institute of Mental Health/Early Psychosis Intervention Programme (EPIP) clinical database.	Patients with first-episode psychotic disorder and with no prior or minimal treatment (<12wks)	1,724	Positive and negative symptoms.	PANSS	Five times	2 years	Not clearly reported	Latent class growth analysis (LCGA)	Two for positive symptom: early response and stable trajectory, and delayed response trajectory. Four for negative symptom: early response and stable trajectory, early response and relapse trajectory, slower response and no response trajectory and delayed response trajectory	Younger age, male gender, unemployed and economically inactive status, lower education, longer duration of untreated psychosis and diagnosis of schizophrenia spectrum and delusional disorders
Austin 2015[29]	Denmark	Centre for psychiatric research/OPUS trial trail	Patients with first-episode schizophrenia spectrum disorder and had received less than 12 weeks of antipsychotic medication	496	Positive and negative symptoms	SAPS and SANS	Five times	10 years	Composite score using global scores	Latent class analysis (LCA)	Five for positive symptom: response, delayed response, relapse, non-response and episodic response. Four for negative symptom: response, delayed response, relapse and non-response	Positive symptom: longer duration of untreated psychosis, poor global functioning, SCZ diagnosis and substance abuse Negative symptom: poor social functioning, disorganized symptoms and schizophrenia diagnosis
Jager 2014[71]	Germany	ELAN study, psychiatric hospitals	Patients with schizophrenia or schizoaffective disorder and receiving treatment for more than one year	268	Positive and negative symptoms	PANSS	Five times	2 years	Sum score	Latent class growth analysis (LCGA)	Two: amelioration/decrease in all symptoms, stable positive and negative symptoms and deteriorating general psychopathology symptoms)	Global functioning (GAF score), gender, age, living situation and involuntary admission
Levine 2010[70]	12 countries	International cohort/ Johnson & Johnson Pharmaceutical Research	Patients with early episode psychosis and	263	Positive and negative symptoms	PANSS	More than six times	2 years	Sum score	Mixed mode latent class	Five: all amelioration	Absence of a diagnosis of schizophrenia,

		and Development	receiving treatment for more than three months							regression modelling		better premorbid functioning, higher cognitive scores, higher PANSS baseline scores
Chang et al 2018[69]	China	Public psychiatric units	Patients with first-episode nonaffective psychosis and not received any antipsychotics more than one week	138	Negative symptoms	HEN	Four times	3 years	Sum score	Latent class growth analysis (LCGA)	Three: minimal-stable, mild-stable, and high-increasing trajectories	Premorbid adjustment, more severe global cognitive impairment, and severe depressive symptoms
Islam et al 2018[58]	Netherlands	Four medical centres (UMCG, UMCM, UMCU, UMCA)/ GROUP cohort study	Patients with nonaffective psychosis, siblings, and controls	1119 patients, 1059 siblings, 586 controls	Neurocognition	8 tools*	Three times	6 years	Gender and age adjusted z-score and then averaging	Group-based trajectory modelling	Five trajectories in patients: severely altered, moderately altered, mildly altered, normal, and high performer Four trajectories in siblings: moderately altered, mildly altered, normal, and high performer	Patients: education, IQ, premorbid functioning, and psychotic symptoms Siblings: age, gender, education, ethnicity, IQ, premorbid functioning, positive symptoms, frequency of psychotic experiences, and all cognitive performances except CPT variance
Stiekema et al 2017[68]	Netherlands	Four medical centres (UMCG, UMCM, UMCU, UMCA)/ GROUP cohort study	Patients with nonaffective psychosis	1,067	Negative symptoms	PANSS	Three times	6 years	Sum score	Group-based trajectory modelling	Four for social amotivation domain: low, decreased low, increased, decreased high Four for expressive deficit domain: low, decreased, increased and high	Age, gender, educational status, ethnicity, marital status, functioning, quality of life, diagnosis, neurocognitive, negative and psosive symptoms
Thomspson et al 2013[30]	USA	University of California, San Diego Advanced Centre in Innovation in Services and Interventions	Old community-dwelling patients with	201 patients and 67 controls	Neurocognition	MDRS	Four time)	3.5 years	Sum score	Latent growth curve model	Three: high and stable, low and modestly declining, low and rapidly declining	Negative symptoms, living situation, years of education, global

	Research (ACISIR)	schizophrenia and controls	cognition

*= Word Learning Task (i.e. immediate recall and delayed recall), Continuous Performance Test-HQ (CPT-HQ) (CPT performance index and CPT variability), WAIS-III Digit Symbol Substitution Test, WAIS-III Information, WAIS-III Calculation, WAIS-III Block Design; PANSS = Positive and Negative Syndrome Scale; SAPS = Scale for the Assessment of Positive Symptoms; SANS=Scale for the Assessment of Negative Symptoms; HEN = High Royds Evaluation of Negativity Scale; MDRS = Mattis Dementia Rating Scale

Evidence from cross-sectional studies

Of the 46 included studies, 31 were cross-sectional studies. The total sample size per study ranges from 62 to 6,137 individuals irrespective of participants diagnosis status. The reported clustering methods were K-means clustering or non-hierarchical analysis[26,59,61-63,75-80], Ward's method or hierarchical analysis[81-86], K-means clustering and Ward's method[32,33,60,87-92], latent class analysis[27,93] and two-step cluster analysis.[94,95] One study[96] identified clusters using a combination of clinical/empirical method, K-means clustering method and Ward's method. The model selection criteria or similarity metrics were visual inspection of dendrogram, Pearson correlation, squared Euclidean distance, agglomeration coefficients, Dunn index, Silhouette width, Duda and Hart index, elbow test, variance explained, inverse scree plot, average proportion of non-overlap, Akaike information criterion (AIC), Bayesian information criterion (BIC), sample size adjusted Bayesian (ABIC), Schwarz's Bayesian information criterion (BIC), Lo–Mendell–Rubin (LMR) test, adjusted LMR and the bootstrap likelihood ratio test (BLRT). Squared Euclidean distance was the most common index used to determine the number of clusters.

Twenty-one studies[32,33,59,60,76,77,79-85,87,89-91,93-96] reported clusters in patients and their siblings based on neurocognitive and/or social cognitive function. Of these 21 studies, 15 studies[33,59,60,76,77,80-85,87,94-96] found out three clusters, five studies[32,79,89,90,93] reported four clusters and one study[91] discovered five clusters of patients and siblings. Clusters were predicted by multidimensional factors including age, gender, socioeconomic status, educational status or years of education, employment status, ethnicity, risky drinking, obstetric complications, family history of mental disorders, premorbid adjustment, premorbid and current IQ, age of illness onset, duration of illness, clinical diagnosis, cortical thickness, neural activity, general psychopathology, severity of positive schizotypy, severity of negative and positive symptoms, neurocognitive and social cognitive performance, anxiety, disorganization, depression, stigma, state mania, neurological soft signs, antipsychotics dosage, adherence to treatment, global functioning, community functioning, socio-occupational functioning and quality of life.[32,33,59,60,76,77,79-85,87,89-91,93-96] (Table 3)

Likewise, two studies[27,88] reported three clusters of patients based on the negative symptom dimension. The predictors were gender, season of birth, ethnicity, years of education, age of illness onset, hospitalization, treatment history, social anhedonia, severity of positive and negative symptoms, general psychopathology, attitude, neurocognitive and social cognitive performance, global functioning, premorbid adjustment, and psychosocial functioning.[27,88] Regarding positive symptoms, only one study[86] identified three clusters of patients and two clusters in the general population based on hallucination symptom. (Table 3)

One study[92] found three clusters of patients based on social cognition and negative symptom that were predicted by marital status, hospitalization, quality of life, severity of negative symptoms and social cognitive performance whereas another study[78] found four clusters of patients based on neurocognition (attention domain) and negative symptom, which were predicted by self-esteem, attention performance, acceptance of stigma, severity of positive and negative symptoms and social functioning. In addition, one study[26] reported three clusters while another study[75] found out four clusters based on positive and negative symptoms

that predicted by IQ, meta-cognition, age of illness onset, global functioning, comorbid diseases, and severity of positive and negative symptoms. (Table 3)

Moreover, three studies[61-63] consistently reported four clusters of unaffected siblings or general population based on positive and negative schizotypy dimensions. The predictors of cluster membership were gender, severity of positive and negative schizotypy, pleasure experiences, somatic symptoms, substance use and abuse, neurocognitive functioning, social functioning, psychotic-like experiences, depression, schizoid and negative symptoms, personality, proneness to positive and negative symptoms, social adjustment and emotional expressivity.[61-63] (Table 3)

Table 3: Characteristics of included cross-sectional studies (n = 31).

Authors' and publication year	Country	Research centre/Cohort	Participants	Sample size	Symptom dimension	Assessment tool	Method of calculating score	Method of clustering	Number of clusters identified	Predictors
Ahmed 2018[27]	USA	Maryland Psychiatric Research Center (MPRC)	Patients with chronic schizophrenia	706	Negative symptom	SDS	Sum score	Latent class analysis with prior hypothesis	Three: deficit, persistent, transient	Sex, season of birth, ethnicity, years of education, illness onset, treatment history, as well as positive symptoms, profiles of cognitive impairment, premorbid adjustment, psychosocial functioning
Bechi 2018[94]	Italy	IRCCS San Raffael Scientific Institute	Patients with stable schizophrenia	452	Global cognition IQ	BACS, WAIS-R	Global cognition: mean score adjusted to age and education IQ: sum score	Two-step cluster analysis (both scores together)	Three for whole sample: high, medium, low Two for subsamples with high pre-morbid IQ: high, medium	Age, years of education, age of onset, negative and positive symptoms, IQ, cognition
Bell 2010[80]	USA	Community mental health center (CMHC)	Patients with schizophrenia or schizoaffective disorder - clinically stable	151	Cognition: memory	HVLT-R	Sum score	K means cluster analysis (with prior hypothesis)	Three: nearly normal, subcortical, cortical	Educational status
Bell 2013[92]	USA	Community mental health center (CMHC)	Patients with stable schizophrenia or schizoaffective disorder	77	Social cognition and negative symptom	SANS, PANSS, MSCEIT	Sum score	K-means cluster analysis plus Ward's method)	Three: high negative symptom, low negative symptom with higher social cognition, low negative symptom with poorer social cognition	Quality of life, hospitalization, marital status, negative symptoms, social cognition
Chang 2015[86]	Korea	Seoul National University Hospital and Boramae Medical Center	Patients with schizophrenia and nonclinical sample	111 patients and 223 nonclinical population	Positive symptom – hallucination	LSHS-R	Sum score	Hierarchical cluster analysis/Ward's method	Three for clinical sample: Cluster 1, Clusters 2, Cluster 3 Two for nonclinical sample: Cluster 1, Cluster 2	Not reported. It explores only clusters
Craddock 2018[26]	USA	National Institute of Mental Health (NIMH)/Childhood-onset schizophrenia (COS) cohort	Patients with childhood-onset schizophrenia (COS)	125	Positive and negative symptoms	SAPS, SANS	Factor score (CFA)	K-means cluster analysis	Three: low positive and negative, high negative low positive, high positive and negative	IQ, age of onset, global functioning, comorbid diseases, positive and negative symptoms

Dawes 2011[91]	USA	University of California/San Diego (UCSD) Advanced Center for Innovation in Services and Interventions Research (ACISIR)	Patients with schizophrenia or schizoaffective disorder	144	Neurocognition	Comprehensive neuropsychological test battery (seven tests)	Sum of deviation scores adjusted to age, gender, education and ethnicity	Hierarchical cluster analysis plus K-means cluster analysis	Five: K1, K2, K3, K4, K5	Educational status, ethnicity
Geisler 2015[79]	USA	Four research centers (MGH, UI, UMN, UNM)/Mind Clinical Imaging Consortium (MCIC) study of schizophrenia	Patients with schizophrenia and healthy controls	129 patients and 165 healthy controls	Neurocognition	Comprehensive neuropsychological test battery (18 tests)	PC score (PCA)	K-means cluster analysis	Four: diminished verbal fluency, diminished verbal memory and poor motor control, diminished face memory and slowed processing, diminished intellectual function	Duration of illness, severity of positive symptoms, years of education, premorbid adjustment, cortical thickness, neural activity
Gilbert 2014[85]	Canada	Institut en santé mentale de Québec	Patients with schizophrenia	112	Neurocognition	Cognitive battery test (more than eight tests)	Average Z-scores	Hierarchical cluster analysis	Three: generally impaired, selectively impaired, near-normal	IQ, gender, socioeconomic status, cognition
Lewandowski 2014[32]	USA	McLean Hospital/Schizophrenia and Bipolar Disorder Program (SBDP)	Patients with psychosis	167	Neurocognition	Neuropsychological battery (five tests)	Z-scores adjusted to age or age and education	Ward's method, plus K-means cluster analysis	Four: globally normal, normal processing speed/executive function, normal visuospatial function, globally impaired	Cognition, age, educational attainment, antipsychotics dosage, severity of positive and negative symptoms, community functioning
Lewandowski 2018[90]	USA	McLean Hospital/Schizophrenia and Bipolar Disorder Program (SBDP)	Patients with psychosis and healthy controls	120 patients and 31 healthy controls	Neurocognition	MATRICS Consensus Cognitive Battery (MCCB) (10 subtests)	Age and gender adjusted T-scores	Ward's method, plus K-means cluster analysis	Four: normal, mildly impaired, moderately impaired, significantly impaired	Educational status, premorbid IQ, state mania, severity of positive and negative symptoms, antipsychotic dosage, cognition, community functioning
Lui et al 2018[63]	China	Castle Peak Hospital	Unaffected first-degree relatives of patients with schizophrenia	194	Schizotypy (positive and negative together)	Chapman Psychosis Proneness Scales (four subscales)	Sum score	K-means clustering analysis	Four: high positive schizotypy, high negative schizotypy, mixed schizotypy, low schizotypy	Severity of positive and negative schizotypy, everyday life pleasure experiences, emotional expressivity
Lysaker et al 2009[78]	USA	Roudebush VA Medical Center and Community Mental Health Center (CMHC)	Patients with stable schizophrenia or schizoaffective disorder and on treatment	99	Negative symptoms and attention (CPT)	PANSS, CPT	Normalized z-scores	K-Means cluster analysis	Four groups: low negative/relatively better attention, low negative/relatively poor attention, high negative/ relatively poor attention, and high negative/relatively better attention	Self-esteem, attention performance, acceptance of stigma, severity of positive and negative symptoms, social functioning

Barrantes-Vidal et al 2010[62]	USA	University of North Carolina at Greensboro (UNCG)	Healthy college students	6,137	Schizotypy (positive and negative together)	Chapman Psychosis-Proneness Scales	Normalized component score (PCA)	K-means clustering analysis	Four: low (nonschizotypic), high positive, high negative, and mixed (high positive and negative) schizotypy	Severity of positive and negative schizotypy, gender, social functioning, psychotic-like experiences, depression, substance use and abuse, schizoid and negative symptoms, personality, social adjustment
Wang et al 2012[61]	China	Neuropsychology and Applied Cognitive Neuroscience Laboratory	Healthy college students	418	Schizotypy (positive and negative together)	Chapman Psychosis-Proneness Scales	Normalized component score (PCA)	K-means clustering analysis	Four: low (nonschizotypic), high positive, high negative, and mixed (high positive and negative) schizotypy	Psychotic-like symptoms, depression, and social function, emotional expression, pleasure experiences, somatic symptoms, neurocognitive functioning, proneness to positive and negative symptoms
Ochoa et al 2013[77]	Spain	Hospital and community psychiatric services	Patients with a first-episode psychosis	62	Neurocognition	Neuropsychological battery (five tests)	Demographically-adjusted score	K-means clustering analysis	Three: higher neurodevelopment contribution, higher genetic contribution, lower neurodevelopment contribution	Neurocognition performance, premorbid IQ, neurological soft signs, premorbid adjustment, family history of mental disorders, obstetric complications
Crouse et al 2018[84]	Australia	Brain and Mind Research Institute	Patients with a psychosis-spectrum illness	135 patients and 50 controls	Neurocognition	Cambridge Neuropsychological Test Automated Battery (CANTAB) (nine tests)	Age-adjusted Z-scores	Ward's hierarchical cluster analysis	Three: normal-range, mixed, grossly-impaired performance	Socio-occupational functioning, neurocognitive performance, gender, diagnosis, risky drinking, employment status, educational status, premorbid IQ, severity of negative symptoms
Sauve et al 2018[33]	Canada	Douglas Mental Health University Institute (DMHUI)/ PEPP-Montreal program	Patients with psychosis (first- and multiple-episode) receive treatment	80 FEP, 121 MEP and 125 healthy controls for both matching	General cognition	CogState Schizophrenia Battery (13 tests)	Composite scores standardized to controls	Hierarchical and K-means cluster analyses	Three: no impairment, generally impaired, intermediately impaired	IQ, severity of positive symptoms, age, years of education, stage of illness, antipsychotics dosage
Ohi et al 2017[59]	Japan	Kanazawa Medical University Hospital/ Kanazawa Medical University	Patients with schizophrenia, relatives and healthy controls	81 patients, 20 relatives, 25 healthy controls	Neurocognition	Brief Assessment of Cognition in Schizophrenia (BACS) neuropsychological battery (six subscales)	Age- and gender-corrected raw scores	K-means clustering analysis	Three: neuropsychologically normal, intermediate impaired, widespread impaired	Clinical diagnosis, neurocognitive performance, years of education, premorbid IQ, antipsychotics dosage
Potter et al	USA	University of	Patients with	73 patients	Neurocognition-	Neuropsychological	Scaled	K-means	Three: intellectually	Severity of negative

2010[76]		Massachusetts	schizophrenia and controls	and 74 controls	IQ and Oral reading	tests (six tests)	scores	clustering analysis	compromised, intellectually deteriorated, intellectually preserved	symptoms, neurocognitive performance, educational status, general psychopathology
Prouteau et al 2017[83]	France	Public psychiatric hospitals	Patients with schizophrenia-spectrum disorders	69	Neurocognitive insight	Objective: Neuropsychological tests (six tests) Subjective: SSTICS	Standardized Z-scores	Ward's method cluster analysis	Three: high cognitive impairment/moderate cognitive complaints, good cognitive functioning/moderate cognitive complaints, moderate cognitive impairment/high cognitive complaints	Age, educational status, severity of negative symptoms, quality of life, anxiety, depression, stigma, neurocognitive performance
Quee et al 2014[60]	Netherlands	UMCG, UMCU, UMCM, UMCA/GROUP cohort	Health siblings of patients with schizophrenia	654	Neurocognition	Neuropsychological battery test (8 tests)	Mean score of gender and age-adjusted z-scores	Ward's method plus K-means clustering	Three: normal, mixed, impaired	Age, educational status, IQ, premorbid adjustment, severity of positive schizotypy
Rangel et al 2015[93]	Colombia	Universities of Antioquia, Pontificia Bolivariana, Nacional of Colombia	Patients with schizophrenia	253	Neurocognition	Neuropsychological tests (five tests)	Not reported	Latent classes analysis	Four: global cognitive deficit, memory and executive function deficit, memory and facial emotion recognition deficit, without cognitive deficit	Gender, age, severity of negative symptoms, global functioning, employment status, adherence to treatment, neurocognitive performance, depression
Reser et al 2015[89]	Australia	Early Psychosis Prevention and Intervention Centre (EPPIC)	Patients with a first-episode psychosis	128	Neurocognition	Comprehensive cognitive battery test (15 tests)	Range standardized test scores	Ward's method cluster analysis plus K-means clustering analysis	Four: cluster 1, cluster 2, cluster 3, cluster 4	IQ (premorbid and current), years of education, severity of negative symptoms, neurocognitive performance
Rocca et al 2016[95]	Italy	Multicentre study/Italian Network for Research on Psychoses (NIRP)	Patients with schizophrenia and controls	809 patients and 780 controls	Social cognition	MCCB (three tests)	Z-scores of scales	Two-step cluster analysis	Three: unimpaired, impaired, very impaired	Age, educational status, cognitive performance, functioning, severity of positive and negative symptoms, disorganization
Rodrigez et al 2017[82]	Czech	National Institute of Mental Health	Patients with first-episode schizophrenia spectrum disorders and healthy controls	28 patients and 91 controls	Neurocognition	Neuropsychological battery tests (15 tests)	Z-scores standardized using controls	Ward`s method cluster analysis	Three: generalized severe, partial mild, near-normal	Neurocognitive performance
Strauss et al 2013[88]	USA	Veterans Affairs Greater Los Angeles Healthcare	Patients with schizophreni	199	Negative symptoms	SANS	Mean factor scores (PCA)	Ward's method	Three: diminished expression, avolition–	General psychopathology, severity of positive and

		System	a					cluster analysis plus K-means clustering analysis	apathy, low negative symptoms	negative symptoms, social anhedonia, attitude, global functioning, social cognition, hospitalization
Trauelsen et al 2016[75]	Denmark	OPUS	Patients with first-episode non-affective psychosis and controls	97 patients and 101 controls	Positive and negative symptoms	Positive and Negative Symptom Scales (PANSS)	Z-scores	K-means cluster analysis	Four: low positive/low negative, high positive/low negative, low positive/high negative, high positive/high negative	Metacognition
Uren et al 2017[87]	Australia	Early Psychosis Prevention and Intervention Centre (EPPIC)	Patients with first episode psychosis and controls	133 patients and 46 controls	General cognition	Comprehensive battery test (14 tests)	Z-scores	Ward's method plus k-means clustering analysis	Three: severe global impairment, moderate impairment, intact	Premorbid IQ, severity of negative symptoms, cognitive performance, years of education, employment status, functioning
Wells et al 20115[96]	Australia	Australian Schizophrenia Research Bank (ASRB)	Patients with schizophrenia or schizoaffective disorder and healthy controls	534 patients and 635 controls	Neurocognition	Neuropsychological tests (five tests)	Z-scores standardized by healthy controls	Ward's method plus K-means clustering analysis and clinical method	Three: preserved, deteriorated, compromised	Age, years of education, age onset of illness, gender, neurocognitive performance, severity of positive and negative symptoms, functioning
Wu et al 2010[81]	Taiwan	Psychiatric rehabilitation hospital	Patients with schizophrenia	76	Neurocognition	Brief Neuropsychological Cognitive Examination (BNCE) (10 subscales)	Mean scores	Ward's method cluster analysis	Three: near normal, deteriorated conceptual thinking, anomia and impaired executive function	Severity of negative symptoms

Summary of clusters/trajectories and predictors

As illustrated in Table 4, two to five clusters/trajectories and 57 predictors were identified by longitudinal and cross-sectional studies in all group of participants across the three symptom dimensions. In patients or siblings or health subjects, three to four subgroups based on the three symptom dimensions were identified irrespective of study design and duration of follow-up. In addition, both longitudinal and cross-sectional studies identified five subgroups only in patients based on all symptom dimensions. In patient clusters or trajectories based on the three symptom dimensions, age, gender, ethnicity, educational status, age of illness onset, diagnosis of schizophrenia, depressive symptoms, general psychopathology, severity of positive and negative symptoms, cognitive performance, premorbid functioning, quality of life and global functioning were important predictors identified by both longitudinal and cross-sectional studies. Likewise, in sibling clusters or trajectories, age, gender, ethnicity, educational status, severity of positive and negative schizotypy, cognitive performance and premorbid functioning were relevant predictors.

Table 4: Summary of clusters/trajectories and predictors

	Participants					Symptom dimensions					Type of study		
	Patients	Siblings	Healthy subjects	Patients and siblings	Patients and healthy controls	Cognitive impairment	Negative symptoms	Positive symptoms	Negative and positive symptoms/ schizotypy	Negative symptoms and cognitive impairment	Longitudinal study		Cross-sectional study
											< 2 years follow-up	≥ 2 years follow-up	
Number of clusters/trajectories													
2			√				√	√				√	√
3	√	√		√	√	√	√	√	√	√	√	√	√
4	√	√	√			√	√		√	√	√	√	√
5	√					√	√	√	√		√	√	√
Predictors													
Sociodemographic													
Age	√	√				√	√	√			√	√	√
Gender	√	√	√			√	√	√	√		√	√	√
Season of birth	√						√						√
Ethnicity	√	√				√	√	√	√		√	√	√
Marital status	√						√			√		√	√
Educational status	√	√		√	√	√	√	√				√	√
Premorbid or current IQ	√	√		√		√			√			√	√
Family history of psychosis or any mental disorders	√					√	√						√
Living situation	√				√		√	√				√	
Employment status	√					√	√					√	√
Socioeconomic status	√					√							√
Clinical													
Cannabis use	√							√			√		
Substance abuse	√		√					√	√			√	
Risky drinking	√					√							√
Stigma (acceptance of)	√					√				√			√
Self-esteem	√									√			√
Pleasure experiences		√	√						√				√
Emotional expression		√	√						√				√
Obstetric complications	√					√							√
Cortical thickness	√					√							√
Neural activity	√					√							√
Age of illness onset	√					√	√	√	√		√		√
Diagnosis (of schizophrenia)	√			√		√	√	√	√		√	√	√
Duration of untreated psychosis	√						√	√			√	√	
Duration of illness	√					√			√		√		√
Frequency of psychotic experiences		√				√						√	

Previous hospitalizations	√							√		√	√	√			√	
Involuntary admission	√							√	√					√		
Extrapyramidal symptoms	√							√	√	√		√				
Depressive symptoms	√		√			√		√	√	√		√		√	√	
Disorganized symptoms	√					√		√						√	√	
State mania	√					√									√	
Attitude	√							√							√	
Personality			√							√					√	
Social anhedonia	√							√							√	
Neurological soft signs	√					√									√	
General psychopathology	√					√		√	√	√		√			√	
Psychotic-like experiences			√							√					√	
Somatic symptoms			√							√					√	
Comorbid diseases	√									√					√	
Types of antipsychotic medication	√							√	√			√				
Antipsychotics dosage	√			√		√									√	
Adherence to treatment	√					√									√	
Treatment history	√							√							√	
Severity of positive and negative symptoms/schizotypy	√	√	√		√	√		√	√	√	√	√	√	√	√	
Severity of positive schizotypy		√				√									√	
Cognitive performance	√	√	√	√		√		√	√	√	√	√	√	√	√	
Meta-cognition	√									√					√	
Premorbid functioning	√	√				√		√	√	√		√	√			
Premorbid adjustment	√	√				√		√				√	√	√		
Social adjustment			√							√					√	
Quality of life	√					√		√	√	√	√	√	√	√	√	
Social functioning	√		√					√		√	√			√	√	
Community functioning	√					√									√	
Socio-occupational functioning	√					√									√	
Psychosocial functioning	√							√							√	
Global functioning	√				√	√		√	√	√				√	√	

Discussion

To our knowledge, this is the first comprehensive systematic review based on recent cluster- and trajectory-based studies of positive symptoms, negative symptoms and cognitive deficits. The reviewed studies involved various groups of study population including patients with first-episode psychosis or chronic schizophrenia, antipsychotics naïve patients or patients who were on antipsychotic treatment for a month or longer, patients from different age groups and ethnicities, and healthy siblings and controls. In this review, we summarized the number of clusters or trajectories, predictors clusters or trajectories and statistical methods based on the evidence in the last decade.

Longitudinal trajectory-based studies discovered two to five patient trajectory groups based on positive and negative symptoms, and four to five patient and sibling trajectory groups identified based on cognitive deficit. Based on cross-sectional cluster-based studies, three clusters of patients were identified for positive and negative symptoms while four clusters of siblings were identified based on positive and negative schizotypy. Regarding cognitive deficits, three to five clusters were reported in patients and their unaffected siblings. Overall, three to four subgroups were discovered across all the three symptom dimensions and study population irrespective of study design and duration of follow-up. This implicates schizophrenia symptoms are inherently heterogeneous and clinicians should treat clients based on group-level or severity of illness instead of using the broad hallmark symptoms. Age, gender, ethnicity, educational status, age of illness onset, diagnosis of schizophrenia, depressive symptoms, general psychopathology, severity of positive and negative symptoms, cognitive performance, premorbid functioning, quality of life and global functioning were important predictors among patients and their unaffected siblings. These factors could be used for developing clinical risk prediction model as well as machine learning.

In this review, we showed that longitudinal studies have used slightly different latent growth modelling methods, such as growth mixture modelling (GMM)[31,65,73], latent class growth analysis (LCGA)[29,30,66,69,71,72,74], mixed mode latent class regression modelling[25,64,70] and group-based trajectory modelling (GBTM).[58,67,68] Latent growth mixture models (LGMMs) are a generalization of linear mixed-effects models (LMEs) to identify categories based on temporal patterns of change by assuming the existence of latent classes or subgroups of subjects exhibiting similarity with regard to unobserved (latent) variables.[29,97] LGMMs often providing more realistic estimates of heterogeneity in longitudinal trajectories. With LGMMs, latent classes are defined as unobserved groups within which the random effects and error terms are normally distributed with constant mean and variance. GMMs have four potential advantages for modelling longitudinal data compared to other methods, such as LMEs. First, it enables flexible, data-driven estimates of the random effect and error distributions separately that can more accurately reflect observed heterogeneity. Second, it allows for classification of individual subjects into latent classes based on the largest probability of class membership. Third, it is sensitive to the pattern of change over time and robust in the presence of missing data. Fourth, subject-level factors can be directly assessed for association with class membership and hence with different trajectory subtypes.[29,30,97]

Among the reviewed 31 cross-sectional studies, we observed that 26 studies identified meaningful clusters using either K-means clustering analysis[26,59,61-63,75-80] or Ward's method clustering analysis[81-86] or both K-means and Ward's method[32,33,60,87-92] clustering analysis. Cluster analysis, applied to investigate symptoms in schizophrenia for several decades, is an atheoretical, data-driven approach to classify people into homogeneous groups by determining clusters of participants that displays less within-cluster variation relative to the between-cluster variation.[84] K-mean cluster analysis is a non-hierarchical form of cluster analysis appropriate when previous evidence or hypotheses exist regarding the number of clusters contained in a sample. It produces the number of clusters initially called for by minimizing variability within clusters and maximizing variability between clusters.[78] In K-means clustering analysis, participants pass through several iterative processes, assigned to a cluster and moved from one cluster to another until terminating conditions are met. Ward's method is a hierarchical cluster analysis aiming to determine group assignment without knowing ahead of time or prior hypothesis.[78] K-means iterative cluster analyses handle larger data sets better than hierarchical agglomerative methods.[62] Longitudinal studies are scarce as shown by our review with only one-third of the included studies being longitudinal. In this circumstance, researchers can conduct cross-sectional studies can be an option and homogenous subgroups can be identified using either K-means or Ward's method clustering analysis methods or both.

Following the rigorous review, we identified several gaps that could be helpful for future neuroscience and behavioral science researchers. First, only two longitudinal studies[30,58] investigated the trajectory of cognitive impairment in patients and siblings which may limit our knowledge regarding the change in cognitive function over time. This may be because neuropsychological assessment is resource intensive, time-consuming, and needs specialized training to collect the data as well as study participants' commitment. Therefore, additional longitudinal studies are warranted to examine the long-term trajectories of cognitive deficits in schizophrenia. Second, we observed limited use of data from siblings and healthy controls to unravel clusters or trajectories, which would be relevant to validate the heterogeneity of clinical symptoms detected in patients. For example, less than ten studies (mostly cross-sectional studies) examined clusters in siblings. Besides, most studies used healthy controls to standardize patients test scores and other few studies used to compare the distribution of controls across patient cluster or trajectory groups. Comparing patient clusters or trajectories with healthy siblings and controls could provide an accurate means of disentangling the causes of heterogeneity of schizophrenia symptoms. Third, even though all reviewed studies identified cluster or trajectory groups based on their own models, subtle differences between-researchers were noted in terms of constructing composite scores, use of model selection criteria and method of parameter estimates. Fourth, only five longitudinal studies evaluated the long-term effect of trajectories on functioning and quality of life. Given the relevance of functioning and quality of life as an outcome measure of treatment effect and prognosis of patients, it is worthwhile to explore the heterogeneity and trajectories. Fifth, we noted several ways of subtyping and nomenclature for subgroups, which may be confusing for clinicians to use the evidence. Therefore, even though statistical subtyping of SCZ based on its symptoms is important, the output must be translatable to clinical practice. Statisticians must also work together with clinicians and create a common understanding. Finally, none of the reviewed

studies used a single or an aggregated effect of genetic susceptibility that potentially helps to accurately predict clusters or trajectories. Genetic markers are believed to be a specific and sensitive biomarker that shows the inherent heterogeneity in the course of illness. Genetic and epigenetic susceptibility may also influence differences between groups, between subjects, or within subjects over time in patients, their siblings and healthy controls. Through identifying subgroups, gene fine-mapping and enrichment analysis can be done for those groups with severe impairment.[55]

Even though the interpretability and validity of findings are debatable, our comprehensive review synthesized the up-to-date evidence from cross-sectional and longitudinal studies that identified data-driven clusters or trajectories. The results of statistical subtyping approaches, such as cluster or trajectory analysis depends on the mathematical assumptions that do not have a direct relationship to clinical reality, type of data, number of variables or tests, sample size and sampling characteristics. Therefore, the results can be unstable and clinical symptoms may not converge on a consistent set of subgroups.[71,90,98] For example, intermediate clusters or trajectories vary substantially from study to study.[90] So that, to ensure comparability and interpretability of identified clusters or trajectories, it is recommended that future psychiatric researchers validate their model using additional comparable statistical method. In our review, we found that nine cross-sectional studies[32,33,60,87-92] cross-validate their model using K-means and Ward's clustering analysis method though none of the reviewed longitudinal studies does compare their model. It is also relevant to combine statistical methods of subtyping with empirical method. To this end, only one study[96] used a combination of clustering and clinical experience to identify homogeneous subgroups. Additionally, replication of cluster or trajectory groups using separate samples, different assessment tools that measure the same construct and different linkage methods (e.g. cluster analysis) is highly relevant for establishing the validity and generalizability of identified subgroups.[33,99]

Identification of meaningful clusters or trajectories with greater homogeneity based on clinical features or endophenotypes, such as neuropsychological markers, neural substrates, and other neurological soft signs among patients, siblings, and healthy subjects require the use of advanced statistical modeling techniques including machine learning approaches. This could facilitate efforts to identify common etiology, examine the patterns of clinical symptoms, understanding the inherent course of the disease and developing new treatment strategies specific to that subgroup to improve recovery and functional outcomes.[29,30,74] In addition, statistical methods that accurately identify clusters or trajectories and describe within and between-variation can help clinicians and statisticians to characterize the relationship of schizophrenia with various clinical and functional outcomes, treatment history, treatment response, and imaging patterns that inform neuropathology. To accomplish these goals, numerous efforts have been undertaken by carefully designing cross-sectional and longitudinal studies, and developing statistical programming language and software.[30] As a result, identification of clinically and statistically relevant clusters or trajectories became promising from time to time.

In general, researchers claim that schizophrenia clinical symptoms are heterogeneous though the existing evidence is still divergent. Further work is expected from psychiatric researchers targeting a longitudinal study design, unaffected siblings and utilizing genetic markers as a predictor. Despite this, our review may help clinicians to optimize the efficacy of evidence-based personalized medicine by providing personalized assessment, initiating early intervention strategies, and by selecting treatments relevant for subgroups of patients with similar characteristics. Our review would also help the translation of the statistical findings into clinical practice and using clustering and trajectory analysis methods in precision medicine to treat subgroup of patients with poor outcome and to prevent prodromal symptoms in their relatives and the general population.

Funding

Tesfa Dejenie supported by the University of Groningen Scholarship

References

1. Jablensky A. The diagnostic concept of schizophrenia: Its history, evolution, and future prospects. *Dialogues Clin Neurosci*. 2010;12(3):271-287.

2. Owen MJ, Sawa A, Mortensen PB. Schizophrenia. *The Lancet*. 2016;388(10039):86-97.

3. Ripke S, O'Dushlaine C, Chambert K, et al. Genome-wide association analysis identifies 13 new risk loci for schizophrenia. *Nat Genet*. 2013;45(10):1150-1159.

4. Hilker R, Helenius D, Fagerlund B, et al. Heritability of schizophrenia and schizophrenia spectrum based on the nationwide danish twin register. *Biol Psychiatry*. 2017.

5. American Psychiatric Association, ed. *Diagnostic and statistical manual of mental disorders : DSM-5.* 5th ed. American Psychiatric Publishing, 1000 Wilson Boulevard, Suite 1825, Arlington, VA 222093901.; 2013.

6. Xavier RM, Vorderstrasse A. Genetic basis of positive and negative symptom domains in schizophrenia. *Biol Res Nurs*. 2017;19(5):559-575.

7. Pai NB, Vella SC. Negative symptoms in schizophrenia: The prevailing challenges. . 2015.

8. Mäkinen J, Miettunen J, Isohanni M, Koponen H. Negative symptoms in schizophrenia—a review. *Nordic Journal of Psychiatry*. 2008;62(5):334-341.

9. Andreasen NC. The scale for the assessment of negative symptoms (SANS): Conceptual and theoretical foundations. *The British Journal of Psychiatry*. 1989;155(S7):49-52.

10. Kay SR, Fiszbein A, Opfer LA. The positive and negative syndrome scale (PANSS) for schizophrenia. *Schizophr Bull*. 1987;13(2):261-276.

11. Kumari S, Malik M, Florival C, Manalai P, Sonje S. An assessment of five (PANSS, SAPS, SANS, NSA-16, CGI-SCH) commonly used symptoms rating scales in schizophrenia and comparison to newer scales (CAINS, BNSS). *J Addict Res Ther*. 2017;8(3):10.4172/2155-6105.1000324. Epub 2017 May 11.

12. Overall JE, Gorham DR. The brief psychiatric rating scale (BPRS): Recent developments in ascertainment and scaling. *Psychopharmacol Bull*. 1988.

13. Palmer BW, Dawes SE, Heaton RK. What do we know about neuropsychological aspects of schizophrenia? *Neuropsychol Rev*. 2009;19(3):365-384.

14. Petrova N, Dorofeikova M. Cognition in schizophrenia: Selective impairment and factors that influence it. *Eur Psychiat*. 2017;41:S193.

15. Shmukler AB, Gurovich IY, Agius M, Zaytseva Y. Long-term trajectories of cognitive deficits in schizophrenia: A critical overview. *Eur Psychiatry*. 2015;30(8):1002-1010.

16. Faraone SV, Kremen WS, Lyons MJ, Pepple JR. Diagnostic accuracy and linkage analysis: How useful are schizophrenia spectrum phenotypes? *Am J Psychiatry*. 1995;152(9):1286-1290.

17. Faraone SV, Tsuang MT, Tsuang DW. *Genetics of mental disorders: A guide for students, clinicians, and researchers.* New York: Guilford Press; 1999.

18. Kremen WS, Seidman LJ, Pepple JR, Lyons MJ, Tsuang MT, Faraone SV. Neuropsychological risk indicators for schizophrenia: A review of family studies. *Schizophr Bull*. 1994;20(1):103-119.

19. Chen WJ, Faraone SV. Sustained attention deficits as markers of genetic susceptibility to schizophrenia. *Am J Med Genet*. 2000;97(1):52-57.

20. Trandafir A, Méary A, Schürhoff F, Leboyer M, Szöke A. Memory tests in first-degree adult relatives of schizophrenic patients: A meta-analysis. *Schizophr Res*. 2006;81(2):217-226.

21. Krukow P, Karakuła-Juchnowicz H, Juchnowicz D, Morylowska-Topolska J, Flis M, Jonak K. Processing speed is associated with differences in IQ and cognitive profiles between patients with schizophrenia and their healthy siblings. *Nord J Psychiatry*. 2017;71(1):33-41.

22. Walker AE, Spring JD, Travis MJ. Addressing cognitive deficits in schizophrenia: Toward a neurobiologically informed approach. *Biol Psychiatry*. 2017;81(1):e1-e3.

23. Kern RS, Nuechterlein KH, Green MF, et al. The MATRICS consensus cognitive battery, part 2: Co-norming and standardization. *Am J Psychiatry*. 2008;165(2):214-220.

24. Nuechterlein KH, Green MF, Kern RS, et al. The MATRICS consensus cognitive battery, part 1: Test selection, reliability, and validity. *Am J Psychiatry*. 2008;165(2):203-213.

25. Levine SZ, Rabinowitz J. Trajectories and antecedents of treatment response over time in early-episode psychosis. *Schizophr Bull*. 2010;36(3):624-632.

26. Craddock KES, Zhou X, Liu S, Gochman P, Dickinson D, Rapoport JL. Symptom dimensions and subgroups in childhood-onset schizophrenia. *Schizophr Res*. 2017.

27. Ahmed AO, Strauss GP, Buchanan RW, Kirkpatrick B, Carpenter WT. Schizophrenia heterogeneity revisited: Clinical, cognitive, and psychosocial correlates of statistically-derived negative symptoms subgroups. *J Psychiatr Res*. 2018;97:8-15.

28. Nagin DS, Odgers CL. Group-based trajectory modeling in clinical research. *Annu Rev Clin Psychol*. 2010;6:109-138.

29. Austin SF, Mors O, Budtz-Jørgensen E, et al. Long-term trajectories of positive and negative symptoms in first episode psychosis: A 10year follow-up study in the OPUS cohort. *Schizophr Res*. 2015;168(1):84-91.

30. Thompson WK, Savla GN, Vahia IV, et al. Characterizing trajectories of cognitive functioning in older adults with schizophrenia: Does method matter? *Schizophr Res*. 2013;143(1):90-96.

31. Case M, Stauffer VL, Ascher-Svanum H, et al. The heterogeneity of antipsychotic response in the treatment of schizophrenia. *Psychol Med*. 2011;41(6):1291-1300.

32. Lewandowski K, Sperry S, Cohen B, Öngür D. Cognitive variability in psychotic disorders: A cross-diagnostic cluster analysis. *Psychol Med*. 2014;44(15):3239-3248.

33. Sauvé G, Malla A, Joober R, Brodeur MB, Lepage M. Comparing cognitive clusters across first-and multiple-episode of psychosis. *Psychiatry Res*. 2018;269:707-718.

34. Shmukler AB, Gurovich IY, Agius M, Zaytseva Y. Long-term trajectories of cognitive deficits in schizophrenia: A critical overview. *Eur Psychiatry*. 2015;30(8):1002-1010.

35. Szoke A, Trandafir A, Dupont ME, Meary A, Schurhoff F, Leboyer M. Longitudinal studies of cognition in schizophrenia: Meta-analysis. *Br J Psychiatry*. 2008;192(4):248-257.

36. Alfimova MV, Kondratiev NV, Golimbet VE. Results and promises of genetics of cognitive impairment in schizophrenia: Molecular-genetic approaches. *Zh Nevrol Psikhiatr Im S S Korsakova*. 2016;116(11):137-144.

37. Misiak B, Stanczykiewicz B, Kotowicz K, Rybakowski JK, Samochowiec J, Frydecka D. Cytokines and C-reactive protein alterations with respect to cognitive impairment in schizophrenia and bipolar disorder: A systematic review. *Schizophr Res*. 2018;192:16-29.

38. Bortolato B, Miskowiak KW, Kohler CA, Vieta E, Carvalho AF. Cognitive dysfunction in bipolar disorder and schizophrenia: A systematic review of meta-analyses. *Neuropsychiatr Dis Treat*. 2015;11:3111-3125.

39. Mesholam-Gately RI, Giuliano AJ, Goff KP, Faraone SV, Seidman LJ. Neurocognition in first-episode schizophrenia: A meta-analytic review. *Neuropsychology*. 2009;23(3):315-336.

40. Snitz BE, MacDonald III AW, Carter CS. Cognitive deficits in unaffected first-degree relatives of schizophrenia patients: A meta-analytic review of putative endophenotypes. *Schizophr Bull*. 2005;32(1):179-194.

41. Bozikas VP, Andreou C. Longitudinal studies of cognition in first episode psychosis: A systematic review of the literature. *Aust N Z J Psychiatry*. 2011;45(2):93-108.

42. Ventura J, Wood RC, Hellemann GS. Symptom domains and neurocognitive functioning can help differentiate social cognitive processes in schizophrenia: A meta-analysis. *Schizophr Bull*. 2011;39(1):102-111.

43. Fett AJ, Viechtbauer W, Dominguez M, Penn DL, van Os J, Krabbendam L. The relationship between neurocognition and social cognition with functional outcomes in schizophrenia: A meta-analysis. *Neuroscience & Biobehavioral Reviews*. 2011;35(3):573-588. doi: https://doi.org/10.1016/j.neubiorev.2010.07.001.

44. Buchanan RW. Persistent negative symptoms in schizophrenia: An overview. *Schizophr Bull*. 2006;33(4):1013-1022.

45. Boutros NN, Mucci A, Diwadkar V, Tandon R. Negative symptoms in schizophrenia: A comprehensive review of electrophysiological investigations. *Clinical schizophrenia & related psychoses*. 2013;8(1):28-35B.

46. Waters F, Fernyhough C. Hallucinations: A systematic review of points of similarity and difference across diagnostic classes. *Schizophr Bull*. 2017;43(1):32-43.

47. Moher D, Liberati A, Tetzlaff J, Altman DG, Prisma Group. Preferred reporting items for systematic reviews and meta-analyses: The PRISMA statement. *PLoS med*. 2009;6(7):e1000097.

48. Beller EM, Glasziou PP, Altman DG, et al. PRISMA for abstracts: Reporting systematic reviews in journal and conference abstracts. *PLoS Med*. 2013;10(4):e1001419.

49. Guyatt GH, Oxman AD, Vist GE, et al. GRADE: An emerging consensus on rating quality of evidence and strength of recommendations. *BMJ*. 2008;336(7650):924-926.

50. Cheah S-, Lurie JK, Lawford BR, Young RM, Morris CP, Voisey J. Interaction of multiple gene variants and their effects on schizophrenia phenotypes. *Compr Psychiatry*. 2016;71:63-70.

51. Cocchi A, Cerati G, Lora A, et al. Patients with first-episode psychosis are not a homogeneous population: Implications for treatment. *Clin Pract Epidemiol Ment Health*. 2014;10:1-8.

52. Hall M-, Smoller JW, Cook NR, et al. Patterns of deficits in brain function in bipolar disorder and schizophrenia: A cluster analytic study. *Psychiatry Res*. 2012;200(2-3):272-280.

53. Kavanaugh BC, Dupont-Frechette JA, Tellock PP, Maher ID, Haisley LD, Holler KA. Neurocognitive phenotypes in severe childhood psychiatric disorders. *J Nerv Ment Dis*. 2016;204(10):770-777.

54. Lim J, Chin R, Ho NF, et al. Elucidation of shared and specific white matter findings underlying psychopathology clusters in schizophrenia. *Asian J Psychiatry*. 2017;30:144-151.

55. Lin S-, Liu C-, Liu Y-, et al. Clustering by neurocognition for fine mapping of the schizophrenia susceptibility loci on chromosome 6p. *Genes Brain Behav*. 2009;8(8):785-794.

56. Nordon C, Rouillon F, Azorin JM, Barry C, Urbach M, Falissard B. Trajectories of antipsychotic response in drug-naive schizophrenia patients: Results from the 6-month ESPASS follow-up study. *Acta Psychiatr Scand*. 2014;129(2):116-125.

57. Silver H, Shmoish M. Analysis of cognitive performance in schizophrenia patients and healthy individuals with unsupervised clustering models. *Psychiatry Res*. 2008;159(1-2):167-179.

58. Islam MA, Habtewold T, van Es F, et al. Long-term cognitive trajectories and heterogeneity in patients with schizophrenia and their unaffected siblings. *Acta Psychiatr Scand*. 2018.

59. Ohi K, Shimada T, Nemoto K, et al. Cognitive clustering in schizophrenia patients, their first-degree relatives and healthy subjects is associated with anterior cingulate cortex volume. *NeuroImage-Clin*. 2017;16:248-256.

60. Quee PJ, Alizadeh BZ, Aleman A, van den Heuvel E, GROUP Investigators. Cognitive subtypes in non-affected siblings of schizophrenia patients: Characteristics and profile congruency with affected family members. *Psychol Med*. 2014;44(2):395-405.

61. Wang Y, Neumann D, Shum DHK, Chan RCK. A cross-validation study of clustering of schizotypy using a non-clinical chinese sample. *Psychiatry Res*. 2012;200(1):55-58.

62. Barrantes-Vidal N, Lewandowski KE, Kwapil TR. Psychopathology, social adjustment and personality correlates of schizotypy clusters in a large nonclinical sample. *Schizophr Res*. 2010;122(1-3):219-225.

63. Lui SSY, Hung KSY, Wang Y, et al. Clustering of schizotypal features in unaffected first-degree relatives of schizophrenia patients. *Schizophr Bull*. 2018.

64. Levine SZ, Rabinowitz J, Faries D, Lawson AH, Ascher-Svanum H. Treatment response trajectories and antipsychotic medications: Examination of up to 18months of treatment in the CATIE chronic schizophrenia trial. *Schizophrenia Research*. 2012;137(1):141-146. doi: https://doi.org/10.1016/j.schres.2012.01.014.

65. Stauffer V, Case M, Kollack-Walker S, et al. Trajectories of response to treatment with atypical antipsychotic medication in patients with schizophrenia pooled from 6 double-blind, randomized clinical trials. *Schizophrenia Research*. 2011;130(1):11-19. doi: https://doi.org/10.1016/j.schres.2011.03.015.

66. Schennach R, Meyer S, Seemüller F, et al. Response trajectories in "real-world" naturalistically treated schizophrenia patients. *Schizophr Res*. 2012;139(1-3):218-224.

67. Pelayo-Teran J, Diaz FJ, Perez-Iglesias R, et al. Trajectories of symptom dimensions in short-term response to antipsychotic treatment in patients with a first episode of non-affective psychosis. *Psychol Med*. 2014;44(1):37-50.

68. Stiekema AP, Islam MA, Liemburg EJ, et al. Long-term course of negative symptom subdomains and relationship with outcome in patients with a psychotic disorder. *Schizophr Res*. 2017.

69. Chang WC, Ho RWH, Tang JYM, et al. Early-stage negative symptom trajectories and relationships with 13-year outcomes in first-episode nonaffective psychosis. *Schizophr Bull*. 2018.

70. Levine SZ, Rabinowitz J, Case M, Ascher-Svanum H. Treatment response trajectories and their antecedents in recent-onset psychosis: A 2-year prospective study. *J Clin Psychopharmacol*. 2010;30(4):446-449.

71. Jäger M, Weiser P, Becker T, et al. Identification of psychopathological course trajectories in schizophrenia. *Psychiatry Res*. 2014;215(2):274-279.

72. Gee B, Hodgekins J, Fowler D, et al. The course of negative symptom in first episode psychosis and the relationship with social recovery. *Schizophr Res*. 2016;174(1):165-171.

73. Chen L, Johnston JA, Kinon BJ, et al. The longitudinal interplay between negative and positive symptom trajectories in patients under antipsychotic treatment: A post hoc analysis of data from a randomized, 1-year pragmatic trial. *BMC Psychiatry*. 2013;13.

74. Abdin E, Chong SA, Vaingankar JA, et al. Trajectories of positive, negative and general psychopathology symptoms in first episode psychosis and their relationship with functioning over a 2-year follow-up period. *PloS one*. 2017;12(11):e0187141.

75. Trauelsen AM, Gumley A, Jansen JE, et al. Metacognition in first-episode psychosis and its association with positive and negative symptom profiles. *Psychiatry Res*. 2016;238:14-23.

76. Potter AI, Nestor PG. IQ subtypes in schizophrenia: Distinct symptom and neuropsychological profiles. *J Nerv Ment Dis*. 2010;198(8):580-585.

77. Ochoa S, Huerta-Ramos E, Barajas A, et al. Cognitive profiles of three clusters of patients with a first-episode psychosis. *Schizophr Res*. 2013;150(1):151-156.

78. Lysaker PH, Vohs JL, Tsai J. Negative symptoms and concordant impairments in attention in schizophrenia: Associations with social functioning, hope, self-esteem and internalized stigma. *Schizophr Res*. 2009;110(1-3):165-172.

79. Geisler D, Walton E, Naylor M, et al. Brain structure and function correlates of cognitive subtypes in schizophrenia. *Psychiatry Res Neuroimaging*. 2015;234(1):74-83.

80. Bell MD, Johannesen JK, Greig TC, Wexler BE. Memory profiles in schizophrenia: Categorization validity and stability. *Schizophrenia Research*. 2010;118(1):26-33. doi: https://doi-org.proxy-ub.rug.nl/10.1016/j.schres.2009.12.037.

81. Wu M, Chan F, Wang T-, Chen S-. Neurocognitive profiles of rehabilitation clients with schizophrenia in taiwan. *J Rehabil*. 2010;76(3):10-14.

82. Rodriguez M, Fajnerová I, Sedláková K, et al. Cluster analysis and correlations between cognitive domains: Cognitive performance in a czech sample of first episodes schizophrenia spectrum disorders - preliminary results. *Psychiatrie*. 2017;21(1):4-11.

83. Prouteau A, Roux S, Destaillats J-, Bergua V. Profiles of relationships between subjective and objective cognition in schizophrenia: Associations with quality of life, stigmatization, and mood factors. *J Cogn Educ Psychol*. 2017;16(1):64-76.

84. Crouse JJ, Moustafa AA, Bogaty SE, Hickie IB, Hermens DF. Parcellating cognitive heterogeneity in early psychosis-spectrum illnesses: A cluster analysis. *Schizophr Res*. 2018;202:91-98.

85. Gilbert E, Mérette C, Jomphe V, et al. Cluster analysis of cognitive deficits may mark heterogeneity in schizophrenia in terms of outcome and response to treatment. *Eur Arch Psychiatry Clin Neurosci*. 2014;264(4):333-343.

86. Chang JS, Kim Y, Kim SH, et al. Differences in the internal structure of hallucinatory experiences between clinical and nonclinical populations. *Psychiatry Res*. 2015;226(1):204-210.

87. Uren J, Cotton SM, Killackey E, Saling MM, Allott K. Cognitive clusters in first-episode psychosis: Overlap with healthy controls and relationship to concurrent and prospective symptoms and functioning. *Neuropsychology*. 2017;31(7):787-797.

88. Strauss GP, Horan WP, Kirkpatrick B, et al. Deconstructing negative symptoms of schizophrenia: Avolition-apathy and diminished expression clusters predict clinical presentation and functional outcome. *J Psychiatr Res*. 2013;47(6):783-790.

89. Reser MP, Allott KA, Killackey E, Farhall J, Cotton SM. Exploring cognitive heterogeneity in first-episode psychosis: What cluster analysis can reveal. *Psychiatry Res*. 2015;229(3):819-827.

90. Lewandowski KE, Baker JT, McCarthy JM, Norris LA, Öngür D. Reproducibility of cognitive profiles in psychosis using cluster analysis. *J Int Neuropsychol Soc*. 2018;24(4):382-390.

91. Dawes SE, Jeste DV, Palmer BW. Cognitive profiles in persons with chronic schizophrenia. *J Clin Exp Neuropsychol*. 2011;33(8):929-936.

92. Bell MD, Corbera S, Johannesen JK, Fiszdon JM, Wexler BE. Social cognitive impairments and negative symptoms in schizophrenia: Are there subtypes with distinct functional correlates? *Schizophr Bull*. 2013;39(1):186-196.

93. Rangel A, Muñoz C, Ocampo MV, et al. Neurocognitive subtypes of schizophrenia. *Actas Esp Psiquiatr*. 2015;43(3):80-90.

94. Bechi M, Spangaro M, Agostoni G, et al. Intellectual and cognitive profiles in patients affected by schizophrenia. *J Neuropsychol*. 2018.

95. Rocca P, Galderisi S, Rossi A, et al. Social cognition in people with schizophrenia: A cluster-analytic approach. *Psychol Med*. 2016;46(13):2717-2729.

96. Wells R, Swaminathan V, Sundram S, et al. The impact of premorbid and current intellect in schizophrenia: Cognitive, symptom, and functional outcomes. *NPJ Schizophr*. 2015;1(1).

97. Muthén B, Shedden K. Finite mixture modeling with mixture outcomes using the EM algorithm. *Biometrics*. 1999;55(2):463-469.

98. Marquand AF, Wolfers T, Mennes M, Buitelaar J, Beckmann CF. Beyond lumping and splitting: A review of computational approaches for stratifying psychiatric disorders. *Biological Psychiatry: Cognitive Neuroscience and Neuroimaging*. 2016;1(5):433-447. doi: https://doi-org.proxy-ub.rug.nl/10.1016/j.bpsc.2016.04.002.

99. Stroebe W, Strack F. The alleged crisis and the illusion of exact replication. *Perspectives on Psychological Science*. 2014;9(1):59-71.