TECHNICAL COMMENTARY

Early detection

Introduction

Detection of the early stages of bipolar disorder may help develop interventions that prevent or delay the onset of the disorder. Should the disorder develop, early intervention can result in improved clinical outcomes. To achieve early detection, accurate identification of individuals at highest risk of onset of symptoms is paramount.

Potential high risk states include having a family history of bipolar disorder and/or subclinical mood symptoms. As many people with bipolar disorder have a depressive onset, this is also a risk state.

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 2010 that report results separately for people with a diagnosis of bipolar or related disorders. 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, and PsycINFO. Hand searching reference lists of identified reviews was also conducted. When multiple reviews assessing the same topic were found, only the most recent and/or comprehensive were 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[1]. 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 (RCTs) 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)[2]. 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).

Results

We found four systematic reviews that met our inclusion criteria[3-6].

- Moderate to high quality evidence suggests large effects of having psychotic symptoms or a family history of bipolar disorder as risk factors for transition to bipolar disorder in people with major depression. There was a medium-sized effect of higher risk with earlier rather than later age of onset of depression, and a small effect of having a family history of any mood disorder. The risk of transition to bipolar disorder was greatest in the early stages of depression (up to 5 years).

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

Early detection

- Moderate quality evidence suggests subclinical symptoms preceding an initial mood episode last around 27 months, and subclinical symptoms preceding a recurrent mood episode last around 1 month.
- Common subclinical symptoms (in order of decreasing prevalence) are; too much energy, diminished ability to think, indecisiveness, pressured speech, being talkative, elated mood, academic or work difficulties, insomnia, and depressed mood.
- Less common subclinical symptoms (in order of decreasing prevalence) are; over-productive/goal directed behaviour, agitation, rage attacks, racing thoughts, anxiety, decreased need for sleep, irritable mood, fatigue, distractibility, sleep disturbance, disinhibition, weight loss/loss of appetite, hyperactivity, suicidal thoughts, feeling of worthlessness, mood swings, delusions, unkempt or bizarre appearance, feelings of guilt, and auditory hallucinations.
- Rare subclinical symptoms (in order of decreasing prevalence) are loss of interest, somatic complaints, being over-sensitive, hypersexuality, flight of ideas, hypersomnia, weight gain, self-harm, suicide attempts, and visual hallucinations.
- Low quality evidence is unable to determine the accuracy of instruments for early detection. Authors concluded that the Child Behavioral Checklist – Pediatric Bipolar Disorder Phenotype and the General Behavioral Inventory – Revised had the better validity and utility than the Hypomanic Personality Scale, Behavioral Activation Scale or the Family History Scale, and that more studies are required.

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

Ratheesh A, Berk M, Davey CG, McGorry PD, Cotton SM

Instruments that prospectively predict bipolar disorder - A systematic review

Journal of Affective Disorders 2015; 179: 65-73

View review abstract online

Comparison	**Accuracy of instruments for early detection of bipolar disorder.**
Summary of evidence	**Low quality evidence (small samples, unable to assess consistency, mostly imprecise, direct) is unable to determine the accuracy of instruments for early detection.** **Authors concluded that the Child Behavioral Checklist – Pediatric Bipolar Disorder Phenotype and the General Behavioral Inventory – Revised had the better validity and utility than the Hypomanic Personality Scale, Behavioral Activation Scale, or the Family History Scale, and that more studies are required.**

Instruments for early detection

Child Behavior Checklist (CBCL)

A large, significant effect of increased risk of developing bipolar disorder over 23 years with a high score on the pediatric bipolar disorder phenotype subscale (cut-off 60);

1 study, N = 101 offspring of people with bipolar, unipolar, or no disorder

OR = 8.75, 95%CI 1.89 to 40.58, $p = 0.01$

A small, significant effect of increased risk of developing bipolar disorder over 8 years with a high score on the mania subscale (cut-off 1 SD from population mean);

1 study, N = 2230 representative community participants

OR = 1.06, 95%CI 1.03 to 1.08, $p < 0.001$

General Behavior Inventory – Revised (GBI-R)

A small, significant effect of increased risk of developing bipolar disorder over 5-year follow-up with a high score on the depression subscale (cut-off 21);

1 study, N = 140 offspring of people with bipolar disorder

OR = 1.12, 95%CI 1.04 to 3.85, $p < 0.01$

No significant effect of increased risk of developing bipolar disorder over 5-year follow-up with a high score on the mania/biphasic subscale (cut-off 5);

1 study, N = 140 offspring of people with bipolar disorder

OR = 0.44, 95%CI 0.18 to 1.05, $p = 0.99$

Hypomanic Personality Scale (HPS)

A large, significant effect of increased risk of developing bipolar disorder over 13-year follow-up with

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

a high score on the total scale (cut-off >35); 1 study, N = 80 university students OR = 21.76, 95%CI 1.21 to 391.41, *p* = 0.04 <u>Behavioral Activation Scale (BAS)</u> *A small, significant effect of increased risk of developing bipolar disorder over 4.5-year follow-up with a high score on the fun-seeking subscale;* 1 study, N = 57 university studies with bipolar not otherwise specified or cyclothymia OR = 1.40, 95%CI 1.05 to 1.87, *p* = 0.02 <u>Family History Scale (FHS)</u> *A medium-sized, significant effect of increased risk of developing bipolar disorder over 5-year follow-up with a high score on the depression subscale;* 1 study, N = 140 children with bipolar not otherwise specified OR = 2.69, 95%CI 0.99 to 7.31, *p* = 0.05 *A small, significant effect of increased risk of developing bipolar disorder over 5-year follow-up with a high score on mania subscale;* 1 study, N = 140 children with bipolar not otherwise specified OR = 2.58, 95%CI 1.30 to 5.12, *p* = 0.006	
Consistency in results‡	Not applicable; one study per subscale.
Precision in results§	Precise for CBCL mania subscale only.
Directness of results‖	Direct

Ratheesh A, Davey C, Hetrick S, Alvarez-Jimenez M, Voutier C, Bechdolf A, McGorry PD, Scott J, Berk M, Cotton SM

A systematic review and meta-analysis of prospective transition from major depression to bipolar disorder

Acta Psychiatrica Scandinavica 2017; 135: 273-84

View review abstract online

Comparison	**Predicting transition to bipolar disorder in people with major depression.**
Summary of evidence	**Moderate to high quality evidence (large samples, consistent, mostly imprecise, direct) suggests large effects of having psychotic symptoms or a family history of bipolar disorder as risk factors for transition to bipolar disorder in people with major depression. There was a medium-sized effect of higher risk with earlier rather than later age of depression onset, and a small**

	effect of having a family history of any mood disorder. The risk of transition was greatest in the early stages of depression (up to 5 years).
Predictors for transition to bipolar disorder	
Psychotic symptoms (large effect): 5 studies, N = 1,326, OR = 4.76, 95%CI 1.79 to 12.66, I^2 = 23.5% Family history of bipolar disorders (large effect): 7 studies, N = 3,231, OR = 2.89, 95%CI 2.01 to 4.14, I^2 = 13% Early age of onset of depression (medium-sized effect): 6 studies, N = 9,920, *g* = -0.32, 95%CI -0.42 to -0.21, $p < 0.05$, I^2 = 15% Family history of mood disorders (small effect): 5 studies, OR = 1.65, N = 2,602, 95%CI 1.34 to 2.04, I^2 = 0% Authors report that although longer length of study follow-up was associated in higher prevalence of bipolar disorder, the risk of transition was greatest in the early stages of depression (up to 5 years). Gender, family history of depressive disorders, and antidepressant use had no significant effect on transition rates.	
Consistency in results	Consistent
Precision in results	Precise for early age of onset only.
Directness of results	Direct

Scott J, Kallestad H, Vedaa O, Sivertsen B, Etain B

Sleep disturbances and first onset of major mental disorders in adolescence and early adulthood: A systematic review and meta-analysis

Sleep Medicine Reviews 2021; 57: 101429

View review abstract online

Comparison	**Sleep disturbance prior to a first episode of bipolar disorder.** **Follow-up periods ranged from 4-12 years.**
Summary of evidence	**Moderate to high quality evidence (large sample, consistent, imprecise, direct) finds a small effect of increased rates of first-episode bipolar disorder in people with prior sleep disturbance compared to people with no prior sleep disturbance (insomnia and/or hypersomnia).**
Sleep disturbance	

A small effect of increased risk of first-episode bipolar disorder in people with prior sleep disturbance compared to people with no prior sleep disturbance; 5 studies, N = 4,332, OR = 1.72, 95%CI 1.41 to 2.12, p = 0.05, I^2 = 11% Authors report this effect was marginally higher than for depressive disorders (OR = 1.62). Subgroup analysis showed both insomnia and hypersomnia were present prior to a first episode of bipolar disorder.

Consistency in results	Consistent
Precision in results	Imprecise
Directness of results	Direct

Van Meter AR, Burke C, Youngstrom EA, Faedda GL, Correll CU

The Bipolar Prodrome: Meta-Analysis of Symptom Prevalence Prior to Initial or Recurrent Mood Episodes

Journal of the American Academy of Child and Adolescent Psychiatry 2016; 55: 543-55

View review abstract online

Comparison	**Prevalence of premorbid subclinical symptoms.**
Summary of evidence	**Moderate quality evidence (large samples overall, inconsistent, appears imprecise, direct) suggests subclinical symptoms preceding an initial mood episode lasts around 27 months, and preceding a recurrent mood episode lasts around 1 month.** **Common subclinical symptoms (in order of decreasing prevalence70-50%); too much energy, diminished ability to think, indecisiveness, pressured speech, talkative, elated mood, academic or work difficulties, insomnia, and depressed mood.** **Less common subclinical symptoms (in order of decreasing prevalence 50-20%); over-productive/goal directed, agitation, rage attacks, racing thoughts, anxiety, decreased need for sleep, irritable mood, fatigue, distractibility, sleep disturbance, disinhibition, weight loss/loss of appetite, hyperactive, suicidal thoughts, feeling of worthlessness, mood lability, delusions, appearance (unkempt, bizarre appearance, guilt, and auditory hallucinations.** **Rare subclinical symptoms (in order of decreasing prevalence 20-0%) loss of interest, somatic complaints, sensitive, hypersexuality, flight of ideas, hypersomnia, weight gain, self-harm, suicide attempts, and visual hallucinations.**

Subclinical symptoms
11 studies, N = 1,078, found the prodrome preceding an initial mood episode lasted 27.1 ± 23.1 months (range 4.6 to 130 months).
10 studies, N = 1,000, found subthreshold symptoms preceding a recurrent mood episode lasted 1.0 ± 0.9 months (range 0.5 to1.3 months).
Premorbid symptoms in order of prevalence (most common to least common);
Too much energy: 2 studies, prevalence = 68%, 95%CI 26% to 98%
Diminished ability to think: 3 studies, prevalence = 63%, 95%CI 47% to 78%
Indecisiveness: 1 study, prevalence = 62%, 95%CI 47% to 76%
Pressured speech: 1 study, prevalence = 60%, 95%CI 40% to 78%
Talkative: 2 studies, prevalence = 60%, 95%CI 26% to 89%
Elated mood: 2 studies, prevalence = 58%, 95%CI 45% to 71%
Academic or work difficulties: 3 studies, prevalence = 56%, 95%CI 40% to 72%
Insomnia: 3 studies, prevalence = 54%, 95%CI 35% to 72%
Depressed mood: 9 studies, prevalence = 53%, 95%CI 39% to 66%
Over-productive/goal directed: 1 study, prevalence = 50%, 95%CI 36% to 64%
Agitation: 3 studies, prevalence = 46%, 95%CI 10% to 84%
Rage attacks: 4 studies, prevalence = 45%, 95%CI 22% to 69%
Grandiosity: 5 studies, prevalence = 45%, 95%CI 22% to 69%
Racing thoughts: 5 studies, prevalence = 45%, 95%CI 18% to 72%
Anxiety: 9 studies, prevalence = 43%, 95%CI 29% to 58%
Decreased need for sleep: 6 studies, prevalence = 42%, 95%CI 27% to 59%
Irritable mood: 7 studies, prevalence = 42%, 95%CI 26% to 58%
Fatigue: 4 studies, prevalence = 41%, 95%CI 14% to 71%
Distractibility: 3 studies, prevalence = 37%, 95%CI 1% to 84%
Sleep disturbance: 7 studies, prevalence = 36%, 95%CI 15% to 60%
Disinhibited: 4 studies, prevalence = 36%, 95%CI 27% to 46%
Weight loss/loss of appetite: 3 studies, prevalence = 36%, 95%CI 21% to 52%
Hyperactive: 3 studies, prevalence = 35%, 95%CI 28% to 41%
Suicidal thoughts: 3 studies, prevalence = 34%, 95%CI 20% to 50%
Feeling of worthlessness: 3 studies, prevalence = 34%, 95%CI 25% to 43%
Mood lability: 10 studies, prevalence = 32%, 95%CI 20% to 45%
Delusions: 4 studies, prevalence = 23%, 95%CI 9% to 41%
Appearance (unkempt, bizarre) : 1 study, prevalence = 23%, 95%CI 6% to 45%

Early detection

Guilt: 2 studies, prevalence = 22%, 95%CI 13% to 32% Auditory hallucinations: 3 studies, prevalence = 22%, 95%CI 14% to 31% Loss of interest: 2 studies, prevalence = 19%, 95%CI 0% to 62% Somatic complaints: 3 studies, prevalence = 18%, 95%CI 11% to 28% Sensitive: 3 studies, prevalence = 18%, 95%CI 1% to 47% Hypersexuality: 5 studies, prevalence = 17%, 95%CI 6% to 33% Flight of ideas: 1 study, prevalence = 17%, 95%CI 3% to 38% Hypersomnia: 2 studies, prevalence = 17%, 95%CI 0% to 51% Weight gain: 2 studies, prevalence = 16%, 95%CI 9% to 25% Self-harm: 3 studies, prevalence = 12%, 95%CI 0% to 47% Suicide attempts: 4 studies, prevalence = 6%, 95%CI 2% to 12% Visual hallucinations: 1 study, prevalence = 0%, 95%CI 0% to 11%	
Consistency in results	Authors report data are largely inconsistent.
Precision in results	Appears imprecise.
Directness of results	Direct

Explanation of acronyms

CI = confidence interval, I^2 = the percentage of the variability in effect estimates that is due to heterogeneity rather than sampling error (chance), N = number of participants, OR = odds ratio, *p* = statistical probability of obtaining that result ($p < 0.05$ generally regarded as significant), vs. = versus

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[7].

† 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 randomized trial there is an assumption that both groups are comparable on this measure prior to treatment. Standardized 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 treatment effect[7].

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, an 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. An 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[8]. lnOR stands for logarithmic OR where a lnOR of 0 shows no difference between groups. Hazard ratios 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 are an indication of

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

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 dependent 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 treatment 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 substantial heterogeneity and 75% to 100%: considerable heterogeneity. I^2 can be calculated from Q (chi-square) for the test of heterogeneity with the following formula;

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

§ Imprecision refers to wide confidence intervals indicating a lack of confidence in the effect estimate. Based on GRADE recommendations, a result for continuous data 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, this criteria should be relaxed[9].

‖ 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 so is 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. 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.
2. GRADEWorkingGroup (2004): Grading quality of evidence and strength of recommendations. *British Medical Journal* 328: 1490.
3. Ratheesh A, Berk M, Davey CG, McGorry PD, Cotton SM (2015): Instruments that prospectively predict bipolar disorder - A systematic review. *Journal of Affective Disorders* 179: 65-73.
4. Ratheesh A, Davey C, Hetrick S, Alvarez-Jimenez M, Voutier C, Bechdolf A*, et al.* (2017): A systematic review and meta-analysis of prospective transition from major depression to bipolar disorder. *Acta Psychiatrica Scandinavica* 135: 273-84.
5. Van Meter AR, Burke C, Youngstrom EA, Faedda GL, Correll CU (2016): The Bipolar Prodrome: Meta-Analysis of Symptom Prevalence Prior to Initial or Recurrent Mood Episodes. *Journal of the American Academy of Child and Adolescent Psychiatry* 55: 543-55.
6. Scott J, Kallestad H, Vedaa O, Sivertsen B, Etain B (2021): Sleep disturbances and first onset of major mental disorders in adolescence and early adulthood: A systematic review and meta-analysis. *Sleep Medicine Reviews* 57: 101429.
7. CochraneCollaboration (2008): Cochrane Handbook for Systematic Reviews of Interventions. Accessed 24/06/2011.
8. Rosenthal JA (1996): Qualitative Descriptors of Strength of Association and Effect Size. *Journal of Social Service Research* 21: 37-59.
9. GRADEpro (2008): [Computer program]. Jan Brozek, Andrew Oxman, Holger Schünemann. *Version 32 for Windows*