

Validity of the Birchwood Insight Scale in patients with schizophrenia spectrum- and bipolar disorders

Camilla Bakkalia Büchmann^{a*}, Geir Pedersen^{a,b}, Sofie Ragnhild Aminoff^{a,c}, Jannicke Fjæra Laskemoen^a, Elizabeth Ann Barrett^c, Ingrid Melle^a, Trine Vik Lagerberg^a

^a NORMENT and K.G. Jebsen Center for Psychosis Research, Institute of Clinical Medicine, University of Oslo, Norway; Division of Mental Health and Addiction, Oslo University Hospital, Oslo, Norway.

^b Oslo University Hospital, Department of Personality Psychiatry, Division of Mental Health and Addiction, Oslo, Norway.

^c Oslo University Hospital, Early Intervention in Psychosis Advisory Unit for South East Norway, Division of Mental Health and Addiction, Oslo, Norway

*** Corresponding Author:**

NORMENT and KG Jebsen Psychosis Research Centre
Institute of Clinical Medicine, University of Oslo, Norway; Division of Mental Health and Addiction, Oslo University Hospital, Oslo, Norway.

TOP – Psychosis Research Unit, Building 49, Ullevål hospital

P.O. Box 4956

Nydalen, 0424 Oslo, Norway

Email: c.b.buchmann@medisin.uio.no

Highlights

- Findings support a three-factor subscale structure of the Insight Scale (IS)
- Findings support the construct validity of the IS across all three diagnostic groups
- The IS total score showed good convergent validity in SCZD and BD I
- The IS total score showed poor convergent validity in BD II

Abstract

The aim of this study is to investigate the validity of the Norwegian version of the Insight Scale (IS) in large and representative samples of patients with schizophrenia spectrum disorders, bipolar I disorder and bipolar II disorder. A total of 997 participants were included (schizophrenia spectrum disorders: 557; bipolar I disorder: 282; bipolar II disorder: 138). Confirmatory factor analysis was conducted to investigate the construct validity and bivariate correlational analysis was applied to investigate convergent validity. Confirmatory factor analyses indicated a reasonable model fit to the original three-factor subscale structure of the IS in all three diagnostic groups. The IS total score and its subscales correlated significantly with both the insight items in the Young Mania Rating Scale and the Positive and Negative Syndrome Scale in both schizophrenia spectrum disorders and bipolar I disorder. In the bipolar II disorder group, however, the IS subscales correlated poorly with both the observer-rated measures. Our study supports the construct validity of the IS in both schizophrenia spectrum disorder- and bipolar disorder populations. The study also demonstrates that patients' self-reports of insight correspond to observer-based single item ratings of insight in bipolar I disorder and schizophrenia spectrum disorders.

Key words: Construct validity; Convergent validity; Psychometric; Clinical insight; Transdiagnostic; Affective disorder; Severe mental disorders; Self-report; Questionnaire; Measurement; Scale

1. Introduction

Impaired insight into illness is a common clinical feature across various severe mental disorders, especially schizophrenia (SCZ) and bipolar disorder (BD). Studies imply that impaired insight is a key feature in SCZ with prevalence rates varying from 30 – 80 % (Lincoln et al., 2007; Mintz et al., 2003; Wiffen et al., 2010). Less research has been done in individuals with BD, but studies confirm that impaired insight is frequently observed also in this patient population (Cassidy et al., 2001; Varga et al., 2006; Yen et al., 2007a).

Insight is viewed as a multidimensional concept comprised of a number of distinct, but overlapping aspects. Clinical insight, which is the focus of the current study, is often defined as the individual's recognition of having a mental illness, compliance with treatment, and the ability to re-label unusual mental events as pathological (Beck et al., 2004; David, 1990). Furthermore, insight exists along a continuum and may also vary within individuals at different time points over the illness course (Ayesa-Arriola et al., 2014; Buckley et al., 2007). It is also assumed to be modality specific, as patients may have insight into some, but not all, signs of their illness (Ayesa-Arriola et al., 2014; Bayard et al., 2009).

Complex findings highlight the need for clinical attention to the diverging effects insight might have in severe mental disorders, especially when developing and applying specific interventions targeting insight. At one hand insight is positively associated with psychosocial adjustment (Yen et al., 2007b; Yen et al., 2009) and social functioning (van Baars et al., 2013). On the other hand, studies also find that insight is associated with reduced quality of life (Boyer et al., 2012; van Baars et al., 2013), lower self-esteem and increased hopelessness (Cooke et al., 2007; Hasson-Ohayon et al., 2006), as well as increased risk for depression (Mintz et al., 2004) and suicidal behavior (Barrett et al., 2010). There are however conflicting results including opposite findings concerning the relationship between insight and several of these aspects, such as suicidality (Massons et al., 2017), depressive symptoms (Cobo et al., 2016) and quality of life (Montemagni et al., 2014). These contradictory findings may be related to differences in the study populations, but may also depend on the use of different instruments to assess insight.

In addition, a growing number of studies support the relevance of insight to the treatment process itself. The presence of insight facilitates the establishment of a therapeutic alliance

(Novick et al., 2015), and has a significant effect on treatment adherence (Wiffen et al., 2010), while impaired insight has a major negative impact on treatment satisfaction (Bø et al., 2016). Assessment of insight also has implications for clinical decisions, such as whether a patient can be safely discharged from an inpatient setting (Weiler et al., 2000) and is part of risk assessments (Bjørkly et al., 2009). Moreover, recent changes in the role of service users in mental health practice and research have led to increasing emphasis on patient-defined recovery and active involvement of patients in the treatment decision-making process. This makes the evaluation of decision-making capacity increasingly important for clinical practice. Insight has been found to be strongly associated with this capacity in psychotic disorders and manic episodes of BD (Adida et al., 2008; Larkin and Hutton, 2017; Owen et al., 2009; Ruissen et al., 2012). We thus need methods for evaluating insight that are feasible within clinical practice.

Several semi-structured interviews assessing insight have been developed for both research and clinical purposes (SUMD; Amador et al., 1993; SAI; David, 1990; SAIQ; Marks et al., 2000; ITAQ; McEvoy et al., 1989). In a clinical context however, the use of self-report questionnaires has certain potential benefits over the observer-rated format. They are less time consuming and can be a useful low-cost method to secure repeated measurements of insight over the course of illness. Moreover, with self-report methods, patient response bias that could result from social influences from the patient–clinician interaction is eliminated (Jovanovski et al., 2007). In addition, potential observer biases, such as the tendency to rate patients with lower intelligence or poorer communication skills as having less insight, are avoided (Marks et al., 2000).

Several self-report questionnaires have been developed for use in psychotic disorders (IS; Birchwood et al., 1994; VAGUS; Gerretsen et al., 2014; Markova and Berrios, 1992; Markova et al., 2003), BD (Wong et al., 1999), and to measure insight in psychotherapy (IOS; Gori et al., 2015). It is, however, of interest to investigate insight transdiagnostically; i.e. compare levels and dimensions of insight across different severe mental disorders. Based on the symptom-unspecific wording of questions, the Insight Scale (IS; Birchwood et al., 1994) appears to be the currently available self-report questionnaire that best fits with this requirement. The IS was developed on the basis of David's (1990) multidimensional view, and developed to measure change in insight over time. Despite widespread use of the instrument, there has been limited research on its psychometric properties. The construct

validity of the scale was first examined by its developers, who performed a principal components analysis. This yielded a single component accounting for 60% of the total variance, suggesting that the items of the IS reflect a single dimension. However, in a later study, the three dimensional structure was supported by a confirmatory factor analysis based on responses from 218 patients with predominantly SCZD (Trauer and Sacks, 2000). Yet more recently, Cleary et al. (2014) found that a single-factor solution yielded best fit to the data. In the present study, we aim to clarify whether previous findings of factor structure can be replicated in a large and representative sample of participants with SCZD. Although BD shares clinical features with SCZD such as positive psychotic symptoms (Goghari et al., 2013; Yalincetin et al., 2017), negative symptoms (Strauss et al., 2016), depression (Upthegrove et al., 2017), and impaired neurocognition (Simonsen et al., 2011), there are several limitations associated with the use of measurement scales in populations for which they were not designed. With respect to the IS, it may not reflect the symptomology or course of illness among individuals with other severe mental disorders, such as BD. To our knowledge, the dimensional properties of the IS for this population have not been examined. In the current study, we also aim to investigate the factor structure of the IS in participants with BD I and BD II.

Finally, in the current study we aim to examine the convergent validity of the IS through investigating its correspondence with two commonly used clinical measures of insight; the Positive and Negative Syndrome Scale (PANSS; Kay et al., 1987) item G12 ('Lack of judgment and insight') and the Young Mania Rating Scale (YMRS; Young et al., 1978) item 11 ('Insight'). The PANSS was developed for assessing symptoms in SCZD but has also been found valid for use in patients with affective disorders (Purnine et al., 2000) while the YMRS was specifically designed for BD. Several lines of research have demonstrated convergent validity between IS and other observer-based measures of insight (Francis and Penn, 2001; Garavan et al., 1998; Jonsdottir et al., 2008; Karow et al., 2008; van Baars et al., 2013; Young et al., 2003) for patients with SCZD, or mixed mental disorders with psychoses. We are however aware of only one study that has examined convergent validity of the IS in patients with BD. In a previous study from our group, Jonsdottir et al. (2008) compared the observer-rated insight item PANSS G12 and IS in participants with SCZD, BD I and BD II. The results indicated good validity for the SCZD group, acceptable for the BD I group, and poor for the BD II group according to the IS total score. The IS subscale 'Need for treatment' had fairly good convergent validity for all three groups. However, this study included a small

sample, which may have influenced the results. Furthermore, to which extent item 11 of the observer-rated YMRS correspond to the self-report measure of IS, has to our knowledge never been investigated.

The aim of our study is to investigate the validity of the Norwegian version of the IS in large and representative samples of patients with SCZD, BD I and BD II. First we investigate the construct validity by means of confirmatory factor analysis. Since previous studies regarding factor structure have yielded divergent results, we apply hypotheses of both a one- and a three-factor solution. As for the BD I and II groups, no hypothesis is postulated, as this has not been investigated in previous studies. Secondly, we aim to investigate the convergent validity between IS and the insight items of YMRS and PANSS by means of bivariate correlation procedures. In this second set of analysis, we expect self-reported insight to be significantly correlated with the observer-rated insight measures in all three groups, based on previous findings within the field.

2. Methods

2.1 Participants

A total of 997 participants were included in the current study (SCZD $n = 577$; BD I $n = 282$, BD II $n = 138$). The SCZD group included schizophrenia ($n = 427$), schizophreniform disorder ($n = 45$) and schizoaffective disorder ($n = 105$). Demographic and clinical characteristics are summarized in [Table 1](#). Participants were recruited through the on-going multi-center Thematically Organized Psychosis (TOP) Study at the NORMENT KG Jebsen Centre for Psychosis Research at Oslo University Hospital and the University of Oslo from May 2003 to January 2017. Inclusion criteria were a primary diagnosis of BD I, BD II, or SCZD according to the Diagnostic and Statistical Manual of Mental Disorders, fourth edition (DSM-IV-TR, American Psychiatric Association, 2000). Exclusion criteria were a history of severe head injury, neurological or developmental disorders, IQ below 70, age outside the range of 17-65 years, and inability to speak a Scandinavian language or to give informed consent. All participants received a complete description of the study prior to giving written informed consent. The study was approved by the Regional Committee for Medical Research Ethics and the Norwegian Data Inspectorate and completed in accordance with the Helsinki Declaration.

2.2 Clinical assessments

2.2.1 Current symptoms and functioning

Participants were interviewed by clinical psychologists or medical doctors who had completed training in diagnostic and symptom ratings based on the training program at the University of California, Los Angeles (Ventura et al., 1998) and reliability calibration for the TOP study protocol. For DSM-IV diagnostics, reliability was found satisfactory with an overall kappa of 0.77 (95 % CI: 0.60-0.94). Diagnoses were established using the Structural Clinical Interview for DSM Axis I disorders (SCID-I) modules A-E (American Psychiatric Association, 2000). Severity of (hypo)mania was assessed using the observer-rated YMRS (Young et al., 1978), based on the past 48 hours. The YMRS consists of 11 items with a scoring format of 0-4, where 0 represents no abnormality and 4 represent severe abnormality. Severity of depression was assessed using the 30-item Inventory of Depressive Symptoms – Clinician Rated (IDS-C; Rush et al., 1996). Positive and negative symptoms and general psychopathology were assessed with the observer-rated PANSS, using the original three-factor model (Kay et al., 1987). The PANSS is a 16-item scale where each item is rated from 1 ('Absent') to 7 ('Extreme') assessing the past seven days. A score of 4 ('Moderate') or higher indicates the presence of clinical psychopathology. Inter-rater reliability, measured by the intra-class correlation coefficient (ICC 1.1) was 0.82 (95% CI 0.66-0.94) for the PANSS 'Positive' subscale, 0.76 (95% CI 0.58-0.93) for the PANSS 'Negative' subscale and 0.73 (95% CI 0.54-0.90) for the PANSS 'General' subscale. Psychosocial functioning was evaluated with Global Assessment of Functioning (GAF; American Psychiatric Association, 1994). GAF is an observer based single score ranging from 1-100, representing symptom severity and social and occupational impairment. The reliability of GAF scores was not analysed in the current study. However, in a study by Pedersen et al. (2007) reliability of GAF was found acceptable, with a generalizability coefficient of .84, for relative decision and approximately .82 for absolute decisions.

2.2.2 Illness history

Age at onset, duration of illness and history of affective and psychotic episodes were determined based on the detailed history reported during the diagnostic interview. For patients with BD, age at onset was defined as age at the first SCID-verified affective episode, regardless of polarity. For patients in the SCZD group age at onset was defined as first SCID-verified psychotic episode. Duration of illness was defined as time (in years) from age at

onset to age at inclusion in the current study for all groups. Patients were considered to have a history of psychosis if they met criteria for any SCID-verified psychotic symptoms.

2.2.3 Insight into illness

Insight was assessed with the Norwegian version of the IS (Birchwood et al., 1994; Jonsdottir et al., 2008). This scale was modified from the original 3-point response format (agree - unsure - disagree) to a 5-point response format (agree very much - agree - unsure - disagree - disagree very much, ranging from 0-4). During the translation process, minor changes were also made to the wording of one item where 'My stay in hospital *is* necessary' became 'My stay in hospital *was* necessary' in order to take into account patients who were discharged from an inpatient setting. The author of the scale had no objections to the translation and the modifications of the scale (Jonsdottir et al., 2008). Also, at a later time point of the study (approx. 2008), 'It was necessary that I came to the outpatient clinic' was added to the item 'My stay in hospital *was* necessary' to further accommodate for changes in treatment practices over time, with more patients treated without inpatient stays. The eight items of IS are organized into three subscales; 'Awareness of illness' (2 items), 'Re-labeling of symptoms' (2 items), and 'Need for treatment' (4 items). Each subscale has a mean score from 0 to 4. The subscale scores can be summarized to a total score (range 0-12), where a higher score indicates more insight. Scores ≥ 9 imply good insight (Birchwood et al., 1994). Insight was also measured with item G12 ('Lack of judgement & insight') from the PANSS and item 11 ('Insight') from the YMRS. These are global measures of insight assessing the individual's awareness of symptoms, disease, and the need for treatment. Contrary to the IS, for both PANSS and YMRS higher scores indicate impaired insight.

2.2.4 Statistical analyses

Confirmatory factor analysis was conducted with Mplus 7.11 (Muthén & Muthén, 2012) with estimations based on the Maximum Likelihood Mean adjusted (MLM) function (Curran et al., 1996). Goodness of fit was estimated by Root Mean Square Error of Approximation (RMSEA; Steiger, 1990), the Tucker Lewis index (TLI; Tucker and Lewis, 1973), the Comparative Fit Index (CFI; Bentler, 1990), and the Standardized Root Mean Square Residual (SRMR; Hu and Bentler, 1999).

RMSEA values of 0.05 or below indicate good model fit, values between 0.05 and 0.08 indicate a reasonable fit, values between 0.08 and 0.10 a mediocre fit, and values greater than 0.10 an unacceptable fit (MacCallum et al., 1996). The general consensus is to use a cut-off value close to 0.06 (Hu and Bentler, 1999) or a stringent upper limit of 0.07 (Steiger, 2007). The CFI and TLI are derived from the chi square statistic, which measures the fit of the model compared to the independence model; the values are supposed to be between 0 and 1. Values greater than 0.90 are normally required for a good fit of a model although Hu and Bentler (1999) have suggested $TLI \geq 0.95$ as the threshold. The SRMR is the mean absolute value of the covariance residuals, and it ranges from 0 to 1.0. Well-fitting models should have values lower than 0.05 (Byrne, 1998; Diamantopoulos and Siguaw, 2000), but values up to 0.08 are acceptable (Hu and Bentler, 1999).

Descriptive analyses and bivariate correlational analyses were conducted using IBM Statistical Package for the Social Sciences version 22 (SPSS, Chicago, IL). Since both PANSS G12, YMRS 11 and the IS subscales were significantly skewed, Pearson's Correlation Coefficients were estimated using a simple bootstrapping procedure with iteration set to 2000. For correlation analysis, we excluded cases pairwise. Analyses were two-tailed with the significance level set at 0.01.

3. Results

3.1 Factor structure

Two different models of the IS items were analysed for each clinical group. In the first model, all eight items were modelled into one single factor. In the second model, items were specified according to the original three-factor subscale structure; 'Re-labeling of symptoms' (items 1 and 8), 'Awareness of illness' (items 2 and 7), and 'Need for treatment' (items 3, 4, 5, and 6). From the results shown in [Table 2](#), the one-factor model revealed poor fit to the data, while the three-factor model revealed acceptable fit among all three diagnostic groups. However, from the Mplus' modification indices (suggesting modifications to improve model fit), both models (for all groups) had their fit indices reduced due to high residual covariance between item 3 ('I do not need medication') and 5 ('The doctor is right in prescribing medication for me'). Furthermore, item 1 ('Some of my symptoms are made by my mind') and item 4 ('My stay in hospital was necessary') revealed low R^2 compared to the other items, suggesting that these items might appear somewhat ambiguous or state specific.

The correlations between IS subscale scores revealed low to mediocre estimates, supporting that they reflect different aspects of insight (Table 3).

3.2 Convergent validity

The results of the bivariate correlational analyses are presented in Table 4. In the SCZD group, all IS subscale scores showed significant correlations with both the PANSS G12 and YMRS 11 with medium to large effect sizes ($r = -0.30$ to -0.55 , $p < 0.01$).

In the BD I group, the analyses showed significant correlations with medium effect size between the IS ‘Awareness of illness’ and ‘Need for treatment’-subscales and the PANSS G12 and YMRS 11 ($r = -0.35$ to -0.49 , $p < 0.01$). For the IS ‘Re-labeling of symptoms’-subscale, the correlation with PANSS G12 and YMRS 11 was significant but with a small effect size ($r = -0.29$ and -0.28 , $p < 0.01$).

For the BD II group the results were more divergent. There was a significant but weak correlation between the IS ‘Awareness of illness’-subscale and PANSS G12 and YMRS 11 ($r = -0.21$, $p < 0.05$ and -0.24 , $p < 0.01$, respectively). There was also a significant correlation with medium effect size between the IS ‘Need for treatment’-subscale and PANSS G12 ($r = -0.38$, $p < 0.01$). No significant correlation was found between this IS item and the YMRS 11. There were no significant correlations between the IS ‘Re-labeling of symptoms’-subscale and either the PANSS G12 or the YMRS 11.

A significant correlation was also observed between the PANSS G12 and the YMRS 11 within all three diagnostic groups, with large effect sizes ($r = 0.59$ to 0.66 , $p < 0.001$).

4. Discussion

The main finding of the current study is that the IS represents a multidimensional construct, including three aspects of clinical insight in concordance with the dimensions postulated by David (1990). The weak associations between the three subscales of IS give support for the distinction between these three aspects of insight. However, as two of the three subscales are operationalized by only two indicators, they might not address the depth of their respective construct. As for the third subscale, operationalized by four indicators, our findings of high

residual covariance between item 3 ('I do not need medication') and 5 ('The doctor is right in prescribing medication for me') suggest that this subscale has a clear potential for improvement.

The fit indices of the factor analysis also indicate that, although originally designed for patients with SCZD, the application of the IS also can be extended to patients with BD. However, this has yet to be confirmed as it raises the question about measurement invariance. This refers to the assumption that the construct validity of a measure is the same across different groups or between time spans within a group, and that group comparisons of sample estimates reflect true group differences, not contaminated by group specific biases (Byrne and van De Vijver, 2010; van de Schoot et al., 2013; Vandenberg and Lance, 2000).

As the three subscales of IS were well reproduced across all three diagnostic groups, our findings are consistent with the three-factor solution reported by the developers of the scale, further replicated by Trauer and Sacks (2000) in a mixed sample of patients with severe mental disorders. In contrast, Cleary and colleagues (2014) concluded that the items of the IS reflect a single dimension, and that in order to accurately measure a more multidimensional insight construct, more items should be included. This discrepancy could in part be due to a difference in methods between the studies. That is, Cleary et al. (2014) point out that their analysis relied on data collected in an observer-administered format (to avoid problems associated with low literacy) rather than a truly self-administered format. They argue that although it is conceivable that this could influence responses, it is unlikely to have affected the factor structure. The findings of the current study however indicate that this may still have been the case.

We hypothesized that self-reported insight would correlate significantly with the observer-rated insight measures of YMRS and PANSS in all diagnostic groups. This hypothesis was not entirely confirmed. We found that the IS total scale and the items that assess insight in the YMRS and PANSS were highly correlated in the SCZD group. These findings are consistent with previous research showing that observer-rated and self-report measures of insight have a moderate to high level of agreement (Cuesta et al., 2000; Francis and Penn, 2001; Jovanovski et al., 2007; Sanz et al., 1998; Young et al., 2003). In the BD groups the convergence between IS total scale and the observer-rated measures was more moderate, but still significant. However, the 'Re-labeling of symptoms'-subscale was only weakly or not

correlated at all with the YMRS 11 or the PANSS G12 in the BD I and II groups. One possible explanation for this is that the items comprising this subscale assume psychotic experiences. To clarify this issue, we conducted a follow-up analysis (results not shown, available on request) controlling for psychotic episodes. However, this revealed no substantial difference with respect to the magnitude of correlation coefficients or the correlational pattern between them for any of the three diagnostic groups. This suggests that the weak convergent validity especially for the bipolar II group was not related to previous psychotic episodes. Cleary et al. (2014) also found item 1 ('Some of my symptoms are made by my mind') to have poor performance in their factor analysis and proposed that, unlike the other items which are relatively concrete, the use of 'made by your/my mind' is variably interpreted by respondents. The fact that the correlations between the 'Re-labeling of symptoms'-subscale and the observer-rated insight measures were quite weak also in the SCZD group, supports this suggestion. Taken together with the findings in our follow-up analysis, this indicates that there are discrepancies in the way the different instruments assess the 'Re-labeling of symptoms' dimension of insight and that these discrepancies may be independent from the presence/absence of psychosis.

Although not a specific aim of the current study, we also investigated convergent validity between the PANSS G12 and the YMRS 11. The two insight items correlated significantly for all three diagnostic groups. The moderate strengths of the correlations may be due to differences in the way the two items are operationalized, the level of measurement (seven- versus five-point response scale), and lastly, to differences in timeframe, e.g. the scoring of YMRS is based on the last 48 hours, whereas the PANSS has a time frame of 7 days.

It should also be noted that both the IS, the PANSS and the YMRS all rely solely on the biomedical model of illness. Consequently, individuals who present disease attributions and acceptance of psychiatric treatments score higher on measures of insight (Saravanan et al., 2007). These instruments do not capture alternative explanatory models, and it has been argued that within this framework insight does not promote an understanding of the complexities involved in how an individual makes sense of their illness (McCabe and Quayle, 2002). Research has shown that different and even apparently contradictory explanatory models can exist side by side in the same individual (Marriott et al., 2018; Saravanan et al., 2007). Following this line of reasoning, recent views have regarded the concept of clinical insight as being too narrow (Marriott et al., 2018). Roe and colleagues have developed the

concept of narrative insight to expand the insight construct and suggest that insight can be studied as a story people tell about their illness (Lysaker et al., 2009; Roe et al., 2008; Roe and Kravetz, 2003). More recently, insight have also been understood as a form of metacognition where impaired insight may in part reflect a failure in strategic metacognition (David et al., 2012; Pousa et al., 2017). In the years to come we may very well see a change in the conceptualization of the insight construct.

Some potential limitations of the current study should be mentioned. We used single item measures as observer-rated insight measures. At first glance, these measures might seem superficial and are often referred to as unidimensional. They do however incorporate overlapping dimensions of clinical insight such as awareness of illness, need for treatment and awareness of behavioral changes. The assessments are made along a continuum and especially the PANSS offers an elaborate scoring manual. As such they represent a useful addition to observer-rated insight scales, as both the PANSS and the YMRS are frequently used in research settings and clinical practice. Moreover, the PANSS insight item has previously showed high correlations with multidimensional observer-rated questionnaires assessing insight such as the ITAQ, the SAI and the SUMD (Jørgensen et al., 2015; Karow et al., 2008; Sanz et al., 1998) in research with patients with SCZD.

Moreover, the reliability of an individual's self-assessment is subject to several potential limitations, and self-report of insight could be more challenging than self-report of other psychological phenomena. However, most self-report scales measuring insight, including the scale used in the current study, address insight in an indirect way by asking the respondents to make judgements about their attitude to their symptoms and illness. Based on the common use of self-report questionnaires assessing insight, there seems to be a general agreement in the field that such judgements do not differ from self-reports of other psychological phenomena.

Also, the present study did not counterbalance the administration of the observer-rated measures PANSS G12 and YMRS 11 versus the self-reported IS. Previous studies comparing observer-rated and self-report measures of insight found the reliability of insight measures to be higher when a self-report measure is completed before the observer-rated scale (Jovanovski et al., 2007; Young et al., 2003). The authors of these studies argue that the data obtained are more reliable when individuals make the initial decision regarding their beliefs,

without the effect of social influence. In the current study in cases where the examiner rated the level of insight before the participant completed the self-report measure, we cannot exclude the possibility that this influenced the self-report by increasing the probability to report beliefs already explored during the clinical interview.

Finally, we cannot rule out that the referring clinicians selected candidates with higher levels of insight, or that potential participants who declined to participate may have more impaired clinical insight than subjects that participated, leaving us with a potentially selected group. However, the participants in the study presented insight scores below the cut-off value of 9, indicating impaired insight. This finding is similar to other comparable studies (Braw et al., 2012). Moreover, the study recruited participants through both inpatients and outpatient treatment services in the Norwegian mental health care system, offering public mental health care to all individuals with mental illness within a given catchment area. This ensures a sample with high degree of representability. This is also a major strength of the study, together with the extensive clinical characterization of participants and the large sample size.

To our knowledge, the present study is the first to investigate the structure of IS in patients with BD I and II since its initial development. From the current findings, the validity and generalizability of the dimensions of IS to these disorders have been given partial support. For the SCZD and the BD I groups, the subscales of IS have revealed acceptable convergent validity as compared to other measures of insight. This was however not true for the BD II group. Furthermore, the ‘Re-labeling of symptoms’-subscale showed poor convergent validity within all the three diagnostic groups. Although IS appears to serve its purpose fairly well in its current form, our suggestions for further development of IS would be that the subscales should be operationalized with more items, especially the subscales ‘Awareness of illness’ and ‘Re-labeling of symptoms’, and that the subscale ‘Need for treatment’ would benefit from some refinement. Lastly, we would suggest that further studies of IS – or its revised form should include analyses of measurement invariance.

Funding

This study was funded by grants from the University of Oslo, South-Eastern Norway Health Authority (#2017-112), the Research Council of Norway (#223273) and the KG Jebsen Foundation (#SKGJ-MED-008). The funding sources had no further role in study design, in

the collection, analysis, and interpretation of data, in the writing of the report, or the decision to submit the paper for publication.

Declarations of interest

None

Authors' contributions

IM and TVL have contributed to the conception of the study. CBB, GP, IM and TVL contributed to the study design. CBB, SRA, JFL, EAB, and TVL contributed to the data collection. CBB and GP conducted the statistical analyses, interpreted the data and drafted the manuscript. All authors participated in discussion of findings, critical revision of manuscript drafts and approved the final version.

Acknowledgements

We would like to thank all participants for taking part in the study and our collaborators in the NORMENT K.G. Jebsen Center for psychosis research for their data collection and administrative contribution. We thank Thomas Doug Bjella, Ruth Kristine Vik, and Therese Markl for their invaluable help in preparing the data set and Tiril Østefjells for her kind assistance with methodical matters.

References

- Adida, M., Clark, L., Pomietto, P., Kaladjian, A., Besnier, N., Azorin, J.M., et al., 2008. Lack of insight may predict impaired decision making in manic patients. *Bipolar disord.* 10 (7), 829-837.
- Amador, X.F., Strauss, D.H., Yale, S.A., Flaum, M.M., Endicott, J., Gorman, J., 1993. Assessment of insight in psychosis. *Am. J. Psychiatry* 150 (6), 873-879.
- American Psychiatric Association, 1994. *Diagnostic and Statistical Manual of Mental Disorders : DSM-IV*, 4th ed. ed. American Psychiatric Association, Washington, D.C.
- American Psychiatric Association, 2000. *Diagnostic and Statistical Manual of Mental Disorders: DSM-IV-TR*, 4th ed., text revision. ed. American Psychiatric Association, Washington, DC.
- Ayesa-Arriola, R., Morinigo, J.D., David, A.S., Perez-Iglesias, R., Rodriguez-Sanchez, J.M., Crespo-Facorro, B., 2014. Lack of insight 3 years after first-episode psychosis: an unchangeable illness trait determined from first presentation? *Schizophr. Res.* 157 (1-3), 271-277.
- Barrett, E.A., Sundet, K., Færden, A., Agartz, I., Bratlien, U., Romm, K.L., et al., 2010. Suicidality in first episode psychosis is associated with insight and negative beliefs about psychosis. *Schizophr. Res.* 123 (2-3), 257-262.
- Bayard, S., Capdevielle, D., Boulenger, J.P., Raffard, S., 2009. Dissociating self-reported cognitive complaint from clinical insight in schizophrenia. *Eur. Psychiatry* 24 (4), 251-258.
- Beck, A.T., Baruch, E., Balter, J.M., Steer, R.A., Warman, D.M., 2004. A new instrument for measuring insight: the Beck Cognitive Insight Scale. *Schizophr. Res.* 68 (2-3), 319-329.
- Bentler, P.M., 1990. Comparative fit indexes in structural models. *Psychol. Bull.* 107 (2), 238-246.
- Birchwood, M., Smith, J., Drury, V., Healy, J., Macmillan, F., Slade, M., 1994. A self-report Insight Scale for psychosis: reliability, validity and sensitivity to change. *Acta Psychiatr. Scand.* 89 (1), 62-67.
- Bjørkly, S., Hartvig, P., Heggen, F.A., Brauer, H., Moger, T.A., 2009. Development of a brief screen for violence risk (V-RISK-10) in acute and general psychiatry: An introduction with emphasis on findings from a naturalistic test of interrater reliability. *Eur. Psychiatry* 24 (6), 388-394.

- Boyer, L., Aghababian, V., Richieri, R., Loundou, A., Padovani, R., Simeoni, M., et al., 2012. Insight into illness, neurocognition and quality of life in schizophrenia. *Prog. Neuropsychopharmacol. Biol. Psychiatry* 36 (2), 271-276.
- Braw, Y., Sitman, R., Sela, T., Erez, G., Bloch, Y., Levkovitz, Y., 2012. Comparison of insight among schizophrenia and bipolar disorder patients in remission of affective and positive symptoms: analysis and critique. *Eur. Psychiatry* 27 (8), 612-618.
- Buckley, P.F., Wirshing, D.A., Bhushan, P., Pierre, J.M., Resnick, S.A., Wirshing, W.C., 2007. Lack of insight in schizophrenia: impact on treatment adherence. *CNS drugs* 21 (2), 129-141.
- Byrne, B.M., 1998. Structural equation modeling with LISREL, PRELIS, and SIMPLIS : basic concepts, applications, and programming. L. Erlbaum Associates, Mahwah, N.J.
- Byrne, B.M., van De Vijver, F.J., 2010. Testing for measurement and structural equivalence in large-scale cross-cultural studies: Addressing the issue of nonequivalence. *International Journal of Testing* 10 (2), 107-132.
- Bø, B., Ottesen, O.H., Gjestad, R., Jørgensen, H.A., Kroken, R.A., Løberg, E.M., et al., 2016. Patient satisfaction after acute admission for psychosis. *Nordic journal of psychiatry* 70 (5), 321-328.
- Cassidy, F., McEvoy, J.P., Yang, Y.K., Wilson, W.H., 2001. Insight is greater in mixed than in pure manic episodes of bipolar I disorder. *J. Nerv. Ment. Dis.* 189 (6), 398-399.
- Cleary, S.D., Bhatti, S., Broussard, B., Cristofaro, S.L., Wan, C.R., Compton, M.T., 2014. Measuring insight through patient self-report: an in-depth analysis of the factor structure of the Birchwood Insight Scale. *Psychiatry Res.* 216 (2), 263-268.
- Cobo, J., Nieto, L., Ochoa, S., Pousa, E., Usall, J., Banos, I., et al., 2016. Insight and gender in schizophrenia and other psychoses. *Psychiatry Res.* 243, 268-277.
- Cooke, M.A., Peters, E.R., Greenwood, K.E., Fisher, P.L., Kumari, V., Kuipers, E., 2007. Insight in psychosis: influence of cognitive ability and self-esteem. *Br. J. Psychiatry* 191, 234-237.
- Cuesta, M.J., Peralta, V., Zarzuela, A., 2000. Reappraising insight in psychosis. Multi-scale longitudinal study. *Br. J. Psychiatry* 177, 233-240.
- Curran, P.J., West, S.G., Finch, J.F., 1996. The robustness of test statistics to nonnormality and specification error in confirmatory factor analysis. *Psychol. Methods* 1 (1), 16-29.
- David, A.S., 1990. Insight and psychosis. *Br. J. Psychiatry* 156, 798-808.

- David, A.S., Bedford, N., Wiffen, B., Gilleen, J., 2012. Failures of metacognition and lack of insight in neuropsychiatric disorders. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 367 (1594), 1379-1390.
- Diamantopoulos, A., Siguaw, J.A., 2000. *Introducing LISREL*. Sage Publications, Ltd, London.
- Francis, J.L., Penn, D.L., 2001. The relationship between insight and social skill in persons with severe mental illness. *J. Nerv. Ment. Dis.* 189 (12), 822-829.
- Garavan, J., Browne, S., Gervin, M., Lane, A., Larkin, C., O'Callaghan, E., 1998. Compliance with neuroleptic medication in outpatients with schizophrenia; relationship to subjective response to neuroleptics; attitudes to medication and insight. *Compr. Psychiatry* 39 (4), 215-219.
- Gerretsen, P., Remington, G., Borlido, C., Quilty, L., Hassan, S., Polsinelli, G., et al., 2014. The VAGUS insight into psychosis scale--self-report and clinician-rated versions. *Psychiatry Res.* 220 (3), 1084-1089.
- Goghari, V.M., Harrow, M., Grossman, L.S., Rosen, C., 2013. A 20-year multi-follow-up of hallucinations in schizophrenia, other psychotic, and mood disorders. *Psychol. Med.* 43 (6), 1151-1160.
- Gori, A., Craparo, G., Giannini, M., Loscalzo, Y., Caretti, V., La Barbera, D., et al., 2015. Development of a new measure for assessing insight: Psychometric properties of the insight orientation scale (IOS). *Schizophr. Res.* 169 (1-3), 298-302.
- Hasson-Ohayon, I., Kravetz, S., Roe, D., David, A.S., Weiser, M., 2006. Insight into psychosis and quality of life. *Compr. Psychiatry* 47 (4), 265-269.
- Hu, L.-t., Bentler, P.M., 1999. Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Struct. Equ. Modeling* 6 (1), 1-55.
- Jonsdottir, H., Engh, J.A., Friis, S., Birkenæs, A., Ringen, P.A., Vaskinn, A., et al., 2008. Measurement of insight in patients with bipolar disorder: are self-rated scales developed for patients with schizophrenia applicable? *J. Nerv. Ment. Dis.* 196 (4), 333-335.
- Jovanovski, D., Zakzanis, K.K., Atia, M., Campbell, Z., Young, D.A., 2007. A comparison between a researcher-rated and a self-report method of insight assessment in chronic schizophrenia revisited: a replication study using the SUMD and SAIQ. *J. Nerv. Ment. Dis.* 195 (2), 165-169.
- Jørgensen, R., Licht, R.W., Lysaker, P.H., Munk-Jørgensen, P., Buck, K.D., Jensen, S.O., et al., 2015. Effects on cognitive and clinical insight with the use of Guided Self-

- Determination in outpatients with schizophrenia: A randomized open trial. *Eur. Psychiatry* 30 (5), 655-663.
- Karow, A., Pajonk, F., Reimer, J., Hirdes, F., Osterwald, C., Naber, D., et al., 2008. The dilemma of insight into illness in schizophrenia: Self- and expert-rated insight and quality of life. *Eur. Arch. Psychiatry Clin. Neurosci.* 258 (3), 152-159.
- Kay, S.R., Fiszbein, A., Opler, L.A., 1987. The positive and negative syndrome scale (PANSS) for schizophrenia. *Schizophr. Bull.* 13 (2), 261-276.
- Larkin, A., Hutton, P., 2017. Systematic review and meta-analysis of factors that help or hinder treatment decision-making capacity in psychosis. *Br. J. Psychiatry* 211 (4), 205-215.
- Lincoln, T.M., Lullmann, E., Rief, W., 2007. Correlates and long-term consequences of poor insight in patients with schizophrenia. A systematic review. *Schizophr. Bull.* 33 (6), 1324-1342.
- Lysaker, P., Yanos, P.T., Roed, D., 2009. The role of insight in the process of recovery from schizophrenia: A review of three views. *Psychosis: Psychological, Social and Integrative Approaches* 1 (2), 113-121.
- MacCallum, R.C., Browne, M.W., Sugawara, H.M., 1996. Power analysis and determination of sample size for covariance structure modeling. *Psychol. Methods* 1 (2), 130-149.
- Markova, I.S., Berrios, G.E., 1992. The assessment of insight in clinical psychiatry: a new scale. *Acta Psychiatr. Scand.* 86 (2), 159-164.
- Markova, I.S., Roberts, K.H., Gallagher, C., Boos, H., McKenna, P.J., Berrios, G.E., 2003. Assessment of insight in psychosis: a re-standardization of a new scale. *Psychiatry Res.* 119 (1-2), 81-88.
- Marks, K.A., Fastenau, P.S., Lysaker, P.H., Bond, G.R., 2000. Self-Appraisal of Illness Questionnaire (SAIQ): relationship to researcher-rated insight and neuropsychological function in schizophrenia. *Schizophr. Res.* 45 (3), 203-211.
- Marriott, M.R., Thompson, A.R., Cockshutt, G., Rowse, G., 2018. Narrative insight in psychosis: The relationship with spiritual and religious explanatory frameworks. *Psychol. Psychother.* Doi:10.1111/papt.12178. [Epub ahead of print]
- Massons, C., Lopez-Morinigo, J.D., Pousa, E., Ruiz, A., Ochoa, S., Usall, J., et al., 2017. Insight and suicidality in psychosis: A cross-sectional study. *Psychiatry Res.* 252, 147-153.
- McCabe, R., Quayle, E., 2002. Knowing your own mind. *The Psychologist* 15 (1), 14-16.

- McEvoy, J.P., Apperson, L., Appelbaum, P.S., Ortlip, P., Brecosky, J., Hammill, K., et al., 1989. Insight in schizophrenia: Its relationship to acute psychopathology. *J. Nerv. Ment. Dis.* 177 (1), 43-47.
- Mintz, A.R., Addington, J., Addington, D., 2004. Insight in early psychosis: a 1-year follow-up. *Schizophr. Res.* 67 (2-3), 213-217.
- Mintz, A.R., Dobson, K.S., Romney, D.M., 2003. Insight in schizophrenia: a meta-analysis. *Schizophr. Res.* 61 (1), 75-88.
- Montemagni, C., Castagna, F., Crivelli, B., De Marzi, G., Frieri, T., Macri, A., et al., 2014. Relative contributions of negative symptoms, insight, and coping strategies to quality of life in stable schizophrenia. *Psychiatry Res.* 220, 102-111.
- Novick, D., Montgomery, W., Treuer, T., Aguado, J., Kraemer, S., Haro, J.M., 2015. Relationship of insight with medication adherence and the impact on outcomes in patients with schizophrenia and bipolar disorder: results from a 1-year European outpatient observational study. *BMC Psychiatry* 15, 189.
- Owen, G.S., David, A.S., Richardson, G., Szukler, G., Hayward, P., Hotopf, M., 2009. Mental capacity, diagnosis and insight in psychiatric in-patients: a cross-sectional study. *Psychol. Med.* 39 (8), 1389-1398.
- Pedersen, G., Hagtvet, K.A., Karterud, S., 2007. Generalizability studies of the Global Assessment of Functioning-Split version. *Compr. Psychiatry* 48 (1), 88-94.
- Pousa, E., Ochoa, S., Cobo, J., Nieto, L., Usall, J., Gonzalez, B., et al., 2017. A deeper view of insight in schizophrenia: Insight dimensions, unawareness and misattribution of particular symptoms and its relation with psychopathological factors. *Schizophr. Res.* 189, 61-68.
- Purnine, D.M., Carey, K.B., Maisto, S.A., Carey, M.P., 2000. Assessing positive and negative symptoms in outpatients with schizophrenia and mood disorders. *J. Nerv. Ment. Dis.* 188 (10), 653-661.
- Roe, D., Hasson-Ohayon, I., Kravetz, S., Yanos, P.T., Lysaker, P.H., 2008. Call it a monster for lack of anything else: narrative insight in psychosis. *J. Nerv. Ment. Dis.* 196 (12), 859-865.
- Roe, D., Kravetz, S., 2003. Different ways of being aware of a psychiatric disability: a multifunctional narrative approach to insight into mental disorder. *J. Nerv. Ment. Dis.* 191 (7), 417-424.

- Ruissen, A.M., Widdershoven, G.A., Meynen, G., Abma, T.A., van Balkom, A.J., 2012. A systematic review of the literature about competence and poor insight. *Acta Psychiatr. Scand.* 125 (2), 103-113.
- Rush, A.J., Gullion, C.M., Basco, M.R., Jarrett, R.B., Trivedi, M.H., 1996. The Inventory of Depressive Symptomatology (IDS): psychometric properties. *Psychol. Med.* 26 (3), 477-486.
- Sanz, M., Constable, G., Lopez-Ibor, I., Kemp, R., David, A.S., 1998. A comparative study of insight scales and their relationship to psychopathological and clinical variables. *Psychol. Med.* 28 (2), 437-446.
- Saravanan, B., Jacob, K.S., Johnson, S., Prince, M., Bhugra, D., David, A.S., 2007. Belief models in first episode schizophrenia in South India. *Soc. Psychiatry Psychiatr. Epidemiol.* 42 (6), 446-451.
- Simonsen, C., Sundet, K., Vaskinn, A., Birkenaes, A.B., Engh, J.A., Faerden, A., et al., 2011. Neurocognitive dysfunction in bipolar and schizophrenia spectrum disorders depends on history of psychosis rather than diagnostic group. *Schizophr. Bull.* 37 (1), 73-83.
- Steiger, J.H., 1990. Structural model evaluation and modification: An interval estimation approach. *Multivariate Behavioral Research* 25 (2), 173-180.
- Steiger, J.H., 2007. Understanding the limitations of global fit assessment in structural equation modeling. *Pers. Individ. Dif.* 42 (5), 893-898.
- Strauss, G.P., Vertinski, M., Vogel, S.J., Ringdahl, E.N., Allen, D.N., 2016. Negative symptoms in bipolar disorder and schizophrenia: A psychometric evaluation of the brief negative symptom scale across diagnostic categories. *Schizophr. Res.* 170 (2-3), 285-289.
- Trauer, T., Sacks, T., 2000. The relationship between insight and medication adherence in severely mentally ill clients treated in the community. *Acta Psychiatr. Scand.* 102 (3), 211-216.
- Tucker, L.R., Lewis, C., 1973. A reliability coefficient for maximum likelihood factor analysis. *Psychometrika* 38 (1), 1-10.
- Upthegrove, R., Marwaha, S., Birchwood, M., 2017. Depression and Schizophrenia: Cause, Consequence, or Trans-diagnostic Issue? *Schizophr. Bull.* 43 (2), 240-244.
- van Baars, A.W., Wierdsma, A.I., Hengeveld, M.W., Mulder, C.L., 2013. Improved insight affects social outcomes in involuntarily committed psychotic patients: a longitudinal study in the Netherlands. *Compr. Psychiatry* 54 (7), 873-879.

- van de Schoot, R., Kluytmans, A., Tummers, L., Lugtig, P., Hox, J., Muthen, B., 2013. Facing off with Scylla and Charybdis: A comparison of scalar, partial, and the novel possibility of approximate measurement invariance. *Front. Psychol.* Vol 4 2013, ArtID 770 4.
- Vandenberg, R.J., Lance, C.E., 2000. A review and synthesis of the measurement invariance literature: Suggestions, practices, and recommendations for organizational research. *Organ. Res. Methods* 3 (1), 4-69.
- Varga, M., Magnusson, A., Flekkøy, K., Rønneberg, U., Opjordsmoen, S., 2006. Insight, symptoms and neurocognition in bipolar I patients. *J. Affect. Disord.* 91 (1), 1-9.
- Ventura, J., Liberman, R.P., Green, M.F., Shaner, A., Mintz, J., 1998. Training and quality assurance with the Structured Clinical Interview for DSM-IV (SCID-I/P). *Psychiatry Res.* 79 (2), 163-173.
- Weiler, M.A., Fleisher, M.H., McArthur-Campbell, D., 2000. Insight and symptom change in schizophrenia and other disorders. *Schizophr. Res.* 45 (1-2), 29-36.
- Wiffen, B.D., Rabinowitz, J., Fleischhacker, W.W., David, A.S., 2010. Insight: demographic differences and associations with one-year outcome in schizophrenia and schizoaffective disorder. *Clin. Schizophr. Relat. Psychoses* 4 (3), 169-175.
- Wong, S.S., Lee, S., Wat, K.H., 1999. A preliminary communication of an insight scale in the assessment of lithium non-adherence among Chinese patients in Hong Kong. *J. Affect. Disord.* 55 (2-3), 241-244.
- Yalincetin, B., Bora, E., Binbay, T., Ulas, H., Akdede, B.B., Alptekin, K., 2017. Formal thought disorder in schizophrenia and bipolar disorder: A systematic review and meta-analysis. *Schizophr. Res.* 185, 2-8.
- Yen, C.F., Chen, C.S., Ko, C.H., Yen, J.Y., Huang, C.F., 2007a. Changes in insight among patients with bipolar I disorder: a 2-year prospective study. *Bipolar disord.* 9 (3), 238-242.
- Yen, C.F., Chen, C.S., Yang, S.J., Ko, C.H., Yen, J.Y., Huang, C.F., 2007b. Relationships between insight and psychosocial adjustment in patients with bipolar I disorder. *Bipolar disord.* 9 (7), 737-742.
- Yen, C.F., Cheng, C.P., Huang, C.F., Ko, C.H., Yen, J.Y., Chang, Y.P., et al., 2009. Relationship between psychosocial adjustment and executive function in patients with bipolar disorder and schizophrenia in remission: the mediating and moderating effects of insight. *Bipolar disord.* 11 (2), 190-197.

- Young, D.A., Campbell, Z., Zakzanis, K.K., Weinstein, E., 2003. A comparison between an interview and a self-report method of insight assessment in chronic schizophrenia. *Schizophr. Res.* 63 (1-2), 103-109.
- Young, R.C., Biggs, J.T., Ziegler, V.E., Meyer, D.A., 1978. A rating scale for mania: reliability, validity and sensitivity. *Br. J. Psychiatry* 133, 429-435.

Table 1
Demographic and clinical characteristics

| | Schizophrenia spectrum disorder (<i>n</i> = 577) | Bipolar I disorder (<i>n</i> = 282) | Bipolar II disorder (<i>n</i> = 138) |
|---|---|---|--|
| Demographics | | | |
| Female, <i>n</i> (%) | 239 (41.4 %) | 161 (56.9 %) | 91 (65.9 %) |
| Age, <i>mean</i> (<i>SD</i>) | 30 (9.6) | 34 (12.0) | 34 (11.5) |
| European, <i>n</i> (%) | 474 (82.1 %) | 259 (91.8 %) | 123 (89.1 %) |
| Education, completed years, <i>mean</i> (<i>SD</i>) | 12.7 (2.7) | 14.5 (3.0) | 14.7 (3.2) |
| Illness history | | | |
| Age at first illness episode, <i>median</i> (<i>SD</i>) | 22 (18, 27) | 21 (16, 27) | 18 (15, 24) |
| Duration of illness, years, <i>median</i> (<i>IQR</i>) | 4 (1, 10) | 9 (4, 16) | 11 (6, 21) |
| History of psychosis, <i>n</i> (%) | 577 (100 %) | 230 (81.6 %) | 38 (27.5 %) |
| Clinical insight | | | |
| IS Awareness of illness, <i>median</i> (<i>IQR</i>) | 2.5 (2, 3) | 3 (2, 3.5) | 3 (2, 3.5) |
| IS Re-labeling of symptoms, <i>median</i> (<i>IQR</i>) | 2.25 (2, 3) | 3 (2, 3.5) | 2.5 (2, 3) |
| IS Need for treatment, <i>median</i> (<i>IQR</i>) | 2.75 (2.25, 3.25) | 3 (2.5, 3.5) | 3 (2.5, 3.5) |
| IS Total, <i>median</i> (<i>IQR</i>) | 8 (6.75, 9) | 8.25 (7, 9.75) | 8 (7, 9) |
| PANSS G12, <i>median</i> (<i>IQR</i>) | 3 (1, 4) | 1 (1, 2) | 1 (1, 1) |
| YMRS 11, <i>median</i> (<i>IQR</i>) | 4 (1, 10) | 1 (.0, 4) | 2 (.0, 4) |
| Current symptoms and functioning | | | |
| YMRS, <i>median</i> (<i>IQR</i>) | 4 (1, 10) | 1 (.0, 4) | 2 (.0, 4) |
| IDS-C, <i>median</i> (<i>IQR</i>) | 16 (8, 26) | 12.5 (6, 21) | 19 (11, 28.25) |
| PANSS Positive, <i>median</i> (<i>IQR</i>) | 15 (11, 19) | 9 (7, 11) | 9 (8, 11) |
| PANSS Negative, <i>median</i> (<i>IQR</i>) | 15 (11, 20) | 9 (7, 12) | 10 (8, 11) |
| PANSS General, <i>median</i> (<i>IQR</i>) | 32 (27, 38) | 24 (21, 28) | 25 (21, 30) |
| PANSS Total, <i>median</i> (<i>IQR</i>) | 64 (52, 74) | 43 (38, 50) | 45 (38, 51) |
| GAF Symptom, <i>mean</i> (<i>SD</i>) | 42.1 (11.6) | 57.7 (12.7) | 57.7 (9.1) |
| GAF Function, <i>mean</i> (<i>SD</i>) | 42.7 (11.0) | 54.4 (13.2) | 62.3 (58.6) |

Note. GAF: Global Assessment of Functioning; IDS-C: Inventory of Depressive Symptoms – Clinical Rated; IQR: inter-quartile range; IS: Insight Scale; PANSS: Positive and Negative Symptom Scale; SD: standard deviation; YMRS: Young Mania Rating Scale

Table 2**Confirmatory factor analysis of two models**

| | χ | df | RMSEA (90% CI) | CFI | TLI | SRMS |
|-------------------------------|-----------------------|----|---------------------|-------|-------|-------|
| <i>Model 1: One factor</i> | | | | | | |
| Schizophrenia Spectrum | 137.50 | 20 | 0.101 (0.085-0.117) | 0.885 | 0.839 | 0.057 |
| Bipolar type I disorder | 79.96 | 20 | 0.103 (0.080-0.127) | 0.826 | 0.756 | 0.060 |
| Bipolar type II disorder | <i>Not converging</i> | | | | | |
| <i>Model 2: Three-factors</i> | | | | | | |
| Schizophrenia Spectrum | 65.18 | 17 | 0.070 (0.053-0.089) | 0.953 | 0.922 | 0.042 |
| Bipolar type I disorder | 45.62 | 17 | 0.077 (0.051-0.105) | 0.917 | 0.863 | 0.047 |
| Bipolar type II disorder | 30.38 | 17 | 0.076 (0.027-0.118) | 0.929 | 0.884 | 0.069 |

Note: CFA: confirmatory factor analysis; RMSEA: root mean square error of approximation; CI: confidence interval; CFI: Comparative Fit Index; TLI: Tucker Lewis Index; SRMR: standardized root mean square residual. Chi-square (χ) statistics: p -values < 0.0001.

Table 3**IS subscale inter-correlations**

| | <u>Schizophrenia Spectrum disorder</u> | | <u>Bipolar type I disorder</u> | | <u>Bipolar type II disorder</u> | |
|---------------------------------|--|---------------------------------|-----------------------------------|---------------------------------|-----------------------------------|---------------------------------|
| | IS Re-labeling of symptoms | IS Aawareness of illness | IS Re-labeling of symptoms | IS Aawareness of illness | IS Re-labeling of symptoms | IS Aawareness of illness |
| IS Aawareness of illness | 0.429** | - | 0.388** | - | 0.316** | - |
| IS Need for treatment | 0.416** | 0.538** | 0.342** | 0.460** | -0.015 | 0.335** |

Note. IS: Insight Scale

**Correlation is significant at the 0.01 level (2-tailed)

Table 4**Bivariate correlations between the IS total scale and subscales, YMRS 11 and PANSS G12**

| | <u>Schizophrenia Spectrum disorder</u> | | <u>Bipolar type I disorder</u> | | <u>Bipolar type II disorder</u> | |
|-----------------------------------|--|----------|--------------------------------|----------|---------------------------------|----------|
| | PANSS G12 | YMRS 11 | PANSS G12 | YMRS 11 | PANSS G12 | YMRS 11 |
| IS full scale | -0.554** | -0.549** | -0.490** | -0.427** | -0.275** | -0.149 |
| IS Re-labeling of symptoms | -0.391** | -0.391** | -0.287** | -0.280** | -0.014 | 0.051 |
| IS Awareness of illness | -0.418** | -0.455** | -0.464** | -0.367** | -0.214* | -0.242** |
| IS Need for treatment | -0.527** | -0.474** | -0.387** | -0.350** | -0.383** | -0.134 |
| PANSS G12 | - | 0.661** | - | 0.642** | - | 0.593** |

Note. IS: Insight Scale; PANSS: Positive and Negative Symptom Scale; YMRS: Young Mania Rating Scale

**Correlation is significant at the 0.01 level (2-tailed)

*Correlation is significant at the 0.05 level (2-tailed)