OXFORD UNIVERSITY PRESS | Schizophrenia Bulletin

Bipolar at-risk criteria: an examination of which clinical features have optimal utility for identifying youth at risk of early transition from depression to bipolar disorders.

Journal:	*Schizophrenia Bulletin*
Manuscript ID	SZBLTN-ART-16-0369.R1
Manuscript Type:	Regular Article
Date Submitted by the Author:	n/a
Complete List of Authors:	Scott, Jan; Institute of Neuroscience, Academic Psychiatry Marwaha, Steven; University of Warwick, Mental Health and Wellbeing Ratheesh, Aswin; OrygenYouth Health Macmillan, Iain; NTW NHS Trust Yung, Alison; University of Manchester, Institute of Brain, Behaviour and Mental Health Morriss, Richard; Institute of Mental Health, Hickie, Ian; The University of Sydney, Brain & Mind Research Institute Bechdolf, Andreas; University of Cologne, Department of Psychiatry and Psychotherapy
Keywords:	bipolar, screening, case finding, at risk, uktra high risk, depression, youth

Bipolar at-risk criteria: an examination of which clinical features have optimal utility for identifying youth at risk of early transition from depression to bipolar disorders.

Jan Scott*, Institute of Neuroscience, Newcastle University, and IPPN, King's College, London, UK.

Steven Marwaha, Reader in Psychiatry, Division of Health Sciences, Warwick Medical School, The University of Warwick, UK.

Aswin Ratheesh, Honorary Clinical Lecturer, Orygen, the National Centre of Excellence in Youth Mental Health, and Centre for Youth Mental Health, University of Melbourne, Parkville, Australia

Iain Macmillan, Consultant Psychiatrist, Early Intervention in Psychosis Service, NTW NHS Trust, Newcastle, UK.

Alison R. Yung, Professor of Psychiatry, Faculty of Biology, Medicine and Health, University of Manchester, and Greater Manchester West NHS Mental Health Foundation Trust, Manchester, UK.

Richard Morriss, Professor of Psychiatry, Institute of Mental Health, University of Nottingham, Nottingham, UK.

Ian B. Hickie, Professor and Co-Director, Brain and Mind Centre, The University of Sydney, Australia.

Andreas Bechdolf, Professor, Department of Psychiatry and Psychotherapy, University of Cologne, Cologne, and Department of Psychiatry, Psychotherapy and Psychosomatics, Vivantes Klinikum am Urban and Vivantes Klinikum im Friedrichshain, and Academic Hospital of Charity Medicine Berlin, Germany; and Orygen, the National Centre of Excellence in Youth Mental Health, Parkville, Australia.

*corresponding author: Professor Jan SCOTT, Institute of Neuroscience, Newcastle University, Newcastle upon Tyne, United Kingdom

Tel/Fax: +44 191 2563209

Email: jan.scott@newcastle.ac.uk

Abstract=250
Word count=3963 (excluding references, but including abstract, tables, and main text)
Tables=3
References (main text) = 40
Supplementary Material= 1 Appendix

Key words: screening, case finding, youth, bipolar disorder, ultra-high-risk, at-risk criteria, validity, clinical utility index, number needed to screen

Acknowledgements:

We wish to acknowledge that data collected by Helen Ivatt (when she was a medical student at Newcastle University) was an important component of this study. Despite many attempts we were not able to make contact with Dr Ivatt directly.

Declarations:

Ethical approval for this study came from the Newcastle Joint Hospitals and University Ethics Committee and the Newcastle & North Tyneside NRES Committee (12/NE/0325).

Andreas Bechdolf has received speaker fees and travel support from Otsuka, Janssen and Lundbeck. AB has received research support from the German Research Foundation (DFG), NARSAD, the German Ministry of Education and Research (BMBF), the Faculty of Medicine of the University of Cologne and for an investigator initiated trial from Bristol Myers Squibb and Janssen-Cilag.

Ian B Hickie is a Commissioner in Australia's National Mental Health Commission; a Member of the Medical Advisory Panel for Medibank; a Board Member of Psychosis Australia Trust. IH has received honoraria for presentations of his own work at educational seminars supported by a number of non-government organisations and by the pharmaceutical industry (including Servier, Pfizer, AstraZeneca and Eli Lilly). The University of Sydney (Principal Investigator: IH) received funding from Servier for a study of major depression and sleep disturbance in primary care settings. Other relevant funding for IH in relation to this study includes 'Testing and delivering early interventions for young people with depression' (APP ID: 1046899).

Alison R Yung has received honoraria for presentations of her own work from Janssen, Otsuka and Sunovion.

Iain Macmillan has received grant funding from Northumberland, Tyne and Wear NHS Foundation Trust Research Capacity Funding for a project entitled 'Longitudinal Evaluation of Affective Psychoses Symptoms'.
Steven Marwaha: has received awards from the Mental Health Research Network (UK), and Coventry and Warwickshire Partnership Trust to fund research related to mood instability.

Richard Morriss has received grant funding from the UK National Institute of Health Research (NIHR), European Union and the UK Medical Research Council for research into bipolar disorder. His salary is partly funded by NIHR Collaboration for Leadership in Applied Health Research and Care East Midlands currently to do such work. He chaired the National Institute for Care Excellence Guideline for Bipolar Disorder.

Aswin Ratheesh has no declarations or conflicts in relation to this paper.

Jan Scott has received grant funding from the Stanley Foundation (for work on lithium and medication adherence), from the Medical Research Council UK (including for projects on bipolar II disorders, on CBT and on actigraphic monitoring in bipolar disorders) and from the Research for Patient Benefit programme UK (PB-PG-0609-16166: Early identification and intervention in young people at risk of mood disorders). She has not received any Pharma funding in the last five years.

Abstract (words 250)
Background:

A clinical and research challenge is to identify which depressed youth are at risk of "early transition to bipolar disorders (ET-BD)". This two-part study (i) examines the clinical utility of previously reported BD at-risk (BAR) criteria in differentiating ET-BD cases from unipolar depression controls (UP); and (ii) estimates the Number Needed to Screen (NNS) for research and general psychiatry settings.

Methods: 50 cases with reliably ascertained, early transition to BD I and II cases were matched for gender and birth year with 50 UP who did not develop BD over 2 years during two years of prospective follow-up. We estimated the clinical utility for finding true cases and screening out non-cases for selected risk factors and their NNS. Using a convenience sample (N=80), we estimated the NNS when adjustments were made to account for data missing from clinical case notes.

Results: Sub-threshold mania, cyclothymia, family history of BD, atypical depression symptoms and probable antidepressant-emergent elation), occurred significantly more frequently in ET-BD youth. Each of these 'BAR-Depression' criteria demonstrated clinical utility for screening out non-cases. Only cyclothymia demonstrated good utility for case finding in research settings; sub-threshold mania showed moderate utility. In the convenience sample, the NNS for each criterion ranged from ~4 to 7.

Conclusions: Cyclothymia showed the optimum profile for case finding, screening and NNS in research settings. However, its presence or absence was only reported in 50% of case notes. Future studies of ET-BD instruments should distinguish which criteria have clinical utility for case finding versus screening.

Introduction

Globally, the peak age at onset (AAO) of severe mental disorders such as bipolar (BD) and psychotic disorders is late adolescence and early adulthood[1-3]. Since the turn of the century, researchers have begun to identify subgroups of help-seeking youth who are at ultra-high risk (UHR) of early transition from a late prodromal stage to a first episode of psychosis (FEP). The risk of transition varies between about 15% and 35% over 12-24 months, but it can be predicted by UHR criteria, namely the presence of a combination of a limited set of state, trait and familial characteristics[4-5]. Furthermore, these features can be incorporated into screening tools that can be applied in a range of settings. This enables the early identification of UHR individuals who can be monitored prospectively through a critical period of enhanced risk for the onset of FEP and offered clinical interventions if appropriate[6].

In keeping with the UHR concept in psychotic disorders, several tools to identify young people at increased risk of BD have been applied in research settings, specialist clinics and tertiary referral centres[7-11]. To date, the only instrument with published data on predictive validity in the peak AAO group is the BD at-risk (BAR) assessment tool[7]. The BAR has good reliability (free range kappa 0.83), and incorporates generic risk factors (e.g. being in the peak AAO range for BD onset) alongside a set of specific criteria, namely: cyclothymia co-occurring with depression, sub-threshold mania, and depression co-occurring with genetic risk (i.e. a family history of BD). A case note audit of 173 systematically assessed referrals to the ORYGEN early intervention services (in Melbourne, Australia) found that one in seven youth aged 15-24 years met criteria for at least one BAR subgroup (BAR+). In BAR+ cases, the transition rate to (hypo)mania was about 23% over an observation period of about 250 days compared with 0.7% in the BAR- controls[7]. In a further prospective study undertaken by the same research group (35 BAR+ cases matched with 35 BAR- controls), early transition to BD (14%) occurred in BAR+ cases only[12]. A sub-analysis (N=52) of this case-control study demonstrated that sub-threshold mania was the most significant predictor of early transition to BD in those youth with common mental disorders, such as depression and anxiety[13].

Identifying BAR or UHR-BD criteria is clinically important, but published studies on all the available instruments have been limited to the centres where the assessment tools originated. The generalizability of BAR criteria to other clinical settings and locations, and the clinical utility and discriminant validity of the proposed criteria warrants further examination in larger samples of early transition to BD (ET-BD) cases. Critically, it is important to determine whether the BAR criteria can distinguish which youth with early onset depression are at risk of ET-BD

(i.e. who show transition within about two years)[5-6]. Also, information is needed on how the BAR assessment tool might perform in day to day clinical practice, where the quality of case note recordings may be sub-optimal, and ratings of the presence or absence of specific signs and symptoms may be less reliable than in specialist or research settings that employ systematic assessments. This two-part study aims to address the following:

First, using a case-control methodology we examine the:

a) Discriminant validity of the BAR criteria in differentiating ET-BD youth from unipolar depression controls (UP), who were matched for gender and year of birth
b) Clinical utility of the original BAR criteria and of five additional clinical features for finding BD cases and for screening out non-cases
c) Number Needed to Screen (NNS) using comprehensive, systematic assessments undertaken in research settings.

Second, using a convenience sample we estimate:

d) the NNS if screening of case notes is undertaken in routine clinical practice.

Methods

The Appendix provides further detailed descriptions of the sampling, rationale for selection of risk factors, additional information on statistics and other basic data.

The methods for Part I and Part II of the study are briefly summarized below.

Part I: Case-control study

Sample- This sample comprised of 100 cases (50 ET-BD; 50 UP) who were identified from de-identified data from systematic, comprehensive, clinical assessments that had previously been entered into eight databases designated appropriate for data sharing (in accordance with the recommendations of the Organisation for Economic Co-operation and Development[14]).

The 50 ET-BD cases were selected if they met the following criteria: (a) there was reliable evidence that the first episode of DSM IV[15] mania or hypomania occurred between 15-25 years and (b) that the first (hypo)manic episode occurred within two years of a major depressive episode. These ET-BD cases were frequency matched for gender and year of birth to individuals with a diagnosis of UP. The 50 UP controls also met the criteria of reliable evidence that they had experienced a major depressive episode that met DSM IV criteria between the ages of 15-25 years. The key characteristics of the final sample are shown in Table 1.

Measures: Extended Bipolar at Risk (BAR) criteria - Bechdolf et al's[7] BAR criteria explore the presence or absence of four clinical characteristics prior to the onset of the first episode of (hypo)mania. These four variables are used to identify three at risk subgroups ([1] sub-threshold mania; [2] depression and cyclothymic features; [3] depression and genetic risk of BD (family history of BD). These criteria were supplemented by assessment of the presence or absence of five other clinical features that may be risk factors for the onset of BD I or II (see Appendix for details): probable antidepressant-emergent elation; psychotic symptoms during a mood episode; psychomotor retardation; atypical depression (anergia and/or hypersomnia); and family history of (i) multiple generations (>=2) affected by mood disorders, or (ii) other mood or alcohol and substance misuse disorders (ASUD).

Statistical Analysis- As described in the Appendix, we used several established approaches to examine the statistical significance and clinical utility of the selected risk factors in differentiating between cases with early transition to BD and UP controls. We focus on-

a) Odds Ratio and 95% Confidence Intervals (95% CI) was calculated for each clinical characteristic.

   A review of original studies and meta-analyses of youth with mood disorders indicated that most of the clinical features being tested occurred in at least one in ten participants (except multi-generational family history, for which we could not identify a reliable prevalence rate). Thus, assuming an overall sample prevalence of at least 10% for each variable, we estimated that the size of the sample gave 90% power at a 5% level to detect an OR >=1.98 in the matched case-control analyses.

b) Clinical Utility Index (CUI): Mitchell notes that when using a risk factor or symptom to find true cases or screen out non-cases, the real world clinical relevance of any item will be reduced if it arises infrequently[16]. As such, the Clinical Utility Index (CUI) is increasingly recommended as an alternative to sensitivity or positive predictive values etc. (data shown in the Appendix), as it reflects both the discriminatory ability of a factor or criteria and its overall occurrence in the population being examined[17]. The CUI+ (=Positive Predictive Value x Sensitivity) represents an estimate of the utility of a symptom or risk factor in case finding (the Rule In accuracy). The CUI- (= Negative Predictive Value x Specificity) reflects the utility for screening out non-cases (the Rule Out accuracy).

We calculated the CUI+ and CUI- for each feature selected and report the scores according to Mitchell's[17] grading of utility: poor (0-0.2), fair (0.21-0.39), moderate (0.4-0.59), good (0.6-0.79), or excellent (> 0.8). If the CUI+ exceeds the CUI-, an item is regarded as better for case finding; if the CUI- exceeds the CUI+, the item is better for screening. We report the overall CUI for those factors where either the CUI+ or CUI-were graded as good.

c) Number Needed to Screen (NNS): similar to the Number Needed to Treat, the NNS represents the number of patients that need to be screened to yield one additional, correct identification of a case or non-case, *beyond* those who are misidentified[18,19]. The NNS were estimated for each BAR criterion with significant OR and 95% CIs (see Appendix for the formula).

Part II: Convenience sample study

Sample- With ethical approval, a convenience sample of 80 cases of DSM IV mood disorders (40 individuals with UP and 40 with BD) attending general psychiatry outpatient clinics was identified[20].

Measures-The case notes were screened using an itemized checklist and the frequency with which key clinical information was recorded as present or as absent was noted (see Appendix for details). Data on the prevalence of missing information was extracted[20].

Statistical Analysis: NNS for routine clinical practice- the NNS for each BAR criterion with a good CUI+ or CUI- (in the case-control study) was re-calculated to take into account the rates of missing information in the clinical case notes.

Results

As shown in Table 1, the ET-BD cases and UP controls were more likely to be female (62%), with similar AAO for minor and major mood episodes. There were marginal group differences in number of prior mood episodes (BD>UP) or comorbidity rates (UP>BD).

Table 1 about here

As shown in Table 2, the OR estimates demonstrate that the risk factors that best discriminated between ET-BD cases and UP controls were, in order of magnitude: sub-threshold mania (OR: 16.9; 95% CI: 4.7, 61.8), cyclothymia (OR: 14.2; 95% CI: 5.4, 37.2), atypical depression (OR: 11.5; 95% CI: 3.6, 36.7), family history of BD (OR: 7.6; 95% CI: 1.6, 35.9), and evidence of probable antidepressant-emergent elation (OR: 3.4; 95% CI: 1.2, 4.9).

Table 2 about here

The CUI scores showed that cyclothymia had a good CUI+ (0.62) and CUI- (0.62) grading for discriminating ET-BD cases from UP controls. Also, these gradings suggest that cyclothymia has clinical utility for both case finding and screening. Other items had better utility for screening. Sub-threshold mania had a moderate CUI+ (0.46) and a good CUI- (0.62) grading, whilst probable antidepressant-emergent elation demonstrated a fair CUI+ (0.22) but a good CUI- grading (0.66). Family history of BD had a relatively poor CUI+ grading (0.20), but a moderate CUI- grading (0.54).

Table 3 about here

Table 3 shows that the NNS estimates for the five selected BAR items. In systematically assessed cases the NNS ranged from 1.7 (for cyclothymia) through to 5.0 (for family history of BD). In the convenience sample, the predicted NNS ranged from 3.5 to 6.9. The NNS for family history showed the smallest difference between research and clinical settings (rising from 5.0 to 5.9), reflecting the fact that the presence or absence of this criterion was routinely recorded in clinical practice (84% of case notes).

Discussion

Increasing attention is being given to the identification of youth at risk of a first onset episode of BD[21]. Previous research on screening for BD suggested that self-rating instruments, e.g. the Mood Disorders Questionnaire[22], may help to identify pre-existing, unrecognized cases of BD in older adults. However, they cannot be recommended for use as a screening instrument in individuals at risk of early transition from depression to BD[23-25]. As such, the need to better identify ET-BD has led to the appearance of several new 'BAR' instruments that differ in terms of time for completion, complexity and comprehensiveness[7-11]. We chose to examine Bechdolf et al's[7] BAR assessment tool, which has the benefit of brevity, established reliability, and emerging evidence of predictive validity. The present study builds on the research on the BAR instrument

and its criteria in four important ways. First, we applied the criteria in a new clinical setting and used them for the first time outside of the location where the assessment tool was developed and tested. Second, we assessed a set of extended BAR criteria in a larger number of ET-BD cases than previously studied (the largest subgroup reported was 35 BAR+ cases)[7,12,13]. Third, we identified two additional features, antidepressant-emergent elation and atypical depression, that may enhance the utility of the BAR tool to identify ET-BD in cases of major depression aged 15-25 years. Fourth, we use easily interpretable parameters for describing the performance of each criterion, as the CUI and NNS are easier to understand and potentially more relevant to the planning of screening or case finding than other measures such as odds ratios, sensitivity, specificity, or positive and negative predictive value[16-19].

This study found that the three original BAR items plus two additional variables (a set which we will refer to as BAR-Depression or BAR-D criteria) occurred significantly more often in ET-BD cases compared to UP controls, and that these trait, state and familial markers demonstrated moderate to good clinical utility for screening out non-cases. The NNS for each criterion was highly acceptable for research settings (about 2-5). Although the NNS for each criterion was slightly higher in routine clinical settings (range about 4-7) this finding seems to parallel the original case note audit by Bechdolf et al[7] that revealed that one in seven youth met at least one BAR criterion. Overall, the BAR-D items show lower utility for case finding, and only cyclothymia and sub-threshold manic symptoms showed good or moderate capacity to differentiate ET-BD from UP. Although the current performance of sub-threshold mania was modest, it has previously been found to be a significant predictor of imminent transition to mania in a small scale study using BAR criteria in Australia[13] and a large scale study of offspring of bipolar parents in the USA[26].

The current study suggests that cyclothymia has the optimum profile for case finding, screening and NNS. However, the apparent clinical utility of cyclothmia must be counterbalanced by two observations. First, whilst this temperamental feature can be reliably defined and assessed in the research datasets (e.g. using established personality assessment schedules), its presence or absence was not reported in half of the clinical case notes examined (in the convenience sample). Second, whilst systematic clinical assessments can usually discriminate between cyclothymia and sub-threshold manic symptoms (and other forms of affective instability), it is not clear whether these trait and state phenomena are dependably differentiated in routine clinical practice[25,27]. Third, these variables may co-occur at a rate that is greater than previously anticipated[28]. Taking all these issues into account, we suggest that an important implication of

the current study is that clinicians may need help to develop their skills in detecting cyclothymia as well as encouragement to routinely record its presence or absence in youth with depression.

The findings on family history of BD are worthy of further discussion. Clinical and research evidence suggests that family history of BD is one of the most robust predictors of future onset BD[29,30]; and we found that clinicians recorded information about family history more than any other risk factor for BD. However, the present study confirms that the overall prevalence of a positive family history of BD in general psychiatry datasets is lower that reported in specialist clinics and research environments[31]. Furthermore, the recent National Institute of Healthcare and Clinical Excellence (NICE) guideline on BD suggests that the presence of family history of BD in cases of depression should not be used to identify potential risk of BD[32] as it predicts both recurrent UP as well as BD[33] and genetic loading for BD alone may not be sufficiently discriminatory[33.] Also, recent research suggests that other factors, e.g. AAO of BD in a parent, may play a role in heritability and the likelihood of early onset in offspring[34]. Given these data, our finding that family history of BD in depressed youth is better for screening than for case finding seems to be a conservative, but realistic proposition.

Antidepressant emergent elation appeared to demonstrate sufficient clinical utility for use as a screening item. However, as noted in the Appendix, a significant problem arose in assessing the 'probable' presence of elation that may be associated with antidepressant[35]. The definition we used could be applied with moderate confidence only to the data derived from systematic assessments, and it was clear from scrutiny of the general psychiatry case notes that clinicians apply idiosyncratic criteria or do not document how they have operationalized the term (and they often use the term antidepressant emergent elation inter-changeably with antidepressant-emergent mood instability). As such, we suggest caution in regard to considering 'probable' anti-depressant emergent elation as a BAR criterion until there is greater consensus on how to define and assess it, including agreement about the maximum duration of the time delay between prescription of an antidepressant and the onset of these mood changes, and the level of severity and duration of mood and other symptoms required[35]. Having highlighted these concerns, it is noteworthy that the current edition of DSM (DSM-5) recognizes that (hypo)manic episodes that emerge during antidepressant treatment are indicative of underlying BD, and so we encourage more research on this phenomenon in depressed youth.

Recommending the use of atypical depressive symptoms for screening is less problematic, as increases or decreases in sleep, appetite, activity and energy, are key criteria for the diagnosis of

depressive episodes. Unsurprisingly, the presence or absence of these features was recorded in more than 70% of case notes. Indeed, it is unclear why these features are not employed in screening more often as many, but not all, studies indicate their potential importance in differentiating BD from UP[36-38].

The study has several limitations, most notably that none of the datasets we accessed was derived from studies designed for the purposes of assessing risk of ET-BD in the peak AAO period (15-25 years). The age range was selected because it represents the peak AAO for BD, but it can be argued that these boundaries were somewhat arbitrary. Also, the 24-month time frame for transition from depression to BD can be viewed in the same light. However, we would argue that it represents a pragmatic decision based on research evidence and clinical relevance. First, research on UHR criteria for psychosis and related evidence regarding time to transition suggests that two years is a critical time period[39] and that rates of transition then start to fall. Second, we suggest that it would be feasible and justifiable to monitor depressed youth at risk of ET-BD for this time period to offer the prospect of early interventions as appropriate[40].

Instead of using a prospective cohort study approach, we chose a case-control methodology. The rationale for the sampling strategy was that we wanted to ascertain a large number of ET-BD cases to maximize the statistical power of the clinical utility and NNS analyses. However, a weakness of this approach is that it assumes a degree of homogeneity in the clinical populations recruited into the original datasets we accessed and that the prevalence rates for the BAR risk factors in the case-control sample reflect the true prevalence in other clinical and community settings. Although the base rates for each criterion were within the predicted ranges, they were slightly lower than anticipated for some features (e.g. psychotic symptoms and psychomotor retardation). This reduced the power to detect significant OR and may mean we have prematurely excluded some variables from the NNS analyses. The use of a convenience sample can also be criticised as a potential source of biases, although we emphasize that the data were only used in the prediction of the NNS in routine clinical practice. This calculation, by definition, required access to clinically representative, general psychiatry case records. Lastly, we decided that the recruitment procedure for the case-control study, the sample size and the nature of the available data meant it was inappropriate to explore any additive effects for combinations of risk factors, or to undertake survival analyses of time to transition associated with each risk factor. However, it is important to note that the largest NNS is the rate limiting step for screening (so effects on speed of transition or additive effects) does not change the workload for

screening or case finding. These important issues are being addressed in a larger-scale prospective cohort study.

In conclusion, this study is the first we know of that examines the clinical utility and discriminant validity of each factor included in Bechdolf et al's BAR criteria[7] and of other selected trait, state and familial markers of risk of ET-BD in depressed youth. Cyclothymia in individuals with depression showed the optimum clinical utility, as it is useful for both case finding and screening, and showed the lowest NNS. Unsurprisingly, sub-threshold mania also showed utility. Other clinical features (family history of BD, probable antidepressant emergent elation and atypical depression) had better utility for screening out non-cases than for case finding.

We suggest that future prospective studies of BAR tools should report the clinical utility for screening and case finding of each criterion they include, alongside their NNS. In this way it will be possible to compare these different key aspects across studies and also to determine if some tools are more applicable to selected populations (e.g. specialist mood clinics, early intervention in psychosis services, etc.). For example, the BAR-D tool may be more useful for screening young people with major depression who are in the peak age range for risk of ET-BD than for other populations.

Lastly, we draw attention to the widely held view that youth mental health research would benefit from a trans-diagnostic approach. This is particularly relevant in determining the longer-term trajectories of severe mental disorders, many of which demonstrate at least one episode of depression during the earliest clinical stages[4-6,12,35]. As such, it is of considerable interest that there appears to be convergence in the type of criteria being employed in psychosis and BD to define risk syndromes (combinations of limited sets of state, trait and familial characteristics)[4,5,7]. State characteristics examined in the BAR criteria, such as brief, attenuated of sub-syndromal manic symptoms, clearly parallel the descriptor used in psychosis[4]. There are also similarities in the risk rates for transition to psychosis[4,5] and to BD[7,12]. This would seem to indicate that it may be possible to develop a combined tool that could not only further our understanding not only of who is at risk of ET-BD, but also identify if any characteristics are unique to a 'mood disorder trajectory' and which may be shared with individuals who make a transition to psychosis. Clinically, this may help to plan generic as well as specific interventions and treatments [40]. It would also provide opportunities for research into underlying pathophysiological mechanisms associated with transition[6,35,37].

References

1. Kessler R, Chiu W, Demler O, Walters E. (2005). Prevalence, severity, and comorbidity of 12-month DSM-IV disorders in the National Comorbidity Survey Replication. Archives of General Psychiatry 62, 617–627.

2. Merikangas KR, He JP, Burstein M, Swanson SA, Avenevoli S, Cui L, Benjet C, Georgiades K, Swendsen J (2010). Lifetime prevalence of mental disorders in US adolescents: results from the National Comorbidity Survey Replication– Adolescent Supplement (NCS-A). Journal of the American Academy of Child & Adolescent Psychiatry 49, 980–989.

3. Jones PB (2013). Adult mental health disorders and their age at onset. British Journal of Psychiatry 202(s54), s5–s10.

4. Yung A, Nelson B, Thompson A, Wood S. (2010) The psychosis threshold in Ultra High Risk (prodromal) research: is it valid? Schizophr Res.,120, 1-6.

5. Fusar-Poli P, Bonoldi I, Yung A, Borgwardt S, Kempton M, Valmaggia L, Barale F, Caverzasi E, McGuire P. (2012) Predicting psychosis: meta-analysis of transition outcomes in individuals at high clinical risk. Archives General Psychiatry, 69(3), pp.220-229.

6. McGorry PD, Killackey E, Yung A. Early intervention in psychosis: concepts, evidence and future directions. World Psychiatry. 2008 Oct;7(3):148-56.

7. Bechdolf A, Nelson B, Cotton SM, Chanen A, Thompson A, Kettle J, Conus P, Amminger GP, Yung AR, Berk M, McGorry PD. (2010) A preliminary evaluation of the validity of at-risk criteria for bipolar disorders in help-seeking adolescents and young adults. J Affect Disord.,127, 316-20.

8. Brietzke E, Mansur RB, Soczynska JK, Kapczinski F, Bressan RA, McIntyre RS. Towards a multifactorial approach for prediction of bipolar disorder in at risk populations. J Affect Disord. 2012 Sep;140(1):82-91.

9. Leopold K, Ritter P, Correll CU, Marx C, Özgürdal S, Juckel G, Bauer M, Pfennig A. Risk constellations prior to the development of bipolar disorders: rationale of a new risk assessment tool. J Affect Disord. 2012 Feb; 136(3):1000-10.

10. Youngstrom EA, Murray G, Johnson SL, Findling RL. The 7 up 7 down inventory: a 14-item measure of manic and depressive tendencies carved from the General Behavior Inventory. Psychol Assess. 2013 Dec; 25(4):1377-83.

11. Correll CU, Olvet DM, Auther AM, Hauser M, Kishimoto T, Carrion RE, Snyder S, Cornblatt BA. The Bipolar Prodrome Symptom Interview and Scale-Prospective (BPSS-P): description and validation in a psychiatric sample and healthy controls. Bipolar Disord. 2014 Aug; 16(5):505-22.

12. Bechdolf A, Ratheesh A, Cotton SM, Nelson B, Chanen AM, Betts J, Bingmann T, Yung AR, Berk M, McGorry PD. The predictive validity of bipolar at-risk (prodromal) criteria in help-seeking adolescents and young adults: a prospective study. Bipolar Disord. 2014 Aug; 16(5):493-504.

13. Ratheesh A, Cotton SM, Betts JK, Chanen A, Nelson B, Davey CG, McGorry PD, Berk M, Bechdolf A. Prospective progression from high-prevalence disorders to bipolar disorder: Exploring characteristics of pre-illness stages. J Affect Disord. 2015 Sep 1;183:45-8.

14. Organisation for Economic Co-operation and Development (2007) OECD principles and guidelines for access to research data from public funding. www.oecd.org/dataoecd/9/61/38500813.pdf

15. American Psychiatric Association. (2000). Diagnostic and Statistical Manual for mental Disorders. 4th Edition. Washington: APA.

16. Mitchell A (2011) Sensitivity×PPV is a recognized test called the clinical utility index (CUI+). Eur J Epidemiol. 26,251-2.

17. Mitchell A. (2009) How do we know when a screening test is clinically useful? In: Mitchell AJ, James C, editors. Screening for depression in clinical practice: an evidence-based guide. Oxford: OUP. P 92-107.

18. Bandolier (1996) How Good is that Test? May 1996, p27-32. Oxford Centre for Evidence Based Medicine. http://www.medicine.ox.ac.uk/bandolier/band27/b27-2.html

19. Ostergaard SD, Dinesen PT, Foldager L. (2010) Quantifying the value of markers in screening programmes. European J Epidemiol., 25, 151–154.

20. Ivatt H (2007) Can the antecedents to the diagnosis of bipolar disorder be identified and used to create a clinically applicable screening tool for the early identification of at risk individuals? Master of Medical Science (MMedSci) Thesis, Newcastle University, UK.

21. Scott J, Meyer T (2008) Early intervention in affective disorders: a realistic proposition? Early Intervention in Psychosis. 1, 111-113.

22. Hirschfeld R, Calabrese J, Weissman M, Reed M, Davies M, Frye M, Keck P, Lewis L, McElroy S, McNulty J, Wagner K (2003) Screening for bipolar disorder in the community. J Clin Psychiatry, 64, 53-9.

23. Carvalho AF, Takwoingi Y, Sales PM, Soczynska JK, Köhler CA, Freitas TH, Quevedo J, Hyphantis TN, McIntyre RS, Vieta E. Screening for bipolar spectrum disorders: A comprehensive meta-analysis of accuracy studies. J Affect Disord. 2015 Feb 1;172:337-46.

24. Zimmerman M. Screening for bipolar disorder: confusion between case finding and screening. Psychother Psychosom. 2014;83(5):259-62.

25. Waugh M, Meyer T, Youngstrom E, Scott J. A review of self-rating instruments to identify young people at risk of bipolar spectrum disorders. J Affect Disord. 2014 May;160: 113-21.

26. Axelson D, Goldstein B, Goldstein T, Monk K, Yu H, Hickey MB, Sakolsky D, Diler R, Hafeman D, Merranko J, Iyengar S, Brent D, Kupfer D, Birmaher B. (2015) Diagnostic precursors to bipolar disorder in offspring of parents with bipolar disorder: A longitudinal study. Am J Psychiatry. Jul;172(7):638-46.

27. Marwaha S, He Z, Broome M, Singh SP, Scott J, Eyden J, Wolke D. How is affective instability defined and measured? A systematic review. Psychol Med. 2014 Jul;44(9):1793-808.

28. Zeschel E, Bingmann T, Bechdolf A, Krüger-Oezguerdal S, Correll CU, Leopold K, Pfennig A, Bauer M, Juckel G. Temperament and prodromal symptoms prior to first manic/hypomanic episodes: results from a pilot study. J Affect Disord. 2015 Mar 1; 173:39-44.

29. Duffy A, Horrocks J, Doucette S, Keown-Stoneman C, McCloskey S, Grof P. (2014) The developmental trajectory of bipolar disorder. Br J Psychiatry. Feb; 204(2):122-8.

30. Ratheesh A, Davey C, Hetrick S, Alverez-Jimenez M, Voutier C, Bechdolf A, McGorry P, Scott J, Berk M Cotton S. (2016). A systematic review and meta-analysis of the rates and predictors of prospective transition from major depression to bipolar disorder. Paper submitted.

31. Smoller J, Finn CT. Family, twin, and adoption studies of bipolar disorder. Am J Med Genet C Semin Med Genet. 2003;123C(1):48-58.

32. National Institute of Healthcare and Clinical Excellence (2014). Bipolar disorder (update): the management of bipolar disorder in adults, children and adolescents in primary and secondary care. London: NICE guideline [CG185].

33. National Institute of Healthcare and Clinical Excellence (2015) NICE Quality Standards for Bipolar Disorders. London: NICE Quality Standard [QS 95].

34. Preisig M, Strippoli MP, Castelao E, Merikangas KR, Gholam-Rezaee M, Marquet P, Aubry JM, Vandeleur CL. The specificity of the familial aggregation of early-onset bipolar disorder: A controlled 10-year follow-up study of offspring of parents with mood disorders. J Affect Disord. 2016 Jan 15; 190:26-33.

35. Brichant-Petit Jean C, Scott J, Etain B, et al (2016) A re-examination of anti-depressant treatment emergent mania in bipolar disorder: evidence of gender differences in putative risk and associated factors. Paper submitted.

36. Mitchell P, Goodwin G, Johnson G, Hirschfeld R. (2008) Diagnostic guidelines for bipolar depression: a probabilistic approach. Bipolar Disord, 10,144-52.

37. Hickie IB, Hermens DF, Naismith SL, Guastella AJ, Glozier N, Scott J, Scott EM. (2013) Evaluating differential developmental trajectories to adolescent-onset mood and psychotic disorders. BMC Psychiatry. Nov 12;13:303.

38. Steinan M, Krane-Gartiser K, Morken G, Scott J. (2015) Sleep problems in euthymic bipolar disorders: A review of clinical studies. Curr Psychiatry Rev; 11:1–9.

39. Nelson B, Yuen HP, Wood SJ, Lin A, Spiliotacopoulos D, Bruxner A, et al. Long term follow up of a group at ultra-high risk ("prodromal") for psychosis: the PACE 400 study. JAMA Psychiatry. 2013;70(8):793-802.

40. Vallarino M, Henry C, Etain B, Gehue LJ, Macneil C, Scott EM, et al. An evidence map of psychosocial interventions for the earliest stages of bipolar disorder. Lancet Psychiatry. 2015 Jun;2(6):548-63.

Table 1: Characteristics of cases with early transition to bipolar disorder (ET-BD) and controls with unipolar disorder (UP) who were matched for gender and year of birth

Clinical Characteristics[a]	**ET-BD (N=50)**	**UP (N=50)**
Number of Females	31	31
Median Age (IQ range) in years-		
1st Episode Minor Depression	13.3 (8-16)	13.5 (8-16)
1st Episode Major Depression	17.0 (15-20)	18.0 (16-22)
1st Episode Hypomania or Mania	20.3 (17- 23)	
Median Number of Mood Episodes (IQR)[b]	4 (2-6)	3 (2-5)
Number with a Comorbid Mental Disorder or ASUD	7	8

ET-BD: Early Transition to Bipolar Disorder; UP: Unipolar Disorder; ASUD: Alcohol or Substance Use Disorder.

[a]Median and inter-quartile range (IQR) are reported as the age range of the sample is truncated, or the characteristic was not normally distributed.

[b]Assessment of number of episodes is truncated to age=<25 years (see text for details).

Table 2: Prevalence and performance of each putative risk factors for bipolarity in differentiating between cases with early transition to bipolar (ET-BD) and controls with unipolar disorders (UP).

	ET-BD (N=50)	UP (N=50)	Odds Ratio (95% CI)	Clinical Rule In Accuracy (CUI+)	Clinical Rule Out Accuracy (CUI-)	Overall Clinical Utility[c]
Bipolar at Risk Criteria[a]						
Cyclothymia	39	10	**14.2 (5.4, 37.2)**	good	good	**Case Finding & Screening**
Sub-threshold Mania	26	3	**16.9 (4.7, 61.8)**	moderate	good	**Screening**
Family History of BD	12	2	**7.6 (1.6, 35.9)**	poor	good	**Screening**
Additional Risk Factors[b]						
Probable Antidepressant-emergent Elation	21	8	**3.4 (1.2, 4.9)**	fair	good	**Screening**
Atypical Depression	25	4	**11.5 (3.6, 36.7)**	fair	good	**Screening**
Psychomotor Retardation	6	2	2.6 (0.5, 13.6)	poor	moderate	-
Psychotic Mood Episode	6	1	6.7 (0.8, 57.7)	poor	moderate	-
Family History of other Mood Disorders &/or ASUD	22	14	2.0 (0.9, 4.64)	fair	moderate	-
Multi-generational Family History of Mood Disorders	3	0	3.1 (0.3, 31.1)	poor	moderate	-

[a]Criteria from Bechdolf et al (2010); [b]Factors identified from research literature (see Appendix 1 for details)
ET-BD: Early Transition to Bipolar Disorder; UP: Unipolar Disorder; ASUD: Alcohol or Substance Use Disorder.
CUI= Clinical Utility Index (see text for details and numerical estimate of grading).
[c]Overall clinical utility is only reported if the item received a good grading for either the CUI+ or CUI-.

Table 3: Estimated Number Needed to Screen (NNS) for selected clinical features of early transition from depression to bipolar disorders for individuals assessed by structured systematic clinical interview schedules and the predicted NNS in routine clinical settings.

Clinical Features	Overall Clinical Utility	NNS with Systematic Assessment	Proportion of Case Notes with Missing Data[a]	Predicted NNS in Routine Clinical Practice
Cyclothymia	Case Finding & Screening	1.7	51%	3.5
Sub-threshold Mania	Screening	2.2	68%	6.9
Probable Antidepressant-emergent Elation	Screening	2.8	39%	4.6
Atypical Depression	Screening	4.5	27%	6.2
Family History of Bipolar Disorder	Screening	5.0	16%	5.9

NNS: Number Needed to Screen

[a] The percentage refers to the proportion of clinical case notes that failed to report either the presence or absence of the clinical feature.

Supplementary Material for
Bipolar at-risk criteria: an examination of which clinical features have optimal utility for identifying youth at risk of early transition from depression to bipolar disorders.

Appendix: Further Details on Methodology

Samples

Sample 1- Systematically assessed cases: The sample (n=100) was derived from eight datasets[1-8] that were selected on the basis that the study samples were recruited from general psychiatry settings and some or all of each sample was comprised of cases of bipolar (BD) or unipolar disorders (UP). The databases were chosen as they recorded information from structured, systematic, undertaken by trained clinical researchers that used reliable and valid assessments of current and lifetime DSM IV diagnoses, personality and temperament, comorbid disorders, family history, etc. (for example, the Structured Clinical Interview for DSM IV; the Family Interview for Genetics Studies; Personality Assessment Schedule; the WHO treatment equivalence ratings of medications, etc.). In addition, many of studies that incorporated repeated face to face interviews established inter-rater reliabilities for ratings made by the researchers (e.g. Medical Research Council funded study of cognitive therapy for BD[1]).Further details of the ethical procedures, recruitment processes, and specific clinical assessments and schedules and baseline and follow-up procedures are given elsewhere[2-8] (including a publication that refers to Alison Jackson's PhD at University of Glasgow from which some data was extracted[8]).

The datasets included 691 UP and BD cases where the ratings that were recorded in the dataset allowed the age at onset (AAO) of a first depressive and/or a first (hypo)manic episode to be ascertained. In addition, in BD cases, it was possible to estimate the time between a depressive episode and the first (hypo)manic episode, and the systematic assessments of individuals with UP and BD used assessment tools that recorded information that could be used to rate the presence or absence of each of the BAR criteria when using the assessment tool[9] and of the additional items selected (see next section) that we chose to examine. For example, the use of the Family Interview for Genetics Studies allowed the BAR criteria on genetic risk and other specified details of family history to be identified. It is important to highlight that, in some datasets, the individual participating in the assessment was older than 25 years, but the data reported the AAO and sufficient details on episode polarity, timing and frequency to allow the case to be considered for inclusion in the study sample. This approach is frequently used in long-term cohort studies that use intermittent follow-ups every 2-3 years (e.g. Dunedin study). The approach is referred to as ambi-perspective (i.e. meaning that it is possible to date illness episodes and build a longitudinal picture of events that precede or follow the episode onset).

Using this approach, we identified 50 BD cases (BD I or II) with reliable evidence that their first episode of DSM IV[10] mania or hypomania occurred (i) between 15-25 years and (ii) no more than two years after a major depressive episode. Frequency matching was used to match these BD cases for gender and year of birth to individuals with recurrent UP who met DSM IV[10] criteria for their first major depressive episode (i) between 15-25 years, (ii) did not demonstrate transition to BD during two years of prospective follow-up or by the age of 25 years. We matched cases for birth year rather than age of first episode of any mood or non-affective psychopathology as research suggests that these might be expected to differ between UP and BD e.g. BD may have earlier AAO of depression compared to UP[11,12], and to avoid confounding from any putative birth cohort effects[13].

Sample 2- Clinically assessed cases: Ethical approval was obtained from Newcastle Joint University and Hospital Ethics committee to review the case records of a convenience sample of 40 individuals with UP and 40 with BD attending general psychiatry outpatient clinics. The project was independent of the present study and all the consultant psychiatrists who identified the cases for inclusion in the project and the researchers were blind to the aims of the present study[14].

The case notes were screened to determine the frequency with which the presence or absence of about 40 putative risk factors for BD that were described on a checklist[14]. The recordings did not need to demonstrate the reliability or validity of the assessment, but needed to use a term that allowed the report to be classified into one of the categories. A small number of categories were collapsed into one item as they were often used interchangeably e.g. probable antidepressant emergent elation and antidepressant emergent mood swings were often combined into one category. In other circumstances (e.g. if a case note recording was difficult to interpret), information was allocated to the most appropriate category by consensus between researchers (Ivatt, Scott, or Professor Ferrier or Dr Meyer).

A frequency table was created that recorded the proportion of case notes that had recorded information about each factor (details available from the first author). From this list it was possible to calculate the percentage of case notes with missing data regarding each variable that was of interest in the current study.

Measures: Extended BAR Criteria

Time period for ratings: In this study, we assessed the prevalence of the extended BAR criteria in Sample1. In the ET-BD cases, we noted the occurrence prior to the onset of the first episode of (hypo)mania; in the UP controls, we noted the occurrence prior to the age of 25 years (or the end of the follow-up period if that was sooner).

Bipolar At Risk (BAR) criteria- Bechdolf et al's[9] BAR criteria use a combination of four clinical characteristics to identify three 'at risk' subgroups ([1] sub-threshold mania; [2] depression and cyclothymic features; [3] depression and genetic risk of BD). An individual is regarded as BAR+ if the researcher reliably demonstrates that the presence of any of the state or trait features preceded the onset of the first (hypo)manic episode that met recognized diagnostic criteria (such as DSM IV). If the timing was unclear and no other data were available, the criterion was rated as absent.

The BAR assessment incorporates five exclusion criteria (IQ below normal range; evidence of organic brain disorder; past history of a treated or untreated manic episode; prior treatment with a mood stabilizer; presence of psychotic symptoms); we applied these to our sample selection, with the exception of the presence of psychotic symptoms. We included psychotic symptoms in our study as there is evidence to suggest that these can discriminate between youth at risk of BD compared with recurrent UP[15] and they have been found to predict transition to BD in a meta-analysis of studies of depression[16]. Also, if we wish to screen for ET-BD in early intervention (for psychosis) services, it will be clinically important to explore whether the presence or absence of psychotic symptoms affects the performance of the BAR criteria.

Additional Risk Factors-

Five additional putative risk factors were included in this study as it was hypothesized that these might enhance the ability of the BAR criteria to identify which depressed youth were at risk of early transition to BD (ET-BD). The items were included on the basis that there was robust evidence to suggest they are risk factors for early onset mood disorders and/or may be useful in differentiating BD and UP[16-25]. Furthermore, one factor (probable evidence of anti-depressant emergent elation), was selected as all the transitions to BD showed evidence of this feature in the original study by Bechdolf et al[9]. (However, we would draw attention to the fact that in Sample 2, some clinicians making case note recordings had combined this item and recorded it as 'probable evidence of anti-depressant emergent elation/mood stability', which may affect the reliability or validity of any findings relating to the convenience sample).

The additional risk factors were-

- Psychomotor retardation

- Evidence of psychotic symptoms during one or more mood episodes
- Atypical depression (operationalized by two key features: fatigue and/or hypersomnia)
- Family history of (a) multiple generations (>=2) affected by mood disorders or (b) other mood and alcohol and substance use disorders (ASUD).
- Probable antidepressant emergent elation: to meet this criteria, we used similar definitions to those proposed in other studies of anti-depressant emergent mania (for a discussion see Brichant Petit-John et al[26]), but adapted them for the purposes of this study. Namely, we regarded an individual as meeting this criterion if the mood elation or instability (i) occurred within 90 days of commencing treatment with a recognized antidepressant and (ii) was accompanied by at least one other symptom of mania. However, it is important to note, that this criterion was difficult to assess in some circumstances, and often required access to additional data records (e.g. such as free text). Given these uncertainties (and to be compatible with information recorded for Sample 2), we have reported the item as probable antidepressant emergent elation, but draw attention to the fact that the reliability of the assessment of this criterion was the lowest of any variable studied.

Statistical Analysis

We used several established approaches to examine the statistical significance and clinical utility of the selected factors in differentiating between ET-BD (cases) and UP (controls). However, whilst we calculated the sensitivity (Se) and specificity (Sp), and positive and negative predictive values (PPV; NPV) for each item, we do not report these in the main text as these were used purely to calculate the Clinical Utility Index (CUI)[27,28]. The calculations were as follows: CUI+ =PPVxSe, the CUI- = NPVxSp. (Data for the calculations is shown in the Table below).

	ET-BD (N=50)	UP (N=50)	Se	Sp	PPV	NPV
Cyclothymia	39	10	.78	.80	.80	.78
Sub-threshold Mania	26	3	.52	.94	.89	.66
Family History of BD	12	2	.24	.96	.86	.56
Probable Antidepressant-emergent Elation	21	8	.42	.84	.52	.78
Atypical Depression	25	4	.50	.92	.72	.82
Psychomotor Retardation	6	2	.12	.96	.75	.52
Psychotic Mood Episode	6	1	.12	.98	.86	.53
Family History of other Mood Disorders &/or ASUD	22	14	.44	.72	.61	.56
Multi-generational Family History of Mood Disorders	3	0	.06	.99	.86	.51

The Number Needed to Screen (NNS) is analogous to the number needed to treat (NNT) in clinical treatment trials[29-31]. To estimate the NNS, it is first necessary to calculate the 'fraction correct' (i.e. the proportion of individuals correctly identified as cases or non-cases) and the 'fraction incorrect' (i.e. the error rate or proportion of misidentifications)[30]. The NNS is then calculated by the equation:

NNS= 1/ [Fraction Correct -Fraction Incorrect].

We estimated the NNS for the factors most likely to differentiate between UP and ET-BD when structured systematic assessments are used (using OR findings from Sample 1). The predicted NNS for clinical settings was estimated using the formula: NNS using structured assessments/ proportion of case notes with missing information recording data on the item.

Appendix References

1. Scott J, Paykel E, Morriss R, Bentall R, Kinderman P, Johnston T, Hayhurst H, Abbott R (2006) A randomised controlled trial of CBT versus usual treatment in severe and recurrent bipolar disorders. British Journal of Psychiatry, 188, 313-20

2. Adams J, Scott J (2000) Predicting medication adherence in severe mental disorders. Acta Psychiatr Scand, 101, 119-24.

3. Cornwall P, Garland A, Scott J (2005) Beliefs about depression in patients and partners and their effect on prognosis. Behavioural & Cognitive Psychotherapy. *33, 131-138.*

4. Paykel E, Scott J, Teasdale J, Johnson A, Garland A, Moore R, Jenaway A, Cornwall P, Hayhurst H, Abbott R, Pope M (1999) Prevention of relapse in residual depression by cognitive therapy: A controlled trial. Arch Gen Psychiatry, 56, 829-835.

5. Scott J, Garland A, Moorehead S. (2001) A pilot study of cognitive therapy in bipolar disorders. Psychological Medicine, 31, 459-67.

6. Pope M, Dudley R, Scott J (2007) Determinants of social functioning in severe mood disorders. Bipolar Disord, 9, 38-44.

7. Joseph M, Tacchi M, Scott J (2011) Are you being served? Use of crisis assessment services by youth with severe mood disorders. Conference Proceedings of the General Psychiatry Section Meeting, Royal College of Psychiatrists, Newcastle, September, 2011.

8. Scott J. (2011) Clinical parameters of circadian rhythms in affective disorders. Neuropsycho-pharmacology, Suppl. 4, S671-5.

9. Bechdolf A, Nelson B, Cotton SM, Chanen A, Thompson A, Kettle J, Conus P, Amminger GP, Yung AR, Berk M, McGorry PD. (2010) A preliminary evaluation of the validity of at-risk criteria for bipolar disorders in help-seeking adolescents and young adults. J Affect Disord.,127, 316-20.

10. American Psychiatric Association. (2000). Diagnostic and Statistical Manual for mental Disorders. 4th Edition. Washington: APA.

11. Manzano J, Salvador A. (1993) Antecedents of severe affective (mood) disorders. Patients examined as children or adolescents and as adults. Acta Paedopsychiatr., 56, 11-8.

12. Scott J, Meyer T (2008) Early intervention in affective disorders: a realistic proposition? Early Intervention in Psychosis. 1, 111-113.

13. Golmard JL, Scott J, Etain B, Preisig M, Aubry JM, Henry C, Jamain S, Azorin JM, Leboyer M, Bellivier F. (2016) Using admixture analysis to examine birth-cohort effects on age at onset of bipolar disorder. Acta Psychiatr Scand. Mar;133(3):205-13.

14. Ivatt H (2007) Can the antecedents to the diagnosis of bipolar disorder be identified and used to create a clinically applicable screening tool for the early identification of at risk individuals? Master of Medical Science (MMedSci) Thesis, Newcastle University, UK.

15. Baldessarini RJ, Bolzani L, Cruz N, Jones PB, Lai M, Lepri B, Perez J, Salvatore P, Tohen M, Tondo L, Vieta E. (2010) Onset-age of bipolar disorders at six international sites. J Affect Disord.,121,143-6.

16. Ratheesh A, Davey C, Hetrick S, Alverez-Jimenez M, Voutier C, Bechdolf A, McGorry P, Scott J, Berk M Cotton S. (2016). A systematic review and meta-analysis of the rates and predictors of prospective transition from major depression to bipolar disorder. Paper submitted.

17. Angst J, Gamma A, Endrass J. (2003) Risk factors for the bipolar and depression spectra. Acta Psychiatr Scand, Supplt, 418, 15-19.

18. Bechdolf A, Ratheesh A, Wood SJ, Tecic T, Conus P, Nelson B, Cotton SM, Chanen AM, Amminger GP, Ruhrmann S, Schultze-Lutter F, Klosterkötter J, Fusar Poli P, Yung AR, Berk M, McGorry PD. (2012) Rationale and first results of developing at-risk (prodromal) criteria for bipolar disorder. Curr Pharm Des.18(4):358-75.

19. Bellivier F, Golmard J, Rietschel M, Schulze TG, Malafosse A, Preisig M, McKeon P, Mynett-Johnson L, Henry C, Leboyer M. (2003) Age at onset in bipolar I affective disorder: further evidence for three subgroups. Am J Psychiatry, 160, 999-1001.

20. Correll CU, Olvet DM, Auther AM, Hauser M, Kishimoto T, Carrion RE, Snyder S, Cornblatt BA. The Bipolar Prodrome Symptom Interview and Scale-Prospective (BPSS-P): description and validation in a psychiatric sample and healthy controls. Bipolar Disord. 2014 Aug; 16(5):505-22.

21. Faedda GL, Marangoni C, Serra G, Salvatore P, Sani G, Vázquez GH, Tondo L, Girardi P, Baldessarini RJ, Koukopoulos A. Precursors of bipolar disorders: a systematic literature review of prospective studies. J Clin Psychiatry. 2015 May;76(5):614-24.

22. Kim-Cohen J, Caspi A, Moffitt TE, Harrington H, Milne BJ, Poulton R. (2003) Prior juvenile diagnoses in adults with mental disorder: developmental follow-back of a prospective-longitudinal cohort. Arch Gen Psychiatry. Jul;60(7):709-17.

23. Lapalme, M., Hodgins S LaRoche C. (1997) Children of parents with bipolar disorder: a meta-analysis of risk for mental disorders. Canadian Journal of Psychiatry, 42, 623-31.

24. Mitchell P, Goodwin G, Johnson G, Hirschfeld R. (2008) Diagnostic guidelines for bipolar depression: a probabilistic approach. Bipolar Disord, 10,144-52.

25. Schulze, T, Hedeker D, Zandi P, Rietschel M, McMahon F. (2006) What is familial about familial bipolar disorder? Resemblance among relatives across a broad spectrum of phenotypic characteristics. Arch Gen Psychiatry, **63, 1**368-76.

26. Brichant-Petit Jean C, Scott J, Etain B, et al (2016) A re-examination of anti-depressant treatment emergent mania in bipolar disorder: evidence of gender differences in putative risk and associated factors. Paper submitted.

27. Mitchell A. (2009) How do we know when a screening test is clinically useful? In: Mitchell AJ, James C, editors. Screening for depression in clinical practice: an evidence-based guide. Oxford: OUP. P 92-107.

28. Mitchell A (2011) Sensitivity×PPV is a recognized test called the clinical utility index (CUI+). Eur J Epidemiol. 26,251-2.

29. Sackett D, Haynes R, Guyatt G,Tugwell P (1991) *Clinical epidemiology, a basic science for clinical medicine. 2nd ed. Boston: Little Brown, p83.*

30. Bandolier (1996) How Good is that Test? May 1996, p27-32. Oxford Centre for Evidence Based Medicine. http://www.medicine.ox.ac.uk/bandolier/band27/b27-2.html

31. Ostergaard SD, Dinesen PT, Foldager L. (2010) Quantifying the value of markers in screening programmes. European J Epidemiol., 25, 151–154.