TECHNICAL COMMENTARY

Schizoaffective disorder

Introduction

Schizoaffective disorder is on the schizophrenia spectrum of illnesses. Diagnosis of schizoaffective disorder requires schizophrenia-like symptoms of psychosis, in addition to affective/mood symptoms such as depression. However, there is some debate whether schizoaffective disorder represents a unique diagnosis or an intermediary between schizophrenia and mood disorder[1]. There are also considerable differences among different diagnostic criteria regarding the definition of schizoaffective disorder, particularly the *Diagnostic and Statistical Manual* (DSM) and the *International Classification of Diseases* (ICD) criteria. Specifically, the ICD and also the Research Diagnostic Criteria (RDC) require simultaneous and equally prominent presence of psychotic and affective symptoms; conversely, the DSM requires an additional period (>2 weeks) where the psychotic symptoms alone are present. As a result, the DSM criteria can be applied to fewer patients[2].

Method

We have included only systematic reviews (systematic literature search, detailed methodology with inclusion/exclusion criteria) published in full text, in English, from the year 2000 that report results separately for people with a diagnosis of schizophrenia, schizoaffective disorder, schizophreniform disorder or first episode schizophrenia. Due to the high volume of systematic reviews we have now limited inclusion to systematic meta-analyses. Where no systematic meta-analysis exists for a topic, systematic reviews without meta-analysis are included for that topic. Reviews were identified by searching the databases MEDLINE, EMBASE, CINAHL, Current Contents, PsycINFO and the Cochrane library. Hand searching reference lists of identified reviews was also conducted. When multiple copies of reviews were found, only the most recent version was included.

Review reporting assessment was guided by the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) checklist that describes a preferred way to present a meta-analysis[3]. Reviews with less than 50% of items checked have been excluded from the library. The PRISMA flow diagram is a suggested way of providing information about studies included and excluded with reasons for exclusion. Where no flow diagram has been presented by individual reviews, but identified studies have been described in the text, reviews have been checked for this item. Note that early reviews may have been guided by less stringent reporting checklists than the PRISMA, and that some reviews may have been limited by journal guidelines.

Evidence was graded using the Grading of Recommendations Assessment, Development and Evaluation (GRADE) Working Group approach where high quality evidence such as that gained from randomised controlled trials (RCT) may be downgraded to moderate or low if review and study quality is limited, if there is inconsistency in results, indirect comparisons, imprecise or sparse data and high probability of reporting bias. It may also be downgraded if risks associated with the intervention or other matter under review are high. Conversely, low quality evidence such as that gained from observational studies may be upgraded if effect sizes are large, there is a dose dependent response or if results are reasonably consistent, precise and direct with low associated risks (see end of table for an explanation of these terms)[4]. The resulting table represents an objective summary of the available evidence, although the conclusions are solely the opinion of staff of NeuRA (Neuroscience Research Australia).

Margarete Ainsworth Building, Barker Street, Randwick NSW 2031. Phone: 02 9399 1000. Email: info@neura.edu.au

Results

We found nine systematic reviews that met our inclusion criteria[1, 2, 5-11].

- Moderate to low quality evidence suggests schizoaffective disorder appears to occupy an intermediary position between schizophrenia and mood disorder but is not clearly distinct from either disorder.
- Moderate quality evidence suggests schizoaffective disorder samples diagnosed using RDC/ICD criteria may have had fewer hospitalisations, are more likely to be male and more likely to be older or married than samples using DSM IIIR/IV criteria.
- Moderate quality evidence suggests schizoaffective disorder samples may have more males, less African Americans, more married people, have a longer duration of illness, lower functioning (GAF), better global assessments (GAS), more depression (HDRS) and more negative (SANS) symptoms than samples with schizophrenia
- Compared with bipolar disorder samples, moderate quality evidence suggests schizoaffective disorder samples may be younger, with an earlier age at onset, fewer years of education, with less Caucasians and African Americans, less participants who have ever married, a longer duration of illness, more positive and negative symptoms, more depression, and higher IQ.
- Moderate to high quality evidence suggests the proportion of people with first-episode psychosis that retain a diagnosis of schizoaffective disorder over time is around 72%.
- Moderate quality evidence suggests around 36% of people initially diagnosed with schizoaffective disorder have their diagnosis changed at the second assessment. Conversely, around 55% of people diagnosed with schizoaffective disorder at the second assessment were originally diagnosed with other disorders. Schizophrenia or affective disorders were the most common original or subsequent diagnosis.
- Moderate to high quality evidence suggests lower interrater reliability for diagnosis of schizoaffective disorder than for diagnosis of schizophrenia, bipolar disorder or unipolar depression.
- Moderate to high quality evidence suggests significantly poorer test-retest reliability for schizoaffective disorder diagnosis than for bipolar disorder, schizophrenia or unipolar depression diagnoses.

Margarete Ainsworth Building, Barker Street, Randwick NSW 2031. Phone: 02 9399 1000. Email: info@neura.edu.au

Cheniaux E, Landeira-Fernandez J, Telles LL, Lessa JLM, Dias A, Duncan T, Versiani M

Does schizoaffective disorder really exist? A systematic review of the studies that compared schizoaffective disorder with schizophrenia or mood disorders

Journal of Affective Disorders 2008; 106: 209-217

View review abstract online

Comparison	**Assessment of characteristics of schizoaffective disorder (SCA) compared with schizophrenia (SCZ) and mood disorder (MD).**
Summary of evidence	**Moderate to low quality evidence (direct, unclear sample sizes, unable to assess consistency or precision) suggests schizoaffective disorder appears to occupy an intermediary position between schizophrenia and mood disorder but is not clearly distinct from either disorder.**

Demographic characteristics

Gender

23 of 35 studies reported no difference in the gender distribution in people with SCA compared to SCZ. 32 of 38 studies reported no gender differences between SCA and MD.

12 of 35 studies reported more females with SCA compared to SCZ. 4 of 38 studies found fewer females with SCA compared to MD and 2 of 38 studies found more women with SCA than MD.

Marital status

6 of 11 studies found fewer SCA patients were never married compared to SCZ, 5 studies found no difference.

3 of 12 studies found more SCA patients never married compared to MD patients, though 9 studies found no difference.

Employment

3 of 5 studies found fewer SCA patients were unemployed compared to SCZ, 2 studies found no difference.

1 of 12 studies found more SCA patients were unemployed compared to MD patients, though 4 studies found no difference.

Clinical characteristics

Psychotic symptoms

10 of 20 studies found the intensity/frequency of psychotic symptoms was lower in SCA patients

compared to SCZ, 10 studies found no difference.

6 of 12 studies found the intensity/frequency of psychotic symptoms was higher in SCA patients compared to MD, 6 studies found no difference.

Negative symptoms

8 of 16 studies found the intensity/frequency of negative symptoms was lower in SCA patients compared to SCZ, 8 studies found no difference.

5 of 12 studies found the intensity/frequency of negative symptoms was higher in SCA patients compared to MD, though 8 studies found no difference.

Affective symptoms

4 of 11 studies found the intensity/frequency of affective symptoms was higher in SCA patients compared to SCZ, 7 studies found no difference.

5 of 11 studies found the intensity/frequency of affective symptoms was lower in SCA patients compared to MD, 6 studies found no difference.

Clinical evolution

24 of 39 studies SCA patients had more favourable clinical evolution compared to SCZ, 15 studies found no difference.

21 of 44 studies SCA patients had less favourable clinical evolution compared to MD, 23 studies found no difference.

Treatment response

3 of 6 studies SCA patients had better response to drug treatment compared to SCZ, 3 studies found no difference.

3 of 12 studies SCA patients had worse response to drug treatment compared to MD, 9 studies found no difference.

Functional outcomes

Neuropsychological performance

5 of 19 studies found the severity of cognitive deficit was lower in SCA patients compared to SCZ, though 14 studies found no difference.

5 of 12 studies found the severity of cognitive deficit was higher in SCA patients compared to MD, though 7 studies found no difference.

Insight

2 of 4 studies found SCA patients showed better insight compared to SCZ, 2 studies found no difference in degree of insight deficit.

1 of 4 studies found SCA patients showed better insight compared to MD, though 3 studies found no difference in degree of insight deficit.

Social function

5 of 6 studies found the level of premorbid social adaptation was higher in SCA patients compared to SCZ, though 1 study found no difference.

5 of 6 studies found the level of premorbid social adaptation was lower in SCA patients compared

to MD, though 1 study found no difference. *Age of onset* 2 of 21 studies found SCA patients had a later age of onset compared to SCZ, while 19 studies found no difference. 6 of 23 studies found SCA patients had an earlier age of onset compared to MD, while 17 studies found no difference. *Hospitalisations* 4 of 13 studies found SCA patients had more previous hospitalisations compared to SCZ, while 9 studies found no difference. 5 of 14 studies found SCA patients had more previous hospitalisations compared to MD, while 9 studies found no difference. *Suicide* Seven studies found no differences among the three disorders in frequency of suicide attempts, while 2 studies found more suicide attempts in SCA compared to either SCZ or MD. *Substance use* 5 studies report no difference between disorders in rates of substance use, though 1 study reports higher rates in SCA compared to SCZ, and 1 study reports higher rates in MD compared to SCA.
Physical features
Cortisol and HPA axis function 2 of 2 studies found increased cortisol and associated protein levels in SCA patients compared to SCZ. 3 of 5 studies found no difference in HPA axis protein levels between SCA and MD patients, however 2 studies found lower cortisol levels in SCA patients compared to MD. *Brain structure* 4 of 5 studies found no difference in global brain structure between SCA and SCZ. 3 of 5 studies also found no difference in brain structure between SCA and MD.
Family morbidity
26 studies assessed the risk of SCZ or MD in relatives of patients. *Schizophrenia risk* 4 of 14 studies found that the familial risk of SCZ was lower for relatives of SCA patients compared to relatives of SCZ patients, however 10 studies found no difference in risk of SCZ. 9 of 23 studies found that the familial risk of SCZ was higher for relatives of SCA patients compared to relatives of MD patients, however 14 studies found no difference in risk of SCZ. *Mood disorder risk* 11 of 20 studies found that the familial risk of MD was higher for relatives of SCA patients compared to relatives of SCZ patients, 9 studies found no difference.

4 of 26 studies found that the familial risk of MD was lower for relatives of SCA patients compared to relatives of MD patients, 22 studies found no difference.	
Consistency‡	Unable to assess, appears mostly inconsistent.
Precision§	Unable to assess
Directness‖	Direct

Fusar-Poli P, Cappucciati M, Rutigliano G, Heslin M, Stahl D, Brittenden Z, Caverzasi E, McGuire P, Carpenter WT

Diagnostic stability of ICD/DSM first-episode psychosis diagnoses: meta-analysis

Schizophrenia Bulletin 2016; 42(6): 1395-406

View review abstract online

Comparison	**Diagnostic stability in people with first-episode psychosis using ICD or DSM over 4-5 years.**
Summary of evidence	**Moderate to high quality evidence (unable to assess consistency, precise, direct, large sample) suggests the proportion of people with first-episode psychosis that retain a diagnosis of schizoaffective disorder is around 72% over time.**

Diagnostic stability

42 studies N = 14 484, follow-up average 4.5 years

The proportion of people with first-episode psychosis that retain a diagnosis of schizoaffective disorder is around 72%;

Schizoaffective disorder: 0.72, 95%CI 0.61 to 0.73

Other diagnoses;

Schizophrenia: 0.90, 95%CI 0.85 to 0.95

Schizophreniform disorder: 0.29, 95%CI 0.22 to 0.38

All affective spectrum psychoses: 0.84, 95%CI 0.79 to 0.89

Substance-induced psychotic disorder: 0.66, 95%CI 0.51 to 0.81

Delusional disorder: 0.59, 95%CI 0.47 to 0.71

Acute and transient psychotic disorder/brief psychotic disorder: 0.56, 95%CI 0.62 to 0.60

Psychosis not otherwise specified: 0.36, 95%CI 0.27 to 0.45

About 10% of patients with diagnosis of affective spectrum psychoses changed to schizophrenia

spectrum psychoses over time. Moderator analyses revealed more recent studies and studies of inpatient (vs. mixed) settings had higher diagnostic instability over time. There was no effect on diagnostic stability according to patients' age, gender, substance abuse, baseline ICD-10 or DSM-IV diagnostic criteria, baseline functioning level and duration of follow-up.	
Consistency in results	Unable to assess; no measure of consistency is reported.
Precision in results	Precise
Directness of results	Direct

Pagel T, Franklin J, Baethge C

Schizoaffective disorder diagnosed according to different diagnostic criteria – systematic literature search and meta-analysis of key clinical characteristics and heterogeneity

Journal of Affective Disorders 2014; 156: 111-118

View review abstract online

Pagel T, Baldessarini RJ, Franklin J, Baethge C

Heterogeneity of schizoaffective disorder compared with schizophrenia and bipolar disorder

Acta Psychiatrica Scandinavica 2013; 128: 238-250

View review abstract online

Pagel T, Baldessarini RJ, Franklin J, Baethge C

Characteristics of patients diagnosed with schizoaffective disorder compared with schizophrenia and bipolar disorder

Bipolar Disorders 2013; 15: 229-239

View review abstract online

Comparison 1	**Assessment of patients diagnosed with schizoaffective disorder using narrower (DSM IIIR/IV) vs. broader (RDC/ICD) criteria.**

Schizoaffective disorder

Summary of evidence	**Moderate quality evidence (direct, large samples, some imprecision, unable to assess consistency) suggests schizoaffective disorder samples diagnosed using RDC/ICD criteria may have had fewer hospitalisations, are more likely to be male, and more likely to be older or married than samples using DSM IIIR/IV criteria.**
Demographic characteristics, hospitalisations and symptoms	
18 broad criteria studies (14 RDC and 4 ICD-10, n = 986), and 37 narrow criteria studies (20 DSM-IIIR, and 17 DSM-IV, n = 1269) *Schizoaffective studies employing broader (vs. narrower) criteria report the following:* Fewer previous hospitalizations: 2.2 vs. 5.4, p = 0.016 Older samples: 39.0 vs. 37.7 years, p = 0.035 More males: 51.1% vs. 42.5%, p = 0.002, OR 0.71, 95%CI 0.57 to 0.88 More married participants: 39.7% vs. 20.5%, p = 0.064, OR 0.45, 95%CI 0.17 to 1.13 No significant differences were reported for years of education, age at onset, duration of illness, IQ, or symptomology. Overall, studies employing broader criteria had pooled SDs 13% higher for psychometric scales and 4% higher for clinical and demographic data, indicating more heterogeneous results.	
Consistency	Unable to formally assess, although broader criteria results appear more inconsistent than narrow criteria results.
Precision	Precise for gender, imprecise for marital status, unable to assess other variables.
Directness	Direct
Comparison 2	**Assessment of patients diagnosed with schizoaffective disorder, schizophrenia, and bipolar disorder.**
Summary of evidence	**Moderate quality evidence (direct, large samples, some inconsistency and imprecision) suggests schizoaffective disorder samples may have more males, less African Americans, more married people, have a longer duration of illness, lower functioning (GAF), better global assessments (GAS), more depression (HDRS) and more negative (SANS) symptoms than samples with schizophrenia.** **Compared to bipolar disorder samples, moderate quality evidence suggests schizoaffective disorder samples may be younger, with an earlier age at onset, fewer years of education, with less Caucasians and more African Americans, less participants who have ever married, a longer duration of illness,**

	more positive (BPRS) and negative (SANS) symptoms, more depression (HDRS), and higher IQ (WAIS).
Demographic characteristics, hospitalisations and symptoms	
15 studies used DSM-IIIR, 14 studies used DSM-IV, 4 studies used DSV-III, 11 studies used RDC, 1 study used ICD-9, 1 study used ICD-10, and 4 studies used mixed diagnostic tools. Schizophrenia N = 10814, bipolar disorder N = 4814, schizoaffective disorder N = 2684 *Studies of schizoaffective disorder vs. schizophrenia report:* Less males: 48% vs. 61%, OR 0.57, CI 0.47 to 0.68, $p < 0.0001$, I^2 48%, $p = 0.0003$ Less African Americans: 25% vs. 32%, OR 0.52, CI 0.39 to 0.69, $p < 0.0001$, I^2 47%, $p = 0.08$ More currently married: 37% vs. 12%, OR 3.24, CI 1.12 to 9.39, $p = 0.03$, I^2 79%, $p = 0.0002$ More ever married: 34% vs. 25%, OR 1.61, CI 1.05 to 2.45, $p = 0.03$, I^2 45%, $p = 0.12$ Longer duration of illness: 13.3yrs vs. 11.1yrs, MD 1.77, CI 0.13 to 3.51, $p = 0.03$, I^2 4%, $p = 0.11$ Lower GAF score: 38.6 vs. 43.4, MD 2.15, CI 0.73 to 3.57, $p = 0.003$, I^2 0%, $p = 0.71$ Higher GAS score: 40.0 vs. 35.6, MD 3.00, CI 0.24 to 5.76, $p = 0.03$, I^2 0%, $p = 0.88$ Higher HDRS score: 20.3 vs. 13.3, MD 4.71, CI 0.43 to 8.99, $p = 0.03$, I^2 75%, $p = 0.008$ Lower SANS score: 2.9 vs. 3.3, MD -0.63, CI -1.20 to -0.06, $p = 0.03$, I^2 59%, $p = 0.09$ No significant differences were reported for age at onset, current age, years of education, currently married, Caucasian, number of hospitalizations, age at first hospitalisation, CGI, BDRS, SAPS or WAIS-IQ. *Studies of schizoaffective disorder vs. bipolar disorder report:* Earlier age at onset: 23.3yrs vs. 26.1yrs, MD -2.91, CI -4.52 to -1.29, $p < 0.0004$, I^2 77%, $p < 0.00001$ Younger sample: 42.7yrs vs. 46.7yrs, MD -3.03, CI -4.22 to -1.89, $p < 0.0001$, I^2 59%, $p < 0.00001$ Less education: 12.3yrs vs. 13.3yrs, MD -0.92, CI -1.44 to -0.40, $p = 0.0006$, I^2 22%, $p = 0.25$ Less Caucasians: 52% vs. 60%, OR 0.52, CI 0.40 to 0.69, $p < 0.0001$, I^2 0%, $p = 0.50$ More African Americans: 25% vs.13%, OR 1.50, CI 1.02 to 2.21, $p < 0.04$, I^2 59%, $p = 0.02$ Less ever married: 34% vs. 41%, OR 0.63, CI 0.43 to 0.93, $p = 0.02$, I^2 10%, $p = 0.35$ Longer duration of illness: 13.3yrs vs. 11.5yrs, MD 2.10, CI 0.10 to 4.09, $p = 0.04$, I^2 56%, $p = 0.03$ Higher BPRS score: 46.6 vs. 37.8, MD 3.85, CI 1.94 to 5.87, $p < 0.0001$, I^2 0%, $p = 0.48$ Higher HDRS score: 20.3 vs. 10.8, MD 7.01, CI 1.67 to 12.36, $p = 0.01$, I^2 80%, $p = 0.002$ Higher SANS score: 3.3 vs. 0.9, MD 0.85, CI 0.14 to 1.55, $p = 0.02$, I^2 76%, $p = 0.02$ Higher WAIS-IQ score: 105.5 vs.103.7, MD -7.31, CI -10.22 to -4.08, $p = 0.001$, I^2 0%, $p = 0.64$ No significant differences were reported for gender, currently married, number of hospitalizations, age at first hospitalisation, CGI, GAS, GAF, or SAPS. Overall, SDs tended to be larger in bipolar disorder than in schizophrenia or schizoaffective studies,	

indicating higher heterogeneity in bipolar disorder results, although this finding was not-significant.	
Consistency	Consistent for ever married, duration of illness, GAF, GAS in schizoaffective disorder vs. schizophrenia comparison. Consistent for education, Caucasians, ever married, BPRS and WAIS-IQ in schizoaffective disorder vs. bipolar disorder comparison.
Precision	Precise for gender, African Americans GAS in schizoaffective disorder vs. schizophrenia comparison. Precise for Caucasians in schizoaffective disorder vs. schizophrenia comparison.
Directness	Direct

Salamon S, Santelmann H, Franklin J, Baethge C

Test-retest reliability of the diagnosis of schizoaffective disorder in childhood and adolescence - A systematic review and meta-analysis

Journal of Affective Disorders 2018; 230: 28-33

View review abstract online

Comparison	**Test-retest reliability of a schizoaffective disorder diagnosis in children and adolescents compared to a diagnosis of bipolar disorder, schizophrenia, or unipolar depression.** **Time between test and retest ranged from 1 to 16 years (mean 7.2 years).**
Summary of evidence	**Moderate quality evidence (some inconsistency and imprecision, imprecise, direct, large sample) suggests lower test-retest reliability for schizoaffective disorder than for schizophrenia, bipolar disorder and unipolar depression.**
Test-retest reliability	
7 studies, N = 403 *Lower test-retest reliability for schizoaffective disorder than for schizophrenia, bipolar disorder and unipolar depression;* Schizoaffective disorder: 7 studies, Cohen's kappa = 0.27, 95%CI 0.07 to 0.47, I^2 = 91% Schizophrenia: 7 studies, Cohen's kappa = 0.56, 95%CI 0.29 to 0.83, I^2 = 94% Bipolar disorder: 5 studies, Cohen's kappa = 0.64, 95%CI 0.55 to 0.74, I^2 = 0% Unipolar depression: 3 studies, Cohen's kappa = 0.66, 95%CI 0.52 to 0.81, I^2 = 0%	

Consistency in results	Inconsistent for schizoaffective disorder and schizophrenia.
Precision in results	Appears mostly imprecise
Directness of results	Direct

Santelmann H, Franklin J, Bushoff J, Baethge C

Diagnostic shift in patients diagnosed with schizoaffective disorder: A systematic review and meta-analysis of rediagnosis studies

Bipolar Disorders 2016; 18: 233-46

View review abstract online

Comparison	**Diagnostic stability of schizoaffective disorder over two time points (median = 2 years).**
Summary of evidence	**Moderate quality evidence (inconsistent, appears imprecise, direct, large sample) suggests around 36% of people initially diagnosed with schizoaffective disorder have their diagnosis changed at the second assessment. Conversely, around 55% of people diagnosed with schizoaffective disorder at the second assessment were originally diagnosed with other disorders. Schizophrenia and affective disorders were the most commonly switched to/from diagnoses.**

Diagnostic stability

31 studies, N = 1,866

~36% (95%CI 24% to 49%, I^2 = 89%) of people diagnosed with schizoaffective disorder at the first assessment were rediagnosed at second assessment; ~19% (95%CI 13% to 28%, I^2 = 80%) switched to schizophrenia, ~14% (95%CI 10% to 19%, I^2 = 53%) switched to affective disorders, and the rest switched to other disorders.

~55% (95%CI 46 to 63%, I^2 = 79%) of people diagnosed with schizoaffective disorder at the second assessment were originally diagnosed with other disorders; ~18% (95%CI 13% to 25%, I^2 = 76%) switched from schizophrenia, ~24% (95%CI 18% to 30%, I^2 = 66%) switched from affective disorders and the rest switched from other disorders.

Consistency in results	Inconsistent
Precision in results	Appears imprecise
Directness of results	Direct

Santelmann H, Franklin J, Busshoff J, Baethge C

Interrater reliability of schizoaffective disorder compared with schizophrenia, bipolar disorder, and unipolar depression - A systematic review and meta-analysis

Schizophrenia Research 2016; 176: 357-63

View review abstract online

Comparison	**Interrater reliability of diagnosis of schizoaffective disorder compared to schizophrenia, bipolar disorder and unipolar depression.**
Summary of evidence	**Moderate to high quality evidence (inconsistent, appears precise, direct, large sample) suggests lower interrater reliability for diagnosis of schizoaffective disorder than for diagnosis of schizophrenia, bipolar disorder or unipolar depression.**

Interrater reliability

25 studies, N = 7,912

Interrater reliability for schizoaffective disorder is lower than for schizophrenia, bipolar disorder and unipolar depression;

Schizoaffective disorder: Cohen's kappa = 0.57, 95%CI 0.41 to 0.73, I^2 = 98%

Schizophrenia: Cohen's kappa = 0.80, 95%CI 0.76 to 0.84, I^2 = 70%

Bipolar disorder: Cohen's kappa = 0.82, 95%CI 0.77 to 0.86, I^2 = 38%

Unipolar depression: Cohen's kappa = 0.75, 95%CI 0.70 to 0.81, I^2 = 82%

These results did not change according to diagnostic or kappa method used, sample size, number of differential diagnoses, or year of publication. There was no evidence of publication bias.

Consistency in results	Mostly inconsistent
Precision in results	Appears precise
Directness of results	Direct

Santelmann H, Franklin J, Bushoff J, Baethge C

Test-retest reliability of schizoaffective disorder compared with schizophrenia, bipolar disorder, and unipolar depression-a systematic

review and meta-analysis

Bipolar Disorders 2015; 17: 753-68

View review abstract online

Comparison	**Test-retest reliability of a schizoaffective disorder diagnosis compared to a diagnosis of bipolar disorder, schizophrenia, or unipolar depression.**
Summary of evidence	**Moderate to high quality evidence (inconsistent, appears precise, direct, large sample) suggests significantly poorer test-retest reliability for schizoaffective disorder diagnosis than for bipolar disorder, schizophrenia or unipolar depression diagnoses.**

Test-retest reliability

49 studies, N = 14,314

Significantly lower test-retest reliability for schizoaffective disorder than for bipolar disorder, schizophrenia or unipolar depression;

Schizoaffective disorder: 48 studies, Cohen's kappa = 0.50, 95%CI 0.40 to 0.59, I^2 = 96%

Bipolar disorder: 33 studies, Cohen's kappa = 0.77, 95%CI 0.73 to 0.82, I^2 = 92%

Schizophrenia: 42 studies, Cohen's kappa = 0.69, 95%CI 0.64 to 0.74, I^2 = 90%

Unipolar depression: 35 studies, Cohen's kappa = 0.73, 95%CI 0.66 to 0.79, I^2 = 91%

Assessment of differences between diagnoses;

Schizoaffective disorder vs. schizophrenia: MD = 0.21, 95%CI 0.13 to 0.28, $p < 0.001$

Schizoaffective disorder vs. bipolar disorder: MD = 0.23, 95%CI 0.14 to 0.33, $p < 0.001$

Schizoaffective disorder vs. unipolar depression: MD = 0.18, 95%CI 0.10 to 0.27, $p < 0.001$

In studies of schizoaffective disorder, kappa was significantly lower in consistent vs. inconsistent use of diagnostic interview. There were no differences in kappa in studies according to; similar vs. different rater identity; first-episode vs. chronic illness; inpatient vs. outpatient status; low vs. high risk of bias studies (including blinded vs. non-blinded studies); studies using ICD-10 diagnostic tool vs. DSM 111, DSM 1V or DSM 5 diagnostic tools; or studies with a short vs. long follow-up period (< 2 months vs. > 12 months).

Consistency in results	Inconsistent
Precision in results	Appears precise
Directness of results	Direct

Schizoaffective disorder

Explanation of acronyms

BPRS = Brief Psychiatric Rating Scale, CGI = Clinical Global Impression, CI = confidence interval, DSM = American Psychiatric Association's Diagnostic and Statistical Manual, GAF = Global Assessment of Functioning, GAS = Global Assessment Scale, HDRS = Hamilton Depression Rating Scale, I^2 = the percentage of the variability in effect estimates that is due to heterogeneity rather than sampling error (chance), ICD = World Health Organisation's International Classification of Diseases, MD = mood disorder and mean difference, OR = odds ratio, *p* = probability of rejecting a null hypothesis of no differences between groups, RDC = Research Diagnostic Criteria, SANS = Scale for Assessment of Negative Symptoms, SAPS = Scale for Assessment of Positive Symptoms SCA = schizoaffective disorder, SCZ = schizophrenia, SD = standard deviation, WAIS-IQ = Wechsler Adult Intelligence Scale-Intelligence Quotient

Margarete Ainsworth Building, Barker Street, Randwick NSW 2031. Phone: 02 9399 1000. Email: info@neura.edu.au

Explanation of technical terms

* Bias has the potential to affect reviews of both RCT and observational studies. Forms of bias include; reporting bias – selective reporting of results; publication bias - trials that are not formally published tend to show less effect than published trials, further if there are statistically significant differences between groups in a trial, these trial results tend to get published before those of trials without significant differences; language bias – only including English language reports; funding bias - source of funding for the primary research with selective reporting of results within primary studies; outcome variable selection bias; database bias - including reports from some databases and not others; citation bias - preferential citation of authors. Trials can also be subject to bias when evaluators are not blind to treatment condition and selection bias of participants if trial samples are small[12].

† Different effect measures are reported by different reviews.

Prevalence refers to how many existing cases there are at a particular point in time. Incidence refers to how many new cases there are per population in a specified time period. Incidence is usually reported as the number of new cases per 100,000 people per year. Alternatively some studies present the number of new cases that have accumulated over several years against a person-years denominator. This denominator is the sum of individual units of time that the persons in the population are at risk of becoming a case. It takes into account the size of the underlying population sample and its age structure over the duration of observation.

Reliability and validity refers to how accurate the instrument is. Sensitivity is the proportion of actual positives that are correctly identified (100% sensitivity = correct identification of all actual positives) and specificity is the proportion of negatives that are correctly identified (100% specificity = not identifying anyone as positive if they are truly not).

Weighted mean difference scores refer to mean differences between treatment and comparison groups after treatment (or occasionally pre to post treatment) and in a randomised trial there is an assumption that both groups are comparable on this measure prior to treatment. Standardsed mean differences are divided by the pooled standard deviation (or the standard deviation of one group when groups are homogenous) that allows results from different scales to be combined and compared. Each study's mean difference is then given a weighting depending on the size of the sample and the variability in the data. 0.2 represents a small effect, 0.5 a moderate effect, and 0.8 and over represents a large effect[12].

Odds ratio (OR) or relative risk (RR) refers to the probability of a reduction (< 1) or an increase (> 1) in a particular outcome in a treatment group, or a group exposed to a risk factor, relative to the comparison group. For example, a RR of 0.75 translates to a reduction in risk of an outcome of 25% relative to those not receiving the treatment or not exposed to the risk factor. Conversely, a RR of 1.25 translates to an increased risk of 25% relative to those not receiving treatment or not having been exposed to a risk factor. A RR or OR of 1.00 means there is no difference between groups. A medium effect is considered if RR > 2 or < 0.5 and a large effect if RR > 5 or < 0.2[13]. lnOR stands for logarithmic OR where a lnOR of 0 shows no difference between groups. Hazard ratios

Margarete Ainsworth Building, Barker Street, Randwick NSW 2031. Phone: 02 9399 1000. Email: info@neura.edu.au

measure the effect of an explanatory variable on the hazard or risk of an event.

Correlation coefficients (eg, r) indicate the strength of association or relationship between variables. They can provide an indirect indication of prediction, but do not confirm causality due to possible and often unforseen confounding variables. An r of 0.10 represents a weak association, 0.25 a medium association and 0.40 and over represents a strong association. Unstandardised (*b*) regression coefficients indicate the average change in the dependent variable associated with a 1 unit change in the independent variable, statistically controlling for the other independent variables. Standardised regression coefficients represent the change being in units of standard deviations to allow comparison across different scales.

‡ Inconsistency refers to differing estimates of effect across studies (i.e. heterogeneity or variability in results) that is not explained by subgroup analyses and therefore reduces confidence in the effect estimate. I^2 is the percentage of the variability in effect estimates that is due to heterogeneity rather than sampling error (chance) - 0% to 40%: heterogeneity might not be important, 30% to 60%: may represent moderate heterogeneity, 50% to 90%: may represent considerable heterogeneity and over this is considerable heterogeneity. I^2 can be calculated from Q (chi-square) for the test of heterogeneity with the following formula[12];

$$I^2 = \left(\frac{Q - df}{Q}\right) \times 100\%$$

§ Imprecision refers to wide confidence intervals indicating a lack of confidence in the effect estimate. Based on GRADE recommendations, a result for continuous data (standardised mean differences, not weighted mean differences) is considered imprecise if the upper or lower confidence limit crosses an effect size of 0.5 in either direction, and for binary and correlation data, an effect size of 0.25. GRADE also recommends downgrading the evidence when sample size is smaller than 300 (for binary data) and 400 (for continuous data), although for some topics, these criteria should be relaxed[14].

‖ Indirectness of comparison occurs when a comparison of intervention A versus B is not available but A was compared with C and B was compared with C that allows indirect comparisons of the magnitude of effect of A versus B. Indirectness of population, comparator and/or outcome can also occur when the available evidence regarding a particular population, intervention, comparator, or outcome is not available and is therefore inferred from available evidence. These inferred treatment effect sizes are of lower quality than those gained from head-to-head comparisons of A and B.

Margarete Ainsworth Building, Barker Street, Randwick NSW 2031. Phone: 02 9399 1000. Email: info@neura.edu.au

References

1. Cheniaux E, Landeira-Fernandez J, Lessa Telles L, Lessa JL, Dias A, Duncan T*, et al.* (2008): Does schizoaffective disorder really exist? A systematic review of the studies that compared schizoaffective disorder with schizophrenia or mood disorders. *Journal of Affective Disorders* 106: 209-17.
2. Pagel T, Franklin J, Baethge C (2014): Schizoaffective disorder diagnosed according to different diagnostic criteria - Systematic literature search and meta-analysis of key clinical characteristics and heterogeneity. *Journal of Affective Disorders* 156: 111-8.
3. Moher D, Liberati A, Tetzlaff J, Altman DG, PRISMAGroup (2009): Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *British Medical Journal* 151: 264-9.
4. GRADEWorkingGroup (2004): Grading quality of evidence and strength of recommendations. *British Medical Journal* 328: 1490.
5. Pagel T, Baldessarini RJ, Franklin J, Baethge C (2013): Characteristics of patients diagnosed with schizoaffective disorder compared with schizophrenia and bipolar disorder. *Bipolar Disorders* 15: 229-39.
6. Pagel T, Baldessarini RJ, Franklin J, Baethge C (2013): Heterogeneity of schizoaffective disorder compared with schizophrenia and bipolar disorder. *Acta Psychiatrica Scandinavica* 128: 238-50.
7. Fusar-Poli P, Cappucciati M, Rutigliano G, Heslin M, Stahl D, Brittenden Z*, et al.* (2016): Diagnostic Stability of ICD/DSM First Episode Psychosis Diagnoses: Meta-analysis. *Schizophrenia Bulletin* 42: 1395-406.
8. Santelmann H, Franklin J, Bushoff J, Baethge C (2016): Diagnostic shift in patients diagnosed with schizoaffective disorder: A systematic review and meta-analysis of rediagnosis studies. *Bipolar Disorders* 18: 233-46.
9. Santelmann H, Franklin J, Busshoff J, Baethge C (2016): Interrater reliability of schizoaffective disorder compared with schizophrenia, bipolar disorder, and unipolar depression - A systematic review and meta-analysis. *Schizophrenia Research* 176: 357-63.
10. Santelmann H, Franklin J, Bushoff J, Baethge C (2015): Test-retest reliability of schizoaffective disorder compared with schizophrenia, bipolar disorder, and unipolar depression--a systematic review and meta-analysis. *Bipolar Disorders* 17: 753-68.
11. Salamon S, Santelmann H, Franklin J, Baethge C (2018): Test-retest reliability of the diagnosis of schizoaffective disorder in childhood and adolescence - A systematic review and meta-analysis. *Journal of Affective Disorders* 230: 28-33.
12. CochraneCollaboration (2008): Cochrane Handbook for Systematic Reviews of Interventions. Accessed 24/06/2011.
13. Rosenthal JA (1996): Qualitative Descriptors of Strength of Association and Effect Size. *Journal of Social Service Research* 21: 37-59.
14. GRADEpro (2008): [Computer program]. Jan Brozek, Andrew Oxman, Holger Schünemann. *Version 32 for Windows*.