Chapter 7

The validation of the Seven Up Seven Down (7U7D) among bipolar offspring: screening and prediction of mood disorders. Findings from the Dutch Bipolar Offspring Study

Esther Mesman
Anna de Jong
Naeem K. Juliana
Eric A. Youngstrom
Willem A. Nolen
Manon H.J. Hillegers

In preparation for submission
ABSTRACT

Objective
To validate the Seven Up, Seven Down (7U7D), an abbreviated version of the General Behavior Inventory (GBI) as screener for mood disorders in individuals at risk for bipolar disorder and to test its ability to predict mood disorders in general and bipolar disorder (BD) specifically.

Methods
A total of 108 bipolar offspring from the Dutch Bipolar Offspring Study were assessed at baseline (T1), and at one year (T2), five years (T3) and 12 years (T4) follow-up. Psychopathology was assessed using the K-SADS-PL (T1, T2), the SCID (T3, T4) and the GBI (T1 and T4). Both cross-sectional (T4) and longitudinal analyses (from T1 to T4) were performed using area under the curve (AUC) statistics and logistic regression analyses.

Results
The performance of the 73-item GBI and the 14-item 7U7D was found to be equal. As screener for mood disorders at T4, the 7U7D showed fair diagnostic efficiency for the depression scale (AUC 0.68, \( p < 0.1 \), OR 1.53, 95% CI 1.15-2.03). Among offspring with a mood disorder, positive scores on the hypomania scale were associated with a BD diagnosis (OR 1.64, 95% CI 1.06-2.54). In terms of prediction of future onset of mood disorders between T1 and T4, the depression scale, but not the hypomania scale, was found to be associated with an increased risk for mood disorder onset (AUC 0.67, \( p < 0.5 \); OR 1.47, 95% CI 1.09-2.0). The 7U7D did not achieve statistically significant prediction of bipolar diagnoses.

Limitations
A relatively small sample size and low transition rate to BD.

Conclusions
This study shows further evidence for the potential of the 7U7D as brief screening instrument among individuals at risk for BD. In terms of prediction, the 7U7D may also be beneficial in early detection of mood disorders among bipolar offspring.
INTRODUCTION

Early recognition of bipolar disorder (BD) remains challenging for clinicians. BD presents typically with a (mild) depressive episode during early adolescence/adulthood followed by (hypo)mania years later (e.g. Duffy, Alda, Hajek, & Grof, 2009; Mesman, Nolen, Reichart, Wals, & Hillegers, 2013). The diagnostic delay of BD after the first (hypo)manic episode is on average 5-10 years (Drancourt et al., 2013; Suppes et al., 2001). A prolonged duration of unrecognized and thus untreated BD may have serious consequences, including suicide attempts and poorer long-term outcome (e.g. Drancourt et al., 2013). Therefore, sound methods for early detection of BD, and ultimately prediction of its development, would be beneficial.

Screening instruments are needed to assist clinicians to detect BD in the early phase. Presently, several diagnostic and screening instruments exist to examine mania or depression (Youngstrom, Murray, Johnson, & Findling, 2013). Although, screening instruments for BD are well studied in adult populations screening instruments are less studied during the most critical stage for age of onset of BD, i.e. between 15 and 25 years (Waugh, Meyer, Youngstrom, & Scott, 2014).

The General Behavior inventory (GBI) is a validated self-report instrument used to screen for BD in the general population (Depue et al., 1981; Depue, Krauss, Spoont, & Arbisi, 1989). The GBI is a comprehensive self-report 73-item questionnaire that aims to detect both dimensions of BD (Depue et al., 1981; Depue et al., 1989). The GBI intends to capture the tendency for both threshold and subthreshold affective conditions and their fluctuation over time. In the past decades, the GBI has shown its potential as screening instrument for BD in several adult and adolescent populations in the community and clinic (Danielson, Youngstrom, Findling, & Calabrese, 2003; Depue et al., 1981; Depue et al., 1989; Findling et al., 2002; Klein, Depue, & Slater, 1985; Klein, Depue, & Slater, 1986; Pendergast et al., 2014; Youngstrom et al., 2004; Youngstrom, Findling, Danielson, & Calabrese, 2001). However, its considerable length (approximate completion time 20-40 minutes) reduces its clinical applicability (Youngstrom, Frazier, Demeter, Calabrese, & Findling, 2008; Youngstrom et al., 2013).

Recently, Youngstrom et al. (2013) introduced an abbreviated version of the full-length GBI: the Seven Up, Seven Down GBI (7U7D). The 7U7D is designed and validated for adolescents and adults in the age range 11-86 years. The 7U7D is a 14-item instrument carved from the full-length GBI. Initial findings suggest that the brief 7U7D has good psychometric properties, showing high internal consistency, criterion- and a fair discriminative validity for diagnostic groups among clinical and non-clinical samples. Taken together, the 7U7D appears a promising screening instrument for BD in adolescents and adulthood, however these preliminary findings need replication and have not been validated for high risk populations.

As BD is more prevalent among individuals with a positive family history for BD, offspring of patients affected with BD (hereafter referred to as bipolar offspring) are a
particular interest for screening. In a previous study from our group, we have shown in a Dutch Bipolar Offspring Study that the GBI can both function as a screening instrument to correctly identify mood disorders in bipolar offspring, but also has potential to detect future BD and other mood disorders across a five year interval (Reichart et al., 2004a; Reichart et al., 2005). Scores of the depression scale were significantly higher in offspring who developed a mood disorder or BD within a five-year interval. (Reichart et al., 2005). The purpose of the present study in the Dutch Bipolar Offspring Study is three fold: 1) validation of the 7U7D as compared to the full length GBI 2) to test the utility of the 7U7D to correctly identify offspring with lifetime mood disorders and more specifically BD in a high risk population; and 3) to test the capacity of the 7U7D to predict mood disorders and specifically BD outcome, using the longitudinal design of the Dutch Bipolar Offspring Study with 12-years of follow-up.

METHODS

Population and procedure
Participants originate from the Dutch Bipolar Offspring Study, a prospective study following bipolar offspring from adolescence into adulthood. Details of the study have been described in detail elsewhere (Wals et al., 2001). Briefly, a total of 140 bipolar offspring (mean age 16.1 years, range 12-21) from 86 families with one parent with BD I or II were recruited between 1997 and 1999 and followed for 12 years. A family was only included if all offspring within the age range 12-21 agreed to participate. Exclusion criteria were a severe physical illness or handicap or an IQ below 70. Participants were recruited through the Dutch Association for Manic Depressives and Relatives (62 families; 102 children) and outpatient clinics in nine psychiatric hospitals (24 families; 38 children). All parents with BD were outpatients at time of recruitment. DSM-IV diagnoses of the parents with BD were confirmed by face-to-face interviews using the International Diagnostic Checklist and were confirmed by the clinical diagnosis of the treating psychiatrist. Offspring were assessed at baseline, one-, five- and 12 years of follow-up (T1, T2, T3 and T4 respectively) (Hillegers, 2007; Mesman et al., 2013; Reichart et al., 2004b; Wals et al., 2001). One hundred and thirty-two offspring were reassessed at T2, 129 at T3 and 108 at T4, a retention rate of 77%. At T4, 54% of the offspring were diagnosed with a lifetime mood disorder, including 13% bipolar spectrum disorders. There were no statistically significant demographic or clinical differences between the 108 offspring who completed all 12 years of follow-up and the 32 offspring who dropped out (see Mesman et al. (2013) for details). The Medical Ethics Committee of the University Medical Center Utrecht approved the study. Written informed consent was obtained for both the offspring and their parents.
Instruments

Psychopathology

All psychiatric interviews were administered by intensively trained interviewers with graduate degrees in psychology or by a child and adolescent psychiatrist. All interviews were evaluated with psychiatrists certified in child and adolescent psychiatry as well as adult psychiatry to reach consensus on final diagnoses. At T1 and T2, DSM-IV diagnoses were obtained by a face-to-face interview with both the child and the parent using the Schedule for Affective Disorders and Schizophrenia for School-Age Children-Present and Lifetime Version (K-SADS-PL) (Kaufman et al., 1997). After offspring reached the age of 18, the K-SADS-PL was replaced by the Structured Clinical Interview for DSM-IV Axis I Disorders (SCID) (First, Spitzer, Gibbon, & Williams, 1996). Because the SCID does not cover attention deficit hyperactivity disorder, oppositional defiant disorder, conduct disorder, and tic disorders, sections derived from the K-SADS-PL including the symptoms of these disorders were administered alongside the SCID. Each psychiatric assessment evaluated current and past symptoms during the interim period. Lifetime DSM-IV diagnoses at T4 were based on the psychiatric interviews that took place during all four assessments.

The General Behavior Inventory

The 73-item GBI self-report was administered at each assessment. The GBI entails 73 items and comprises a depression- (46 items) and hypomania-/biphasic scale (21 and 7 items respectively) (Depue et al., 1981; Depue et al., 1989). Each of the 73 items asks the subject to which extent he/she has experienced the symptom or feeling to which the item alludes. The response set is based on a four-point Likert scale; 1 (hardly ever), 2 (sometimes), 3 (often) and 4 (very often). According to Youngstrom et al. (2013), the 14-item 7U7D was extracted from the original full length GBI. The 7U7D follows the same response set as the full length GBI. The seven down (7D) scale and the seven up (7U) reflect the depression and hypomania- and biphasic scale respectively.

Data analysis

In the first set of tests we investigated cross-sectionally the value of the GBI and subsequently the 7U7D as a screener for mood disorders in general and BD specifically in bipolar offspring. Due to the relative low number of transitions to BD it was decided to focus on the latest assessment to capture most transitions (T4). Based upon this assessment offspring were assigned to a diagnostic outcome category, namely: any lifetime mood disorder (AnyMD) versus no lifetime mood disorder diagnosis (NoMD). The mood disorders category was then divided into BD spectrum (BD) and unipolar depression (UD).

In the second set of analyses, we aimed to assess the applicability of the 7U7D to predict development of mood disorders in the future, i.e. during 12 year follow-up (between T1 and T4). As illustrated in Figure 1 a total of 28 offspring developed a first mood disorder between T1 and T4 (NoMD→AnyMD). Twenty-five offspring developed a first unipolar depression and three offspring developed BD between T1 and T4. Overall, during follow-
up 13 cases made a transition to BD between T1 and T4 (NoBD→AnyBD). As the majority of these BD subjects already had a mood disorder diagnosis at T1, this outcome category was compared to offspring with unipolar depression at T1, who did not develop BD during follow-up (UD = UD).

![Figure 1](image_url) Development of psychopathology among bipolar offspring over 12 year follow-up

**Statistical analyses**

Demographic group characteristics were compared using descriptive statistics including one-way ANOVA, $\chi^2$ or Fisher exact tests as appropriate. Using Receiver Operating Characteristic (ROC) analyses, Area Under the Curve (AUC) statistics were calculated to discriminate diagnostic outcome categories and compare the diagnostic efficiency of the GBI and 7U7D. ROC analyses aim to capture the sensitivity and specificity of a test, the AUC quantifies the accuracy of a test. An AUC may be interpreted as the probability that someone with e.g. BD would have a higher score on a test than a randomly selected individual without BD. Some rules of thumb for the interpretation of the diagnostic efficiency of AUCs: $\geq 0.90$ “excellent”, $\geq 0.80$ “good”, $\geq 0.70$ “fair”, $\leq 0.70$ “poor” (Swets, Dawes, & Monahan, 2000; Youngstrom, 2014). Logistic regression analyses were used to calculate odd’s ratios per unit increase of the scores on the depression and hypomanic/biphasic scale of the GBI and seven down and seven up scale of the 7U7D respectively. Moreover, logistic regression analyses allowed us
to check whether there was an incremental value for a combination of the depression and hypomanic/biphasic scale or its abbreviated counterparts. We considered an alpha of $\leq 0.05$ statistically significant. Statistical analyses were performed using IBM SPSS Statistics for Windows, version 20.0 (IBM Corp, Armonk, NY, USA).

RESULTS

The validation of the GBI and 7U7D as a screener for mood disorders in bipolar offspring

Table 1 presents the characteristics, Area Under the Curve (AUCs) statistics and logistic regression analyses on the GBI and 7U7D. No significant differences in gender or age were found across diagnostic outcome categories. The diagnostic efficiency of the GBI and 7U7D in terms of discriminating diagnostic groups were examined using Area Under the Curve (AUCs) statistics. As shown in Table 1, the performance of the GBI and the 7U7D did not differ significantly. As we are especially interested in the performance of a brief screening instrument, further analyses were performed for the 7U7D only.

As screener for lifetime mood disorders in a familial high risk population the diagnostic efficiency of the seven down (7D) scale shows a fair to good performance. Moreover, a fair performance of the seven up (7U) scale was found for identification of BD subjects as compared to offspring not developing BD. However, no significant AUCs were found between participants with BD or unipolar depression.

Additional logistic regression analyses were performed to test whether there was an incremental value for a combination of the scales. The odds ratio for a clinician based mood disorder diagnosis was significant among offspring with a mood disorder 1.53 (95% CI 1.15-2.03) per unit of the 7D scale. No significant odds ratio for the 7U scale was found. Among those offspring with a lifetime mood disorder, a unit increase of the 7U scale increased the odds of lifetime BD diagnosis with 1.64 (95% CI 1.06-2.54), the odds for the 7D scale were not significantly elevated.

Prediction of mood disorders

In this section we explore the predictive value of the GBI and 7U7D for adolescent bipolar offspring to predict BD and mood disorders during follow-up. Descriptive characteristics and AUC analyses are presented in Table 2. Figure 1 illustrates the transition of DSM-IV psychopathology over 12-year follow-up. Again no significant differences across diagnostic outcome categories were found for age or gender. AUC’s showed a poor to fair diagnostic efficiency for prediction of mood disorders over a 12-year interval. Once more, the performance of the GBI and 7U7D was similar and AUC analyses did not differ significantly; hence, further exploration of the data was examined using the 7U7D.
Table 1 | The GBI as screener in a high risk population

<table>
<thead>
<tr>
<th>Demographic characteristics</th>
<th>7U7D</th>
<th>GBI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age at T4</td>
<td>Gender</td>
<td>Seven down (7D)</td>
</tr>
<tr>
<td>All bipolar offspring</td>
<td>102</td>
<td>28.0</td>
</tr>
<tr>
<td>Any mood disorder (AnyMD)</td>
<td>55</td>
<td>28.0</td>
</tr>
<tr>
<td>Bipolar spectrum disorder (BD)</td>
<td>13</td>
<td>27.8</td>
</tr>
<tr>
<td>Unipolar depression (UD)</td>
<td>42</td>
<td>28.5</td>
</tr>
<tr>
<td>Non-mood disorder or no disorders (NoMD)</td>
<td>47</td>
<td>27.4</td>
</tr>
<tr>
<td>Offspring without BD (NoBD)</td>
<td>89</td>
<td>27.9</td>
</tr>
</tbody>
</table>

Comparison of 7U7D and GBI

<table>
<thead>
<tr>
<th>Area Under the Curve (AUC) statistics</th>
<th>7D vs D</th>
<th>7U vs HB</th>
</tr>
</thead>
<tbody>
<tr>
<td>AUC</td>
<td>SE</td>
<td>AUC</td>
</tr>
<tr>
<td>AnyMD vs. NoMD</td>
<td>0.68**</td>
<td>0.05</td>
</tr>
<tr>
<td>BD vs. NoBD</td>
<td>0.73**</td>
<td>0.08</td>
</tr>
<tr>
<td>BD vs. UD</td>
<td>0.66</td>
<td>0.09</td>
</tr>
</tbody>
</table>

Logistic regression

<table>
<thead>
<tr>
<th>Logistic regression</th>
<th>OR</th>
<th>(95% CI)</th>
<th>OR</th>
<th>(95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>AnyMD vs. No MD</td>
<td>1.53*</td>
<td>(1.15-2.03)</td>
<td>0.90</td>
<td>(0.67-1.20)</td>
</tr>
<tr>
<td>BD vs. NoBD</td>
<td>1.26**</td>
<td>(1.01-1.59)</td>
<td>1.35</td>
<td>(0.69-1.89)</td>
</tr>
<tr>
<td>BD vs. UD</td>
<td>1.09</td>
<td>(0.86-1.40)</td>
<td>1.64*</td>
<td>(1.06-2.54)</td>
</tr>
</tbody>
</table>

* p < 0.05, ** p < 0.01, *** p < 0.001
Table 2 | Prediction of future mood disorders

<table>
<thead>
<tr>
<th>Transition from T1 to T4</th>
<th>Demographic characteristics</th>
<th>7U7D</th>
<th>GBI</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Age at T1</td>
<td>Girls</td>
<td>Seven down (7D)</td>
</tr>
<tr>
<td>Descriptives by Dx</td>
<td>n</td>
<td>Mean</td>
<td>SD</td>
</tr>
<tr>
<td>All bipolar offspring</td>
<td>106</td>
<td>16.0</td>
<td>2.7</td>
</tr>
<tr>
<td>NoBD→BD</td>
<td>13</td>
<td>15.8</td>
<td>3.3</td>
</tr>
<tr>
<td>UD = UD</td>
<td>19</td>
<td>16.6</td>
<td>2.2</td>
</tr>
<tr>
<td>NoMD→AnyMD</td>
<td>28</td>
<td>15.9</td>
<td>2.8</td>
</tr>
<tr>
<td>NoMD = NoMD</td>
<td>49</td>
<td>15.8</td>
<td>2.9</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Comparison of 7U7D and GBI</th>
</tr>
</thead>
<tbody>
<tr>
<td>7D vs D</td>
</tr>
<tr>
<td>z</td>
</tr>
<tr>
<td>NoMD→Any MD vs. NoMD = NoMD</td>
</tr>
<tr>
<td>NoBD→BD vs UD = UD</td>
</tr>
</tbody>
</table>

Area Under the Curve (AUC) statistics

<table>
<thead>
<tr>
<th>Logistic regression</th>
</tr>
</thead>
<tbody>
<tr>
<td>Odds Ratio (95% CI)</td>
</tr>
<tr>
<td>NoMD→Any MD vs. NoMD = NoMD</td>
</tr>
<tr>
<td>NoBD→BD vs UD = UD</td>
</tr>
</tbody>
</table>

* p < 0.05, ** p < 0.01, *** p < 0.001
Results from the logistic regression are shown in Table 2. Among offspring (mean age 16) without a mood disorder at baseline (T1), the odds for a transition to mood disorders at follow-up (T4) significantly increased per unit increase at the 7D scale (OR: 1.47, 95% 1.09-2.0). As expected from the AUCs the 7U scale did not significantly contribute to the 7D scale. It was not possible to predict the transition to BD between baseline (T1) and 12-year follow-up (T4) with the 7U7D.

DISCUSSION

In this study we aimed to investigate the validity of the 7U7D as compared to the GBI and its utility to screen for mood disorders and predict future onset of mood disorders in bipolar offspring. Results of this study indicated that the performance of the 73-item GBI and the 14-item 7U7D was found to be equal. As screener for mood disorders among bipolar offspring, the 7U7D showed potential as screener. In terms of prediction of future onset of mood disorders during the 12-year follow-up, the depression scale, but not the hypomania scale, was found to be associated with an increased risk for transitions to mood disorders. Prediction of BD specifically was not possible with the 7U7D.

In line with the study by Youngstrom et al. (Youngstrom et al., 2013) we found that the performance of the full length GBI and 7U7D was found to be equally good. Results of this study provide therefore further validation for the 7U7D as brief screening instrument for use in clinical practice. The first study on the 7U7D focused on a broad clinical and non-clinical population; this study extends the population the 7U7D may be used for, as this study is the first to test the 7U7D among a high risk population for BD.

As screener for mood disorders in bipolar offspring at the mean age of 28 years, we found that the depression scale of the 7U7D was significantly associated with mood disorders among bipolar offspring. This finding is in line with a previous study by our group on the Dutch Bipolar Offspring Study at the mean age of 16 (Reichart et al., 2004a). Among offspring with a mood disorder, especially higher scores on the hypomania scale were indicative for a BD diagnosis. This finding again replicates findings from previous studies and supports the notion that the hypomania scale is beneficial in the identification of BD (Pendergast et al., 2014). The hypomania scale was only associated with BD onset among those offspring with a mood disorder. Suggesting a sequential approach may be useful: use the depression scale to identify those with a mood disorder, and then next use the hypomania scale to detect and search for possible BD.

In terms of prediction, the diagnostic efficiency of the 7U7D was found to be poor to fair over a 12-year interval. However, a positive score on the depression scale was associated with increased odds for a future onset of mood disorders. This finding complements the finding of the 5 year follow-up of the Dutch Bipolar Offspring Study (Reichart et al., 2005). This time, the follow-up interval was 12-years and includes higher numbers of mood disorders in general and BD specifically. As with the 5-year follow-up, the hypomania scale
was not informative in terms of prediction of mood disorders. To the best of our knowledge no other studies on prediction of BD among bipolar offspring and the 7U7D or the GBI exist. Several bipolar offspring studies have now demonstrated that subsyndromal symptoms of mania are an important feature of the BD prodrome (Axelson et al., 2015; Duffy, 2010; Duffy et al., 2014) [Mesman et al., chapter 6]. The fact we do not find an association between the 7U7D hypomania scale and a future transition to BD may have several explanations. At first, the interval of 12-years may be too wide. Moreover, with the relatively low absolute number of BD transitions, the study may be underpowered. However, it is also known that individuals with BD, and probably also offspring during the prodromal phase, normally do not consider symptoms of elevated mood or increased activity as problematic, but perceive them as adaptive and therefore do not report these symptoms spontaneously (i.e. Smith & Ghaemi, 2006). Furthermore, several studies on the GBI have shown that among children and adolescents that parental reports outperform those of adolescents in diagnostic efficiency (Youngstrom et al., 2004), especially in case of hypomania symptoms (Youngstrom et al., 2005). Future studies could benefit from incorporating also the parental GBI. However, until further notice, we recommend that based on the clinical reports as noted above (Axelson et al., 2015; Duffy, 2010; Duffy et al., 2014; Mesman, chapter 6), that even very low scores on the 7U dimension should clinically be alarming and request full clinical assessment and follow-up.

Findings of the present study should be interpreted with caution as sample sizes are limited and only a low absolute number of bipolar offspring developed BD. Therefore larger offspring studies are needed. Also, since we used an enriched population of adolescent bipolar offspring, these findings must be regarded as helpful in identifying those adolescents at high risk to develop mood disorders and BD in particular. Unfortunately, we could not compare the performance of the 7U7D with other BD screening instruments for adolescents and young adults like the Bipolar Spectrum Diagnostic Scale (BSDS) (Ghaemi et al., 2005), Mood Disorder Questionnaire (MDQ) (Hirschfeld et al., 2000) and Hypomania Checklist (HCL-32) (Angst & Cassano, 2005; Waugh et al., 2014). But, the prospective design with long-term follow-up, in combination with the validation of all diagnoses by face to face clinical interview (SCID-I) is a strength of this study.

In conclusion, this study provided further validation for the 7U7D. The 7U7D shows potential as brief screening instrument among individuals at risk for BD. In case of mood disorders, high scores on the hypomania scale may indicate a BD diagnosis. In terms of prediction, the 7U7D may be beneficial in early detection of mood disorders among bipolar offspring, however further research is required.
REFERENCES


Youngstrom, E. A. (2014). A primer on receiver operating characteristic analysis and diagnostic efficiency statistics for pediatric psychology: we are ready to ROC. *J.Pediat.Psychol.*, 39, 204-221.


