Longitudinal and cross-sectional validation of the WERCAP Screen for assessing psychosis risk and conversion

Daniel Mamah
Washington University School of Medicine in St. Louis
Victoria N Mutiso
Africa Mental Health Research and Training Foundation
David M Ndetei
University of Nairobi

Follow this and additional works at: https://digitalcommons.wustl.edu/oa_4

Part of the Medicine and Health Sciences Commons

Please let us know how this document benefits you.

Recommended Citation
https://digitalcommons.wustl.edu/oa_4/1231

This Open Access Publication is brought to you for free and open access by the Open Access Publications at Digital Commons@Becker. It has been accepted for inclusion in 2020-Current year OA Pubs by an authorized administrator of Digital Commons@Becker. For more information, please contact vanam@wustl.edu.
Longitudinal and cross-sectional validation of the WERCAP screen for assessing psychosis risk and conversion

Daniel Mamah a, *, Victoria N. Mutiso b, David M. Ndetei b,c

a Department of Psychiatry, Washington University Medical School, St. Louis, MO, United States of America
b Africa Mental Health Research and Training Foundation, Nairobi, Kenya
c Department of Psychiatry, University of Nairobi, Kenya

ARTICLE INFO

Keywords:
WERCAP
Stress screen
Psychosis
Risk
Kenya
Validation

ABSTRACT

Background: The Washington Early Recognition Center Affectivity and Psychosis (WERCAP) Screen was developed to assess risk for developing psychosis. Its validity has not been investigated in a large population-based study or with longitudinal analyses. Methods: 825 participants, aged 14–25, were recruited from Kenya. Symptoms were assessed using the WERCAP Screen, as experienced over the prior 3-months (3MO), 12-months (12MO) or lifetime (LIF). ROC curve analysis was used to determine the validity of the WERCAP Screen against the Structured Interview of Psychosis-Risk Syndromes. Longitudinal validity was assessed by comparing baseline p-WERCAP scores in psychotic disorder converters and non-converters, and using ROC curve analysis. Relationship of the p-WERCAP was examined against clinical variables. Results: ROC curve analyses against SIPS showed an AUC of 0.83 for 3MO, 0.79 for 12MO and 0.65 for LIF psychosis scores. The optimal cut-point on 3MO was a score of >12 (sens: 0.78; spec: 0.77;ppv: 0.41), and >32 for 12MO (sens: 0.71; spec: 0.74;ppv: 0.24). Baseline 3MO scores (but not LIF scores) were higher in converters compared to high-risk non-converters (p = 0.02). 3MO scores against conversion status had an AUC of 0.75, with an optimal cutoff point of >16 (sens: 1.0; spec: 0.53). All p-WERCAP scores significantly correlated with substance use and stress severity. 12 MO scores were most related to cognitive impairment. Conclusions: The WERCAP Screen is a valid instrument for assessing psychosis severity and conversion risk. It can be used in the community to identify those who may require clinical assessment and care, and for recruitment in psychosis-risk research.

1. Introduction

The schizophrenia prodrome is the period preceding illness onset, which occurs in 75–90% of affected people (Addington et al., 2017; Klosterkotter et al., 2008; Perkins, 2004) and can vary in duration from a few days to several years (Fusar-Poli et al., 2020; Woods et al., 2021). Prodromal symptoms generally involve functional decline and sub-threshold psychotic experiences, as well as depression, anxiety, negative symptoms and/or cognitive deficits. The psychosis clinical high risk (CHR) state was formulated to capture the prodrome and comprises most commonly of attenuated psychotic symptoms (Fusar-Poli et al., 2013). CHR individuals are considered putatively prodromal, although conversion to a psychotic disorder only occurs in 15–30% of cases (Cannon et al., 2008; Ciarleglio et al., 2019). Many CHR youths who do not convert however, continue to have distress and disability. A gold standard for diagnosing those at CHR is the Structured Interview for Psychosis-Risk Syndromes (SIPS), an interviewer-administered assessment which can be time intensive and requires specialized training, characteristics which limit their use in population screening (Kline et al., 2012; Yung et al., 2006). Questionnaire-based methods have the potential to rapidly identify high-risk populations who may require further clinical evaluation. They can also increase the willingness to disclose sensitive information compared with face-to-face interviews (Bowling, 2005). A challenge with self-report instruments, however, is that respondents could incorrectly complete items which are difficult to comprehend. It is therefore essential that questionnaire items

* Corresponding author at: Department of Psychiatry (Box 8134), Washington University School of Medicine, 660 S. Euclid, St. Louis, MO 63110, United States of America.
E-mail address: mamahd@wustl.edu (D. Mamah).

https://doi.org/10.1016/j.schres.2022.01.031
Received 19 August 2021; Received in revised form 5 January 2022; Accepted 16 January 2022
Available online 7 February 2022
0920-9964/© 2022 The Authors. Published by Elsevier B.V. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).
are culturally applicable to the target population.

Several self-report tools have been developed for identifying the psychosis-risk state, which apply varying assessment methods. For example, the Prodromal Questionnaire (PQ) consists of 92-items in a true/false format across positive, negative, disorganized and general symptom dimensions (Loewy et al., 2005); and there is also a 16-item version of the PQ (de Jong et al., 2018). The 21-item Brief Prodromal Questionnaire (PQ-B) consists of 'yes/no' questions, and includes a five-point Likert scale probing associated distress (Loewy et al., 2011). Other screeners using primarily 'yes/no' questions include the Youth Psychosis At-Risk Questionnaire, which has a 92-item version (Ord et al., 2004) and a 28-item version (Fonseca-Pedrero et al., 2017), the 32-item Self-Screen Prodrome (Muller et al., 2010) and the 15-item Composite Psychosis Risk Questionnaire (Liu et al., 2013). The 12-item Prime Screen-Revised (Kline et al., 2012; Kobayashi et al., 2008; Owoso et al., 2014) and the 40-item Eppendorf Schizophrenia Inventory (Niessen et al., 2010) use Likert scales to assess the degree of symptom 'trueness'. Other questionnaires assess symptom severity, such as the 21-item PROD-Screen (Heinimaa et al., 2003), which probes both lifetime and 12-month symptoms, and the 42-item Community Assessment of Psychic Experiences (Mossaheb et al., 2012), which probes positive, depressive and negative symptoms using measures of symptom frequency and degree of distress.

The Washington Early Recognition Center Affectivity and Psychosis (WERCAP) Screen was developed to evaluate the risk for psychotic and bipolar disorders and to be cross-culturally applicable (Mamah et al., 2014). It assesses symptom severity using both frequency of occurrence and degree of functional impairment, to identify even subtle psychotic experiences. The psychotic section of the WERCAP Screen (p-WERCAP) has been validated against the SIPS in a small sample of U.S. subjects, which found very high sensitivity (0.89) and specificity (1.0) (Mamah et al., 2014). It has been used in both U.S. (Hsieh et al., 2016; Mamah et al., 2014) and African (Mamah et al., 2020; Mamah et al., 2016; Mamah et al., 2021b; Ndetei et al., 2019; Owoso et al., 2018) studies, but has not been previously validated in a large community sample.

The current study investigates the psychometric properties of the p-WERCAP in 825 Kenyan adolescents and young adults. We explore the validity of the p-WERCAP cross-sectionally against the SIPS, as well as other clinical variables which are commonly associated with psychotic disorders (including substance use, cognition and perceived stress). In addition, using data from a two-year longitudinal study, we report the utility of the p-WERCAP in predicting psychosis conversion, which is rarely done in questionnaire validation studies.

2. Methods

2.1. Recruitment

The study included two cohorts of Kenyan adolescents and young adults, as depicted in the flowchart in Fig. 1. In Cohort 1, 285 participants were selected from among 2800 students from Machakos county in the 10th–12th grades of study, aged 14–20 years (mean: 17.3 years). All high psychosis scorers (i.e. ≥30) on the WERCAP Screen (Mamah, 2011) were selected for the study. In addition, a comparable number of participants were selected to span the 0–29 score distribution relatively evenly. Cohort 2 participants were recruited from Nairobi county (largely urban) and Machakos, Kitui and Makueni counties (largely rural). 87% of Cohort 2 participants were recruited from tertiary academic institutions (i.e. eight colleges and one public university) and...
13% were recruited directly through community outreach efforts. In this cohort, 540 participants were selected from among 9564 youths using the WERC Screen (Mamah et al., 2021b). Community youth were directed to specific public meeting areas for assessments, with the help of local community leaders. None of the subjects were help-seeking or recruited from a clinical setting. All high scorers were selected for the study. Low scorers were randomly selected across the low score distribution. Cohort 2 participants were aged 15–25 years (mean: 21.2 years). The mean age of community youth was slightly lower (19.2 years) than that of tertiary school students (21.4 years).

Written consent was provided by participants or their guardians, and written assent was obtained from minors. The study was approved by the ethical review boards of the Kenya Medical Research Institute (Cohort 1) and the Maseno University, Kenya (Cohort 2), as well as the Institutional Review Board of Washington University in St. Louis (both cohorts).

2.2. Psychosis assessment with the WERCAP screen

The WERCAP Screen (Mamah, 2011) is a 16-item self-report questionnaire which assesses the severity of mood dysregulation and psychosis. It was developed with the goal of being a cross-culturally applicable questionnaire, by using terminology that can be similarly understood in the United States and Africa (Mamah et al., 2013). For each symptom item, it estimates the frequency of occurrence on a six-point scale (ranging from ‘no’ to ‘almost always’), and for most symptoms, also their effect on functioning on a four-point scale (ranging from ‘not at all’ to ‘severely’) (Hsieh et al., 2016; Mamah et al., 2014; Ndeti et al., 2019). The first eight items probe mood (or ‘affectivity’) symptom severity (a-WERCAP), and the latter eight, psychotic symptom severity (p-WERCAP). Total scores in each of the two symptom domains are derived as a sum of their constituent items. The maximum score on the p-WERCAP is 64.

Symptom time frames are manually specified on the WERCAP Screen. In the current study, symptoms in Cohort 1 were assessed separately over a lifetime (LIF) and over the last 3-months (3MO). In Cohort 2, symptoms were assessed over the last 12-months (12MO) and 3MO.

2.3. Other clinical assessments

The SIPS (McGlashan et al., 2010) was administered to each participant by a trained interviewer. It identifies CHR status based on either attenuated psychotic symptoms (APSS), brief limited intermittent psychotic episodes, and/or a genetic risk and deterioration syndrome. All CHR cases ascertained in this study were found to have exclusively APSS. Previous studies have shown strong to moderate inter-rater reliability in Kenya across SIPS positive symptom items (Owoso et al., 2014).

Lifetime substance use was measured with the WHO Alcohol, Smoking and Substance Involvement Screening Test (ASSIST) (Group, 2002). The WERC Stress Screen, a self-report questionnaire, was used to assess perceived stress severity (Hsieh et al., 2016; Mamah et al., 2014). Disability was measured using the WHO Disability Assessment Schedule (WHODAS 2.0) (WHO, 2010).

The Penn Computerized Neurocognitive Battery (PennCNB) (Gur et al., 2010; Moore et al., 2015) was administered using a portable laptop computer. Test modules (and domains measured) included the: 1) continuous performance test (attention), 2) letter n-back (working memory), 3) word memory test (verbal memory), 4) visual object learning test (visual memory), 5) verbal reasoning test (verbal reasoning), 6) motor praxis test (sensorimotor processing), 7) penn matrix reasoning test (abstraction), and 8) emotion recognition test (emotion recognition).

2.4. Longitudinal assessment

Cohort 1 participants were part of a longitudinal study investigating psychosis conversion over a 20-month period (Mamah et al., 2016). Psychosis conversion was defined as meeting the psychosis syndrome criteria on the SIPS. Five participants converted to a psychotic disorder over that time period. Cohort 2 participants did not have a longitudinal study component.

2.5. Statistical analysis

Most statistical analyses were done using SAS 9.4 (SAS Institute Inc., Cary NC). The Receiver Operating Characteristic (ROC) curve analyses were done using MedCalc, version 12.7.7.0 (MedCalc Software, Ostend, Belgium).

Cross-sectional validity analysis of each timeframe on the p-WERCAP (3MO, 12MO, and LIF) aimed to determine their relative agreement with the SIPS-obtained CHR classification. The area under each ROC curve (AUC) was interpreted as the probability that a randomly chosen respondent with CHR or without CHR would be correctly distinguished based on their screening scale scores (Hanley and McNeil, 1982). Additional validity indices included examination of Spearman correlations to determine the relatedness of scores with lifetime substance use, stress, disability and cognition. Performance on cognitive tests was determined as previously described (Mamah et al., 2021a). Z-scores were calculated separately for each cohort.

Longitudinal validity was assessed in Cohort 1 subjects by comparing baseline p-WERCAP scores (3MO and LIF) in psychotic disorder converters and all nonconverters. AUC of psychosis score ROC curves were used to determine the optimal p-WERCAP cut-point for conversion.

3. Results

3.1. Demographics and distribution of psychosis scores

Table 1 shows the demographics of each psychosis timeframe group. Cohort 1 participants were younger than Cohort 2 participants, and consisted of 59% females compared to Cohort 2 (49%).

Fig. 2 shows the frequency distribution of p-WERCAP scores by psychosis timeframe. In Cohort 1, mean (s.d.) of 3MO p-WERCAP scores was 13.4 (13.5) and median was 11. The mean LIF p-WERCAP score was 24.9 (14.2) and the median 27. In Cohort 2, mean (s.d.) of 3MO p-WERCAP scores was 8.9 (11.1) and median was 4; while the mean 12MO...
A. COHORT 1

3-month scores

Fig. 2. Distribution of p-WERCAP scores across the two youth cohorts. Figures depict the prevalences of different p-WERCAP score ranges. Cohort 1 participants completed WERCAP screens with 3-month and lifetime symptom timeframes (A). Cohort 2 participants completed WERCAP Screens with 3-month and 12-month symptom timeframes (B).

B. COHORT 2

3-month scores

12-month scores

Fig. 3. Comparison of p-WERCAP ROC curves by symptom timeframe. Figures show ROC curves involving p-WERCAP scores against Structured Interview of Psychosis-Risk Syndromes (SIPS)-based clinical high risk (CHR) classification. In (A) ROC curves for 3-month symptom timeframe p-WERCAP scores are compared between Cohort 1 (blue) and Cohort 2 (green) participants. In (B) ROC curves for lifetime symptom timeframe p-WERCAP scores from Cohort 1 (blue) participants are compared to 12-month symptom timeframe p-WERCAP scores from Cohort 2 (green) participants. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)
p-WERCAP score was 18.9 (18.1) and the median 30.

3.2. Internal consistency of the p-WERCAP

Standardized Cronbach's alpha coefficients were 0.90 for LIF p-WERCAP scores \((n = 285)\), 0.91 for 12MO scores \((n = 540)\), and 0.90 for 3MO scores \((n = 825)\).

3.3. ROC curve analyses against SIPS

Fig. 3A, compares the 3MO p-WERCAP ROC curves generated separately from cohorts 1 and 2, and show similar AUC values of 0.81 and 0.83 respectively. When the two cohorts were combined \((n = 825)\), the AUC was 0.83. The optimal cut-point on the 3MO p-WERCAP was a score of 12. At this cut-point, sensitivity was 77.9%; specificity 76.6%, PPV 40.7% and NPV 94.4%.

As seen in Fig. 3B, the ROC curve for the LIF p-WERCAP had an AUC of 0.65, and the 12MO p-WERCAP had an AUC of 0.79. The optimal cut-point on the 12MO p-WERCAP was a score of 32. At this cut point, sensitivity was 71.4% and specificity was 73.8%. The PPV was 24.3% and the NPV was 95.6%.

Criterion values at each 12MO and 3MO p-WERCAP score are shown in Supplementary Tables 1 and 2.

3.4. Conversion to psychotic disorder: ANOVA and ROC analysis

We assessed baseline p-WERCAP scores of the five participants who converted to psychosis within 20-months, high-risk (HR) participants who did not convert, and control participants. As seen in Fig. 4, average 3MO p-WERCAP scores showed a significant group difference \((F = 37.7; p < 0.0001)\). Post-hoc analysis showed significant group effects between converters and either controls \((p < 0.0001)\) or HR non-converters \((p = 0.021)\). Average LIF p-WERCAP scores also showed group differences \((F = 95.3; p < 0.0001)\), with post-hoc analysis finding significant effects between converters and controls \((p = 0.013)\) but not between converters and HR non-converters \((p = 0.7)\).

The ROC curve for the 3MO p-WERCAP scores against conversion status had an AUC of 0.75 \((p = 0.01)\) (Fig. 5). The optimal cut-point on

![ROC curve](image)

Fig. 5. 3-month p-WERCAP ROC curves against psychosis conversion.

The figure depicts the ROC curve generate using scores from the 3-month symptom timeframe p-WERCAP against psychosis conversion over a 20-month longitudinal study. Psychosis converters \((n = 5)\) and non-converters \((n = 272)\) were ascertained using the SIPS and the Diagnostic Interview Schedule.
the 3MO p-WERCAP was a score of 16. At this cut-point, sensitivity was 100%, specificity 52.7%, PPV 3.7% and NPV 100%. Criterion values at each p-WERCAP score are shown in Supplementary Table 3. The ROC curve for the LIF p-WERCAP scores against conversion status had an AUC of 0.68, but did not meet statistical significance \( (p = 0.08) \).

### 3.5. Substance use relationship

We evaluated p-WERCAP relationships with use histories of the four most common substances in Kenya. The 3MO p-WERCAP scores correlated with a lifetime use history of tobacco \( (r_s = 0.20; p < 0.0001) \), alcohol \( (r_s = 0.12; p = 0.006) \), marijuana \( (r_s = 0.22; p < 0.0001) \) and khat \( (r_s = 0.15; p = 0.0008) \). Use history correlated with 12MO p-WERCAP scores for tobacco \( (r_s = 0.16; p = 0.0004) \), alcohol \( (r_s = 0.10; p = 0.03) \), and marijuana \( (r_s = 0.14; p = 0.0006) \). There was no significant relationship between khat use and 12MO p-WERCAP scores \( (p = 0.24) \).

### 3.6. Disability and stress relationships

Disability, assessed using the WHODAS was related to both 3-month p-WERCAP scores \( (r_s = 0.32; p < 0.0001) \) and 12MO p-WERCAP scores \( (r_s = 0.32; p < 0.0001) \).

The WERC Stress Screen \( (\text{Mamah et al., 2014}) \) score (Cohorts 1 and 2) correlated with 3MO p-WERCAP scores \( (r_s = 0.46; p < 0.0001) \), 12MO p-WERCAP scores \( (r_s = 0.61; p < 0.0001) \) and LIF p-WERCAP scores \( (r_s = 0.32; p < 0.0001) \).

### 3.7. Cognitive relationships

Total cognitive scores inversely correlated with 12MO p-WERCAP scores \( (r_s = 0.14; p = 0.016) \), but not with 3MO \( (r_s = 0.05; p = 0.25) \) and LIF \( (r_s = 0.01; p = 0.8) \) p-WERCAP scores. Relationships with 3MO psychosis scores were still not significant when investigated separately in cohort 1 \( (r_s = 0.03; p = 0.6) \) and in cohort 2 \( (r_s = 0.03; p = 0.06) \). Relationships between each cognitive domain performance and p-WERCAP scores are shown in Table 2. The 12MO p-WERCAP scores correlated with verbal memory \( (r_s = 0.12; p = 0.04) \) and sensorimotor processing \( (r_s = -0.12; p = 0.03) \), and showed trend level relationships with verbal reasoning and emotional recognition. LIF p-WERCAP scores correlated significantly with sensorimotor processing, showing better sensorimotor processing with increasing p-WERCAP scores. There were no significant relationships between 3MO p-WERCAP scores and cognitive scores when each dataset was analyzed separately.

Clinical and cognitive characteristics of the three pWERCAP high risk groups compared to the SIPS high risk group are shown in Table 3.

### 4. Discussion

Our study investigated the validity of the p-WERCAP in a large community youth population. Over a 3-month symptom timeframe, it showed an excellent AUC of 0.83 against the SIPS in classifying those at CHR for psychosis. Over a 12-month symptom timeframe, it had a slightly lower AUC of 0.79, and over a lifetime timeframe, the AUC was only 0.65. Taken together, we found that validity of the p-WERCAP for CHR classification is better when symptoms are probed over shorter time frames, with the 3- and 12-month timeframes being optimal. It is notable, that the severity of psychotic experiences reported by participants over a 3-month period was less than half of that reported over longer timeframes. The optimum cut-off score on the 3-month p-WERCAP in this study was 12, compared to that of the 12-month p-WERCAP which was 32 and similar to the 30 cut-off on the lifetime p-WERCAP found in an earlier US study \( (\text{Mamah et al., 2014}) \). Higher scores with symptoms probed over longer time frames may be due to less precision remembering details of distant events, or disproportional weighting given to the most severe symptomatic period within the timeframe. Many of those with very high recent psychosis scores \( (i.e. \text{in last 3 months}) \) may also not be available for the study due to distress or other limitations.

Our longitudinal analysis found that psychosis converters had significantly higher baseline 3-month p-WERCAP scores compared to

### Table 2

<table>
<thead>
<tr>
<th>Cognitive domain</th>
<th>3-month psychosis(^a)</th>
<th>12-month psychosis(^b)</th>
<th>Lifetime psychosis(^c)</th>
<th>( p^a )</th>
<th>( p^b )</th>
<th>( p^c )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Visual memory</td>
<td>–0.08</td>
<td>–</td>
<td>–0.03</td>
<td>0.2</td>
<td>–</td>
<td>0.6</td>
</tr>
<tr>
<td>Dataset 1</td>
<td>0.01</td>
<td>–</td>
<td>–</td>
<td>0.8</td>
<td>0.7</td>
<td>–</td>
</tr>
<tr>
<td>Attention</td>
<td>–0.07</td>
<td>–</td>
<td>–0.11</td>
<td>0.8</td>
<td>–</td>
<td>0.07*</td>
</tr>
<tr>
<td>Dataset 2</td>
<td>–0.05</td>
<td>–</td>
<td>–0.04</td>
<td>0.4</td>
<td>0.6</td>
<td>–</td>
</tr>
<tr>
<td>Verbal memory</td>
<td>–0.09</td>
<td>–</td>
<td>–0.02</td>
<td>0.04**</td>
<td>–</td>
<td>0.8</td>
</tr>
<tr>
<td>Dataset 1</td>
<td>–0.03</td>
<td>–</td>
<td>–0.12</td>
<td>0.6</td>
<td>–</td>
<td>0.04**</td>
</tr>
<tr>
<td>Dataset 2</td>
<td>–0.02</td>
<td>–</td>
<td>–0.07</td>
<td>0.7</td>
<td>–</td>
<td>0.3</td>
</tr>
<tr>
<td>Working memory</td>
<td>0.02</td>
<td>–0.06</td>
<td>–</td>
<td>0.8</td>
<td>0.3</td>
<td>–</td>
</tr>
<tr>
<td>Sensorimotor processing</td>
<td>0.09</td>
<td>–</td>
<td>0.13</td>
<td>0.12</td>
<td>–</td>
<td>0.035**</td>
</tr>
<tr>
<td>Dataset 2</td>
<td>–0.05</td>
<td>–0.13</td>
<td>–</td>
<td>0.4</td>
<td>0.032**</td>
<td>–</td>
</tr>
<tr>
<td>Abstraction</td>
<td>0.02</td>
<td>0.03</td>
<td>0.11</td>
<td>0.7</td>
<td>–</td>
<td>0.07*</td>
</tr>
<tr>
<td>Dataset 1</td>
<td>0.09</td>
<td>0.03</td>
<td>–</td>
<td>0.13</td>
<td>0.6</td>
<td>–</td>
</tr>
<tr>
<td>Verbal reasoning</td>
<td>–0.01</td>
<td>–</td>
<td>–0.04</td>
<td>0.8</td>
<td>–</td>
<td>0.5</td>
</tr>
<tr>
<td>Dataset 2</td>
<td>–0.06</td>
<td>–0.11</td>
<td>–</td>
<td>0.4</td>
<td>0.07*</td>
<td>–</td>
</tr>
<tr>
<td>Emotion recognition</td>
<td>–0.04</td>
<td>–</td>
<td>–0.07</td>
<td>0.5</td>
<td>–</td>
<td>0.2</td>
</tr>
<tr>
<td>Dataset 2</td>
<td>–0.03</td>
<td>–0.11</td>
<td>–</td>
<td>0.6</td>
<td>0.07*</td>
<td>–</td>
</tr>
</tbody>
</table>

Values are given as Spearman’s rank correlation coefficient \( (p) \). Dataset 1 \( (n = 280) \). Dataset 2 \( (n = 296) \).

\(^a\) \( p < 0.1 \).
\(^b\) \( p < 0.05 \) and bolded.
\(^c\) Relates to 3-month psychosis.
\(^*\) Relates to 12-month psychosis.
\(^*\) Relates to lifetime psychosis.
Table 3
Clinical characteristics in those that met high-risk criteria using the p-WERCAP and SIPS.

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>3-month P- WERCAP+ (n = 192)</th>
<th>12-month P- WERCAP+ (n = 155)</th>
<th>Lifetime P- WERCAP+ (n = 108)</th>
<th>SIPS+ (n = 118)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Positive symptoms</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Unusual thought</td>
<td>1.48 (1.3)</td>
<td>0.92 (1.1)</td>
<td>1.16 (1.2)</td>
<td>2.16 (1.3)</td>
</tr>
<tr>
<td>Persecutory</td>
<td>1.35 (1.3)</td>
<td>1.03 (1.2)</td>
<td>1.00 (1.2)</td>
<td>2.23 (1.3)</td>
</tr>
<tr>
<td>Grandiosity</td>
<td>1.17 (1.2)</td>
<td>0.69 (1.0)</td>
<td>1.08 (1.1)</td>
<td>1.83 (1.3)</td>
</tr>
<tr>
<td>Hallucinations</td>
<td>1.40 (1.5)</td>
<td>0.70 (1.2)</td>
<td>1.14 (1.4)</td>
<td>2.21 (1.5)</td>
</tr>
<tr>
<td>Disorg. communication</td>
<td>0.61 (1.0)</td>
<td>0.36 (0.8)</td>
<td>0.46 (0.9)</td>
<td>0.86 (1.2)</td>
</tr>
<tr>
<td>Negative symptoms</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Social anhedonia</td>
<td>1.00 (1.5)</td>
<td>0.71 (1.2)</td>
<td>0.84 (1.4)</td>
<td>1.51 (1.8)</td>
</tr>
<tr>
<td>Avolition</td>
<td>0.47 (1.0)</td>
<td>0.26 (0.8)</td>
<td>0.31 (0.8)</td>
<td>0.64 (1.2)</td>
</tr>
<tr>
<td>Emotion expression</td>
<td>0.41 (0.9)</td>
<td>0.34 (0.9)</td>
<td>0.23 (0.7)</td>
<td>0.61 (1.1)</td>
</tr>
<tr>
<td>Emotion/self-experience</td>
<td>0.30 (0.8)</td>
<td>0.22 (0.7)</td>
<td>0.25 (0.8)</td>
<td>0.46 (1.0)</td>
</tr>
<tr>
<td>Difficulty understanding</td>
<td>0.77 (1.2)</td>
<td>0.50 (1.0)</td>
<td>1.24 (1.3)</td>
<td>0.83 (1.1)</td>
</tr>
<tr>
<td>Occupational functioning</td>
<td>0.45 (0.8)</td>
<td>0.22 (0.5)</td>
<td>0.52 (1.0)</td>
<td>0.57 (0.9)</td>
</tr>
<tr>
<td>General symptoms</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sleep disturbance</td>
<td>0.49 (1.3)</td>
<td>0.33 (0.8)</td>
<td>0.34 (0.8)</td>
<td>0.62 (1.1)</td>
</tr>
<tr>
<td>Dysphoric mood</td>
<td>1.44 (1.8)</td>
<td>0.96 (1.4)</td>
<td>0.98 (1.6)</td>
<td>1.91 (1.9)</td>
</tr>
<tr>
<td>Motor disturbances</td>
<td>0.16 (0.6)</td>
<td>0.09 (0.4)</td>
<td>0.14 (0.5)</td>
<td>0.26 (0.8)</td>
</tr>
<tr>
<td>Stress tolerance</td>
<td>0.45 (0.9)</td>
<td>0.26 (0.7)</td>
<td>0.30 (0.7)</td>
<td>0.59 (1.0)</td>
</tr>
<tr>
<td>Cognitive functioning (z)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Visual memory</td>
<td>−0.05 (0.9)</td>
<td>−0.11 (1.0)</td>
<td>−0.02 (1.1)</td>
<td>0.01 (1.0)</td>
</tr>
<tr>
<td>Attention</td>
<td>−0.13 (1.0)</td>
<td>−0.17 (1.1)</td>
<td>−0.17 (1.0)</td>
<td>−0.02 (0.9)</td>
</tr>
<tr>
<td>Verbal memory</td>
<td>−0.01 (1.0)</td>
<td>−0.05 (1.0)</td>
<td>−0.04 (1.1)</td>
<td>−0.07 (1.0)</td>
</tr>
<tr>
<td>Working memory</td>
<td>0.05 (1.0)</td>
<td>0.07 (0.9)</td>
<td>−0.09 (1.0)</td>
<td>0.19 (0.8)</td>
</tr>
<tr>
<td>Sensorimotor processing</td>
<td>−0.13 (1.0)</td>
<td>0.09 (1.2)</td>
<td>0.17 (0.8)</td>
<td>0.27 (0.8)</td>
</tr>
<tr>
<td>Abstraction</td>
<td>0.20 (1.0)</td>
<td>0.31 (1.0)</td>
<td>0.10 (1.1)</td>
<td>0.47 (1.0)</td>
</tr>
<tr>
<td>Verbal Reasoning</td>
<td>0.14 (1.1)</td>
<td>0.37 (0.9)</td>
<td>−0.02 (1.2)</td>
<td>0.23 (1.0)</td>
</tr>
<tr>
<td>Emotion recognition</td>
<td>0.23 (0.8)</td>
<td>0.34 (0.9)</td>
<td>0.02 (0.8)</td>
<td>0.36 (0.8)</td>
</tr>
<tr>
<td>Lifetime substance use</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tobacco</td>
<td>28 (21.3%)</td>
<td>44 (16.1%)</td>
<td>n/a</td>
<td>16 (23.2%)</td>
</tr>
<tr>
<td>Alcohol</td>
<td>43 (32.3%)</td>
<td>78 (60.5%)</td>
<td>n/a</td>
<td>22 (31.9%)</td>
</tr>
<tr>
<td>Marijuana</td>
<td>24 (18.1%)</td>
<td>33 (12.1%)</td>
<td>n/a</td>
<td>13 (18.8%)</td>
</tr>
<tr>
<td>Khat</td>
<td>11 (8.3%)</td>
<td>15 (5.5%)</td>
<td>n/a</td>
<td>5 (7.3%)</td>
</tr>
<tr>
<td>Stress</td>
<td>52.6 (34.5)</td>
<td>54.3 (34.1)</td>
<td>39.3 (29.8)</td>
<td>49.1 (37.2)</td>
</tr>
<tr>
<td>Disability</td>
<td>14.6 (8.6)</td>
<td>10.1 (8.7)</td>
<td>n/a</td>
<td>13.2 (8.8)</td>
</tr>
</tbody>
</table>

SIPS+ participants are those who met high clinical high-risk (CHR) criteria for psychosis on the SIPS. 3-month p-WERCAP+ participants are those who scored ≥15 on the 3 MO p-WERCAP Screen. 12-month p-WERCAP+ and Lifetime p-WERCAP+ participants are those who scored ≥30 on the 12 MO or LIF p-WERCAP Screens respectively. Values are given as means (s.d.) or number per group (%). n/a = not applicable, data not collected. * 3-month psychosis scores were obtained from both cohorts. Cohort 1 also collected lifetime psychosis scores, and cohort 2 also collected 12-month scores. Thus, participants with lifetime psychosis scores are a sum of participants with 12-month and lifetime psychosis scores.

high-risk non-converters, with a high AUC observed against conversion status. Scores higher than 16 were found to be the optimum cut-point for predicting conversion. Lifetime p-WERCAP scores however, were similar between converters and high-risk non-converters, suggesting that a lifetime symptom timeframe has low utility in risk prediction. Taking cross-sectional and longitudinal studies together, we recommend a cut-off score of 15 on the 3-month p-WERCAP for community screening to identify those at high psychosis risk. For the 12-month p-WERCAP, a cut-point of 30 is recommended, however longitudinal validation against conversion has not been done with this symptom timeframe. The 12-month p-WERCAP may be better suited for identifying cumulative brain insults and those with psychotic disorders (Hsieh et al., 2016) who may have received treatment which can obscure symptoms.

We also investigated other clinical relationships to p-WERCAP scores. 3-month p-WERCAP scores were related to substance use history, consistent with the observed comorbidity of substance use with psychotic disorders (Buchy et al., 2015; Carney et al., 2017) and the CHR state (Khokhar et al., 2018). Some (Addington et al., 2014; Cannon et al., 2008), but not all (Addington et al., 2014; Buchy et al., 2014) authors have also reported substance use as a predictor of psychosis conversion in CHR individuals. The 3-month p-WERCAP had the largest effect size for disability compared to other symptom timeframe groups, while the 12-month p-WERCAP scores had the strongest relationship to stress severity and cognitive functioning. These differential effect sizes across symptom timeframes would have to be replicated, but it suggests that cognitive functioning and HPA axis dysfunction may be markers of a more longstanding illness, while disability is more reflective of the presence of recent psychotic symptoms.

Community psychotic symptom screening has been underutilized, in spite of the known benefits of early intervention for improving long-term disability in psychotic disorders (Haas et al., 1998; Marshall et al., 2005). In the United States, the average duration of untreated psychosis (DUP), the time between the first psychotic break and anti-psychotic treatment, is between 1 and 3 years (Addington et al., 2015). The DUP in developing countries is even longer, and many of those with psychotic illnesses are never treated (Farooq et al., 2009). Universal screening of adolescents and young adults for psychotic symptoms would help identify those at psychosis risk or with an untreated psychotic disorder. This could be done directly within the community or in primary care clinics, where applicable. Behavioral health screening by general medical practitioners usually includes depression, anxiety, attention and suicide risk, but rarely psychotic symptoms. Periodic screening in clinics could be facilitated by the availability of a valid psychosis screening tool with clear symptom thresholds linked to guidelines about further assessment, management and specialist referral (Kennedy et al., 2019). The 3-month WERCAP Screen appears well-suited for this purpose, as it provides quantitative measures of severity and cognitive functioning. These differential effect sizes across symptom timeframes would have to be replicated, but it suggests that cognitive functioning and HPA axis dysfunction may be markers of a more longstanding illness, while disability is more reflective of the presence of recent psychotic symptoms.

Universal psychosis screening tool with clear symptom thresholds linked to guidelines about further assessment, management and specialist referral (Kennedy et al., 2019). The 3-month WERCAP Screen appears well-suited for this purpose, as it provides quantitative measures of severity and cognitive functioning. These differential effect sizes across symptom timeframes would have to be replicated, but it suggests that cognitive functioning and HPA axis dysfunction may be markers of a more longstanding illness, while disability is more reflective of the presence of recent psychotic symptoms.
diagnostic tool, and most high scorers will not convert to a psychotic disorder. The utility of the p-WERCAP lies in rapidly identifying community youth who require further evaluation to ascertain clinical status, and to monitor change in symptoms over time. Thirdly, the items on the p-WERCAP do not include all symptoms relevant to psychosis risk prediction (Ciocco et al., 2014). Other symptom domains such as negative symptoms or cognition are likely relevant and may increase the effectiveness of psychosis-risk screening tools (Cannon et al., 2008; Cannon et al., 2016; Carrion et al., 2016; Ellman et al., 2020; Woods et al., 2009). Improvements in psychotic disorder prediction has been reported by combining clinical symptoms with cognitive markers (Kotsouleris et al., 2012; Riecher-Rossler et al., 2009), electrophysiological measures (van Tricht et al., 2011), specific environmental factors (Dragt et al., 2009; Mechelli et al., 2009; Mechelli et al., 2011) or cortisol secretion (Walker et al., 2013). Risk symptoms used in combination with other measures are therefore likely to be the most useful for predicting psychosis risk.

In summary, our studies demonstrate the validity of the psychosis section of the WERCAP Screen in a large population of adolescents and young adults. We found that psychosis scores reported over 3- or 12-month timeframes were highly related to CHR status, and 3-month symptoms were most predictive of psychosis conversion. Findings support the use of the WERCAP screen for psychosis-risk screening for clinical and research purposes.

Role of funding source

This work was funded primarily by NIMH grant R56 MH111300 and R21 MH095645. Additionally, Dr. Mamah has received funding from Taylor Family Institute, Dept. Psychiatry, Washington University; and the Center for Brain Research on Mood Disorders, Dept. Psychiatry, Washington University.

Contributors

All authors contributed to the drafting and editing of the manuscript.

Declaration of competing interest

The authors do not have any conflicts to report.

Acknowledgments

None.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.schres.2022.01.031.

References


