Journal of Affective Disorders 141 (2012) 34–39

Contents lists available at SciVerse ScienceDirect

Journal of Affective Disorders

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

Research report

A diagnostic profile of those who return a false positive assignment on bipolar screening measures

Gordon Parker [a,b,*], Rebecca Graham [a,b], Anne-Marie Rees [b], Shulamit Futeran [b], Paul Friend [b]

[a] School of Psychiatry, University of New South Wales, Australia
[b] Black Dog Institute, Australia

ARTICLE INFO

Article history:
Received 19 December 2011
Received in revised form 18 January 2012
Accepted 20 February 2012
Available online 25 May 2012

Keywords:
Bipolar disorder
Screening measures
Diagnosis

ABSTRACT

Objectives: Our aim was to identify the diagnostic profile of patients classified as 'false positives' on two bipolar screening measures; the Mood Swings Questionnaire (MSQ) and the Mood Disorders Questionnaire (MDQ).
Methods: A total of 1534 patients attending the Black Dog Institute Depression Clinic completed the MSQ-46, and a smaller subset of 852 completed the MDQ. All patients underwent clinical assessment by one or more Institute psychiatrists.
Results: Using clinical assignment (i.e. bipolar vs. unipolar) as the criterion measure for assessing the screening measures, the overall agreement rates were 84% for the MSQ-46 and 74% for the MDQ. Patients identified as 'false positives' were most likely to be clinically diagnosed as having a unipolar non-melancholic depression (37% for MSQ-46; 46% for MDQ), or a primary anxiety condition with secondary non-melancholic depression (19% for MSQ-46; 15% for MDQ). In addition, within the unipolar non-melancholic group, 46% of the MSQ-46 assigned false positives and 63% of the MDQ assigned false positives had co-morbid anxiety conditions.
Conclusions: These findings suggest that patients with anxiety conditions account for a significant proportion of false positive diagnoses on bipolar screening tests — a finding that should be conceded in the development and refinement of such screening measures and in clinical assessment of the possibility of a bipolar disorder.

1. Introduction

The bipolar disorders are gravid disabling illness, both intrinsically and via risks of suicide and misadventure associated with their mood swings, and with a lifetime prevalence of 2–3% (Merikangas et al., 2011) — albeit distinctly higher if 'bipolar not otherwise specified' (NOS) conditions are included. Despite their prevalence and potential to be controlled with appropriate management strategies, the bipolar disorders are often poorly diagnosed with estimates of ten years or more between initial onset and correct diagnosis — and clearly with such analyses limited to those eventually receiving such a diagnosis (Hirschfeld et al., 2003). In response to increasing recognition about difficulties in detecting and diagnosing the bipolar disorders (I, II and NOS), a number of screening measures have been developed. As noted by Zimmerman et al. (2011), the "most widely studied" and the most widely used measure is the Mood Disorder Questionnaire (MDQ), developed by Hirschfeld et al. (2000). In their paper, Zimmerman and colleagues overviewed evaluative studies of the MDQ in samples of depressed patients, and in more "heterogeneous series of psychiatric outpatients", before observing that "None of the studies determined the psychiatric diagnoses associated with false positive results on the MDQ" — and which they set as their study objective.

Specifically, Zimmerman and colleagues studied a sample of patients referred to a Rhode Island psychiatric out-patient facility, and with 480 completing the MDQ. Patients were interviewed by a trained diagnostic rater who administered

* Corresponding author at: School of Psychiatry, University of New South Wales, Australia. Tel.: +61 293824372; fax: +61 293824343.
E-mail address: g.parker@unsw.edu.au (G. Parker).

0165-0327/$ – see front matter © 2012 Elsevier B.V. All rights reserved.
doi:10.1016/j.jad.2012.02.026

the SCID structured clinical interview (First et al., 1995) thus generating current and lifetime DSM diagnoses. While Zimmerman and colleagues scored the MDQ in two ways; one which respected the original rules defined in developing the measure (i.e. the patient affirming 7 or more symptoms clustering within the same time period and resulting in moderate to severe impairment) and the other which ignored the impairment question — we focus on the former since it is the most widely used scoring system. Focusing on their lifetime data, 52 (14.8%) received a bipolar diagnosis (bipolar I = 18; bipolar II = 21; bipolar NOS = 8; cyclothymia = 5). After excluding the affirmed bipolar patients, the researchers focused on the 65 (15.2%) MDQ 'false positives' — that is, where a bipolar diagnosis was not sustained by clinical interview. In that sub-set, most patients received multiple alternate current and lifetime diagnoses. The authors quantified the most common false positive bipolar assignments being received by those with current diagnoses of major depression, social phobia, generalised anxiety and attention deficit disorder; and lifetime diagnoses of major depression, alcohol use and impulse control disorders.

The authors also noted that, in two of the three reported studies (i.e. Hardoy et al., 2005; Konuk et al., 2007; Zimmerman et al., 2009) the majority of respondents who screened positive on the MDQ did not in fact have a bipolar disorder. They judged that their study showed that, within a heterogeneous outpatient psychiatric sample, a false diagnosis of bipolar disorder was high, and most commonly contributed to by conditions that evidenced impulsivity and affective instability (i.e. anxiety, impulse control, substance abuse and attention deficit disorders). The authors concluded that the accuracy of the MDQ appears compromised when used as a community screening test or in psychiatric samples with heterogeneous conditions, and that examination in samples "with differing demographic characteristics is warranted".

At the Black Dog Institute, we similarly developed a screening measure for bipolar disorder — initially termed the Mood Swings Survey (MSS) and later the Mood Swings Questionnaire (MSQ; Parker et al., 2006) which has a 46-item (MSQ-46) and a refined 27-item version (MSQ-27) as used in the Institute's online bipolar self-test. The intrinsic and comparative properties of both versions were tested in relation to the MDQ in two comparative studies (Parker et al., 2008, 2012). In our earlier 2008 evaluative study, we recruited 247 patients referred to our Depression Clinic (with 28% clinically diagnosed as having a bipolar condition) and reported an overall agreement between the assessing clinical psychiatrist's judgement of lifetime polarity of 82% for the MSQ-46 and 81% for the MSQ-27 (kappa of 0.6 for each analysis). Of the 205 completing both the MSQ and MDQ measures, the overall agreement between MSQ-46 and MDQ polarity allocation was high (84% agreement, kappa of 0.63), while a logistic regression analysis using MSQ-46 and MDQ total scale scores identified the MSQ as superior to the MDQ in predicting bipolar classification. Optimal cut-off scores were calculated by QROC analyses (Kraemer, 1992), and quantified as 35 or more for the MSQ-46, 22 or more for the MSQ-27 and 7 or more for the MDQ. In our more recent 2011 study using a significantly larger sample, we reported similar cut-off scores to that found previously (35 for the MSQ-46, 20 for the MSQ-27 and 8 for the MDQ), with MSQ scores again somewhat superior to the MDQ in predicting bipolar classification (Parker et al., 2012). It should be noted that the criterion measure (clinician diagnosis) used a broader definition of hypomania that went beyond that stipulated by DSM-IV and this may have underestimated the screening properties of the MDQ.

In this report, we address the issue raised by Zimmerman et al. (2011) — the extent to which a bipolar screening measure might generate false positive diagnoses. Examining the diagnostic profile of this group of false positives is important as it may indicate suggestions for interpretation and improvement of screening measures. Furthermore, as noted by Zimmerman and colleagues, an examination of the nature of false positives is useful in assisting in the interpretation of studies that use the MDQ as the sole case-finding measure, although we argue against the use of such screening measures as stand-alone diagnostic or case-finding measures. Our focus on a sample of patients primarily with mood disorders is beneficial in effectively highlighting areas for improvement of these measures — measures which serve an important role in signalling the possibility that a patient previously diagnosed as having unipolar depression may, in fact, have a bipolar disorder. Our objectives are to compare the MDQ and MSQ in terms of both their false positive rates and the diagnostic composition of the 'false positive' subjects, particularly to examine for any consistency of confounding across measures as this might suggest general rather than measure-specific factors. While Zimmerman and colleagues examined both current and lifetime diagnoses, we examine only lifetime data.

2. Methods

The sample was derived from the Sydney-based Black Dog Institute's Depression Clinic which acts as a state-wide tertiary service for psychiatrists and primary physicians who refer patients for clarification of diagnosis (especially of mood disorders) and management advice. Patients complete a series of questionnaires via a computerised assessment (including the MSQ) and a number of forms including the MDQ (and where compliance is lower) prior to a detailed clinical assessment by a psychiatrist. Approximately one-third of patients are also assessed by a second independent psychiatrist to derive a consensus diagnosis, including decisions about a lifetime bipolar or non-bipolar condition, and with formal reliability quantified at 89.5% (kappa = 0.77, $p<0.001$) in independently assigning a bipolar/unipolar diagnosis (Parker et al., 2008). Patients are clinically assigned a lifetime diagnosis of a bipolar disorder if they ever had a manic or hypomanic episode meeting DSM-IV criteria (other than imposing a minimum duration for either hypomania or mania) and additional clinically-weighted judgments. If any such diagnosis is suggested but not firmly identified at the interview of the patient, further information may be sought from a relative or the referring physician. Patients reporting 'highs' in seeming response to an antidepressant drug being introduced or prescribed at high dose (so-called bipolar III) are not allocated a bipolar diagnosis and are assigned to the 'unipolar' sample group.