Accepted Manuscript

The Impact of Age on the Validity of Psychosis-Risk Screening in a Sample of Help-Seeking Youth

Pamela J. Rakhshan Rouhakhtar , Steven C. Pitts , Zachary B. Millman , Nicole D. Andorko , Samantha Redman , Camille Wilson , Caroline Demro , Peter L. Phalen , Barbara Walsh , Scott Woods , Gloria M. Reeves , Jason E. Schiffman

PII: S0165-1781(18)31178-8
DOI: https://doi.org/10.1016/j.psychres.2019.02.020
Reference: PSY 12235

To appear in: *Psychiatry Research*

Received date: 29 June 2018
Revised date: 8 February 2019
Accepted date: 8 February 2019

Please cite this article as: Pamela J. Rakhshan Rouhakhtar , Steven C. Pitts , Zachary B. Millman , Nicole D. Andorko , Samantha Redman , Camille Wilson , Caroline Demro , Peter L. Phalen , Barbara Walsh , Scott Woods , Gloria M. Reeves , Jason E. Schiffman , The Impact of Age on the Validity of Psychosis-Risk Screening in a Sample of Help-Seeking Youth, *Psychiatry Research* (2019), doi: https://doi.org/10.1016/j.psychres.2019.02.020

Highlights

- Younger participants endorsed more risk items on average on the Prime Screen
- Self-report was more predictive of diagnosed risk among older participants
- Optimal threshold for items endorsed to predict CHR/EP status decreased with age

The Impact of Age on the Validity of Psychosis-Risk Screening in a Sample of Help-Seeking Youth

Pamela J. Rakhshan Rouhakhtar [a]

Steven C. Pitts [a]

Zachary B. Millman [a]

Nicole D. Andorko [a]

Samantha Redman [a]

Camille Wilson [a1]

Caroline Demro [a2]

Peter L. Phalen [b]

Barbara Walsh [c]

Scott Woods [c]

Gloria M. Reeves [b]

Jason E. Schiffman [a]

[a] Human Services Psychology Department, University of Maryland Baltimore County, 1000 Hilltop Circle, Baltimore, MD, 21250, USA

[b] Department of Psychiatry, University of Maryland, School of Medicine, 701 West Pratt St., Baltimore, MD, 21201, USA

[c] Department of Psychiatry, Yale University, 300 George St., New Haven, CT 06511, USA

[1] Present affiliation: Nationwide Children's Hospital

[2] Present affiliation: Department of Psychiatry, Harvard Medical School

Corresponding author: Jason Schiffman, schiffma@umbc.edu, phone: 410-455-1574, fax: 410- 455-8991. Human Services Psychology Department, University of Maryland Baltimore County, 1000 Hilltop Circle, Baltimore, MD, 21250, USA

Abstract

Self-report screening instruments offer promise in furthering early identification of at-risk youth, yet current efforts are limited by false positive rates. Identifying moderators of accuracy is a potential step towards improving identification and prevention efforts. We investigated the moderating effect of age on self-reported attenuated positive symptoms from the Prime Screen and clinician diagnosed clinical high-risk/early psychosis (CHR/EP) status. Participants (N = 134) were racially diverse, lower-income, help-seeking adolescents and young adults from a primarily urban community. The overall model predicting CHR/EP status was significant, with results suggesting the presence of a trending interaction between age and Prime Screen symptoms. Analyses indicated that number of items endorsed to predict CHR/EP decreased with age (youngest group [M = 12.99] cut off = 6 items; middle age group [M = 14.97] cut off = 3; oldest age group [M = 18.40] cut off = 1). Although younger participants endorsed more risk items on average, follow up analyses suggested that the Prime Screen was a more accurate predictor of clinician-diagnosed-risk among older participants relative to their younger peers. The current study builds on the literature identifying moderators of psychosis-risk screening measure accuracy, highlighting potential limitations of CHR/EP screening tools in younger populations.

1. Introduction

Designed to complement clinician administered interviews, brief self-report screening instruments offer promise in furthering early identification efforts as a more feasible first step towards identifying people at risk for, or in early phases of, psychosis (Kline and Schiffman, 2014). Several self-report screening tools for psychosis-risk syndromes have been developed, allowing for quick evaluation of large samples not possible with clinical interviews (Daneault and Stip, 2013). Screening to identify appropriate candidates for clinical interview appears effective, as screening tools have shown clinically meaningful agreement with clinician identified risk, as well as the ability to predict later onset of psychosis in these high-risk samples (Addington et al., 2015a; Kline et al., 2012; Kline and Schiffman, 2014; Rietdijk et al., 2012; van der Gaag et al., 2012). For a more detailed review of psychosis risk screening tools and their relative psychometric properties, see Addington, Stowkowy and Weiser (2015a) and Kline and Schiffman (2014).

Among other measures, the Prime Screen has emerged as a viable screening instrument for emerging psychosis (Miller et al., 2004). Developed by the authors of the Structured Interview for Psychosis Risk Syndromes (SIPS; one of the gold-standard interviews for assessment of psychosis risk; McGlashan et al., 2010), the Prime Screen is a 12-item scale that assesses levels of attenuated psychosis symptoms such as unusual thought content, suspiciousness, grandiosity, and perceptual abnormalities. The Prime Screen has been validated in multinational samples (Kline et al., 2012; Kobayashi et al., 2008; Miller et al., 2004) and has favorable or comparable predictive values of psychosis-risk status compared to similar measures, such as the Prodromal Questionnaire Brief (Addington et al., 2015a; Kline et al., 2012). Additionally, as the compendium instrument to the SIPS interview, it is often recommended (although not exclusively) by SIPS certified trainers that Prime Screen administration precede SIPS administration. The brevity of the Prime Screen and its mapping to the SIPS interview items represent pragmatic and conceptual utility consistent with other screening measures (e.g., PQ-B), in that it may be employed in busy clinical settings

with little disruption and with confidence that it is measuring constructs similar to those probed by the SIPS.

As the field attempts to enhance screening accuracy, increased attention is needed to understand individual client characteristics that may moderate screening tool effectiveness. One possibly relevant factor is age. Evidence suggests elevated reports of psychotic-like experiences, as well as elevated psychosis risk screening results, among children compared to adolescents and young adults (e.g., Bartels-Velthius et al., 2011; Brandizzi et al., 2014; Kelleher et al., 2012a; Laurens et al., 2007; Rubio et al. 2012; Scott et al., 2008). Despite self-reporting more psychotic-like symptoms, younger individuals are not diagnosed with psychological disorders at greater rates compared to older participants, suggesting the possibility of over-endorsement, or endorsement of developmentally normative, non-pathological experiences among younger groups (Carol and Mittal, 2015; Kelleher et al., 2012b; Millman and Schiffman, 2018).

Population based, longitudinal examination of self-reported psychotic-like experiences indicates that rates of endorsements decline as individuals age, suggesting the possibility that such subthreshold psychotic experiences may be a part of typical development for younger people (Bourque et al., 2017; Mackie et al., 2013; van Os et al., 2009). Nonetheless, due to the nature of the self-report measures, it is unclear whether increased levels of attenuated psychosis symptoms at younger ages are due to a normative developmental phenomenon or a differential responding style.

With respect to the latter phenomenon, there is some evidence that younger children tend to report on extreme ends of self-report scales in an "all-or-nothing" style. For instance, in a general population sample of children ages 6-13 years, Mellor and Moore (2014) found that younger participants endorsed items at the extreme end of a scale in a manner similar to a yes/no format, rather than endorsing middle range items. This pattern was observed for participants answering Likert-type questions measuring various abstract constructs (e.g. "I feel very lonely," "I sleep very well," "I am able to do things as well as

most other kids."). Additional work has found similar response patterns for young participants (Chambers and Johnston, 2002; Kramer, 2009), suggesting that children tend to use Likert-type scales dichotomously as compared to older peers, particularly for abstract concepts such as mental state. Other work has demonstrated patterns of worse psychometric properties or extreme response styles on self-report measures for younger participants (Davis et al., 2007; Ellis et al., 2011; Gilles et al., 2008; Kramer, Kielhofner and Smith, 2010; Liddle and Carter, 2015; Tomyn et al., 2017; Wei, Stevens and Lan, 2018). These findings are supported by theory indicating that children at earlier developmental stages are more likely to engage in dichotomous thinking (Gelman and Baillargeon, 1983; Rosen, 1985).

Despite the existing literature, the impact of age on the validity of psychosis risk screening tools remains unclear. Although some studies have indicated that psychosis-risk screening measures like the Prodromal Questionnaire Brief show comparable psychometric properties across age groups (e.g., Pelizza et al., 2018b), other work has pointed to the potential effect of age on psychosis-risk measurement (e.g., Brandizzi et al., 2014; Kelleher et al. 2012a; Pelizza et al., 2018a), with authors indicating the need for future work to evaluate and consider the benefit of adding age-tailored scoring norms for psychosis-risk screening measures (Pelizza et al., 2018c). The current study aimed to investigate the potential moderating effect of age on the association between self-reported attenuated positive symptoms from the Prime Screen and clinician diagnosed risk/early psychosis status from the SIPS. Given prior research on the increased frequency of potentially non-pathological endorsement of psychotic-like experiences among younger age groups, we hypothesized that Prime scores would be less accurate in terms of predicting clinical interview identified risk/early psychosis for younger individuals relative to older participants (i.e., the concurrent validity of the Prime Screen would increase with age).

2. Method

2.1 Participants

This study was conducted through the Youth FIRST research program/Strive for Wellness clinic affiliated with the University of Maryland, Baltimore County (UMBC) and the University of Maryland School of Medicine. All procedures received Institutional Review Board approval at the University of Maryland, School of Medicine and University of Maryland, Baltimore County. Participants ($N = 145$) were help-seeking adolescents and young adults between the ages of 12-25, recruited from within the community through providers, clinics, hospitals, and schools in central Maryland and surrounding areas. To be eligible for the study, participants must have been currently receiving services for any mental health concerns. Although some participants were referred to the study due to concerns about psychosis symptoms, others were referred due to diagnostic conditions unrelated to psychotic spectrum disorders. All participants provided written consent to participate in the study; if participants were under the age of 18 years old, parents provided written consent, with the adolescent providing written assent.

2.2 Measures

All clinical interviews were administered by a graduate student in clinical psychology, or a master's level or higher clinician, and all clinical formulations were presented to a team of clinicians with extensive training in the area (two doctoral-level clinicians including one who is one of only three certified SIPS trainers in the United States (JS), a psychiatric nurse practitioner, and several psychology doctoral students, all of whom received certification in SIPS administration by completing a two-day training on the SIPS by the assessment author) to ensure clinical consensus. Participants completed the Prime Screen in person using pencil/paper; following completion of self-report measures, the SIPS interview was administered by a study clinician. Clinicians did not view the Prime Screen prior to interview or interview scoring.

2.2.1 Demographics Measures of binary gender (female, male), birth date, total family/household income, and race were self-reported by participants or, in the case of younger participants, by their guardian. Age in years was calculated as a continuous measure based on the difference between interview date and birthdate.

2.2.2 Structured interview for psychosis-risk syndromes (SIPS) The SIPS is a clinician-administered interview that assesses symptoms associated with the risk state for psychosis (McGlashan et al., 2010; Miller et al., 2003). Three risk syndromes (attenuated psychosis syndrome, brief intermittent psychosis syndrome, and genetic risk syndrome), as well full threshold psychosis, are determined from the SIPS interview based on scores of 3-5 (risk) or 6 (psychotic) for the positive symptom dimension. Consistent with the clinical applicability of identifying people who were heretofore not identified as at-risk or having crossed a diagnosable threshold for psychosis, participants positively diagnosed with a risk disorder or full-threshold psychosis were designated at "clinical high-risk/Early Psychosis" (CHR/EP) and considered as a single group. Given findings that risk of progression to psychosis among adolescents with SPD is similar to the magnitude of risk among peers meeting other SIPS criteria (e.g., 21%; Woods et al., 2009), the largest multisite consortium of CHR researchers in North America has added SPD-only as a risk syndrome if the participant is younger than age 19 (Addington et al., 2015b). Accordingly, and given that youth with SPD are clinically more similar to youth with other CHR syndromes than help seeking controls, we included individuals with SPD only (n = 3; ages = 15.45, 15,46, and 18.29) in the CHR group. In the present sample, n = 27 participants were diagnosed with attenuated psychosis syndrome, n = 3 participants were diagnosed with brief intermittent psychosis syndrome, n = 1 participants were diagnosed with genetic risk syndrome, n = 3 participants were diagnosed with schizotypal personality disorder in the absence of functional decline, n = 18 were diagnosed with more than one CHR diagnosis, and n = 19 were diagnosed with full threshold psychosis.

SIPS administrators were trained either by attending a two-day training with the creators of the SIPS,

interviews administered by staff, being observed by trained staff performing interviews, and achieving a high level of reliability over at least two cases before being cleared for SIPS interviewing (ICC > 0.80). Additionally, participants' clinical presentations were reviewed during team meetings to provide supervision and agreement in ratings. Within the SIPS raters, symptom reliability for the SIPS was ICC = 0.82, and diagnostic agreement was perfect (kappa = 1.0)

2.2.3 Prime screen The Prime Screen is a self-report measure containing 12 Likert-type items describing attenuated symptoms rated on a 7-point scale, ranging from "definitely disagree" to "definitely agree" (Miller et al., 2003). Individuals are categorized as "at risk" by the Prime Screen if they endorse at least two items with a score of 5 ("somewhat agree") or one item with a score of 6 ("definitely agree"). At least two studies have reported good sensitivity, specificity and predictive validity of the Prime Screen relative to SIPS diagnoses (Kline et al., 2015; Miller et al., 2004).

2.2.3.1 Prime screen scoring The Prime Screen was operationalized in two ways for the current study. In addition to the dichotomous measure of "at risk" for early psychosis status as defined by the instrument authors ("Prime Screen diagnosis"), a simple count of the number of items endorsed with a 6 ("definitely agree") was created, which we refer to as "Prime Screen symptoms." A Prime Screen symptom count was calculated in this manner to derive a continuous measure of items endorsed at the extreme positive end of the anchor. This scoring procedure allowed investigators to evaluate the potential effect of dichotomous response patterns across age.

2.3 Statistical Analyses

Analyses aimed to determine the impact of age on the relation between Prime Screen symptoms and SIPS-obtained diagnoses. Data were examined for normality and outliers, and the presence of confounds of age were examined (other DSM diagnoses, gender, income, race, and CHR/EP status). A point-biserial correlation between Prime Screen symptoms and CHR/EP status was examined to determine the

moderated logistic regression was estimated, with Prime Screen symptoms (number of extreme endorsements), age, and their interaction term predicting CHR/EP status. In a second moderated logistic regression, age^2 and its interaction term with Prime Screen symptoms was added to the model given the quadratic nature of the effect. A final moderated logistic regression was estimated with traditional Prime Screen risk symptom count (number of items receiving a score of 5 *or* 6) replacing Prime Screen symptoms (count of items endorsed at a 6) to evaluate the unique effect of extreme response patterns. Finally, the sample was trichotomized into three groups (younger, middle, older); screening tool accuracy within each age group was further examined in a receiver operating characteristic (ROC) curve with dichotomized CHR/EP classification as the 'state' variable. Area under the curve (AUC) and overall classification accuracy of each screening instrument in the ROC plot was examined. The Youden method was used to generate empirically derived "optimal" cut-off scores for each age group within the sample.

3. Results

Participants were excluded from analysis if they were missing data for age or risk/psychosis status (n = 1 excluded). For the Prime Screen, participants who answered fewer than 10 items were excluded from the analysis (n = 10 excluded), yielding an analysis sample of n = 134. Prior to analysis, all continuous measures were examined for normality and outliers with no indication of pronounced non-normality (skewness and kurtosis levels < 2 and < 10, respectively; Curran et al., 1996). Tests of the relation of age with gender (t[132] = 0.77, p = 0.44), race (F[2, 125] = 0.76, p = 0.47), and income level (r_S[N = 119] = -0.03, p = 0.77) suggested no apparent relations of these variables with age. Evaluation of differences in age between diagnostic groups (from Kiddie Schedule for Affective Disorders and Schizophrenia interview; Kaufman et al., 1997) indicated significant differences in age for those diagnosed with externalizing disorders (t[129] = 4.22, p <0.001, d = 0.75) . No significant group differences were observed in age for anxiety, mood, substance use, or other disorders (all ps ≥ 0.05). Due to power concerns in the present study, externalizing diagnosis was not controlled for in analyses. Table 1 presents

descriptive statistics and frequencies for the analysis sample. Given the limited power to detect effects, as well as the theoretical grounding of questions, no corrections were made for multiple comparisons.

The point-biserial correlation suggested a relation between Prime Screen symptoms and CHR/EP status ($r_{pb} = 0.39$, $p < 0.001$); those with positive SIPS diagnoses reported more Prime Screen symptoms. To test the primary hypotheses that this relation was moderated by age, CHR/EP status was regressed on age, Prime Screen symptoms (count of items with a score of 6, "definitely agree"), and their crossproduct in a logistic regression. Though the overall model predicting SIPS diagnostic group was significant, (χ^2 [3, $N = 134$] = 23.41, $p < 0.001$), the interaction term did not achieve statistical significance, $b = 0.03$, χ^2 (1) = 0.21, $p = 0.65$.

Visual inspection of the residual by first-order predictor plots suggested the presence of nonlinearity due to age. Accordingly, a second logistic regression was estimated in which CHR/EP status was regressed on Prime Screen symptoms, age, quadratic age (age squared), the crossproducts of Prime Screen symptoms by age, as well as quadratic age (prior to creation of the crossproduct terms, all continuous measures were mean-centered for easier interpretability). The overall model predicting CHR/EP status was statistically significant (χ^2 [5, $N = 134$] = 28.59, $p < 0.001$). Importantly, results suggested the presence of a linear-by-quadratic interaction between age and Prime Screen symptoms (see Table 2). Though the linear-by-quadratic interaction between age and Prime Screen symptoms did not achieve statistical significance at traditional levels ($p = 0.07$), we elected to probe this interaction. Research suggests that higher order terms (e.g., interactions) are typically low in power and have small effect sizes (Aiken and West, 1991). To explore the interaction, the simple effects of Prime Screen symptoms on CHR/EP status were estimated at three ages. Age was mean-centered at 13, 15, and 18 years to reflect the effect of Prime Screen symptoms on CHR/EP status at younger, middle, and older age (for an in depth discussion of testing and interpreting interactions, see Aiken and West, 1991). Though Prime Screen symptoms demonstrated a relation (analogous to "predictive validity") across age, the relation

To determine whether the moderating effect of age was unique to scores of 6 on the Prime Screen, a moderated logistic regression was estimated in which we changed Prime Screen symptoms from a count of 6's to a count of 5's or 6's, consistent with screening tool author scoring conventions. Though the overall model was significant, χ^2 [5, N = 134] = 29.73, $p < 0.001$, the linear by non-linear interaction did not achieve statistical significance, b = -0.01, χ^2 (1) = 1.08, p = 0.30. This may suggest that over-endorsement of items by younger participants is more prevalent for the anchor (extreme) options.

To help understand the practical implications of the nonlinear effects of age, the sample was trichotomized into three age groups roughly corresponding to the ages used with simple effects. Groups were defined as "younger age" (N = 43, M = 12., 12.06-13.94), "middle age" (N = 45, M = 14.97, 14.14-15.85), and "older age" (N = 46, M = 18.40, 16.05-22.55). Table 4 provides the sensitivity, specificity, positive predictive validity, negative predictive validity, and accuracy of Prime Screen diagnosis (risk/no risk) with regard to CHR/EP status (as defined by SIPS). Current findings fall within ranges of previously reported psychometric values for the Prime Screen (Addington et al., 2015a). The data in the present study suggest younger participants over-endorsed items on the Prime Screen, as they have notably lower specificity (true negative) rates than other age groups.

Exploratory analyses were performed to determine optimal cutoffs for number of risk items (5 and 6) endorsed at three age groups. Using CHR/EP status (risk vs. no risk) and number of Prime items endorsed, ROC curves were calculated for each age group (younger, middle, older) to evaluate the screening tool's accuracy for age groups, and to produce empirically derived "optimal" cutoffs for the Prime Screen for each age group. The optimal cutoff for the Prime Screen decreased with age. For the younger age group (N = 43), the AUC was 0.67, p = 0.04, 95% CI = 0.51-0.83. Using the Youden method, the empirically derived optimal cutoff

point was determined to be 6. This yielded a sensitivity = 0.32, specificity = 0.95, PPV = 0.88, and NPV = 0.57. For the middle age group (N = 45), the AUC was 0.86, $p < 0.001$, 95% CI = 0.76 –0.96. The

empirically derived optimal cutoff point was determined to be 3 or more items, yielding sensitivity = 0.78, specificity = 0.94, PPV = 0.96, and NPV = 0.74. The oldest age group ($N = 46$) had an AUC of 0.68, $p = 0.02$, 95% CI = 0.53-0.83. The empirically derived optimal cutoff point was determined to be 1 or more items, yielding sensitivity = 0.77, specificity = 0.54, PPV = 0.61, and NPV = 0.72.

4. Discussion

Psychosis-risk screening measures are currently limited by high rates of false-positives, leading to a variety of possible clinical and ethical concerns (Kline et al., 2012; Millman and Schiffman, 2018). Identifying factors that account for differential effectiveness in screening tools is essential to reducing false-positives and increasing accuracy. Previous work has established that overall screening accuracy depends on factors such as sample (help-seeking individuals vs. general population; Kline and Schiffman, 2014), race (Millman et al., submitted), and the input of additional informants (Kline et al., 2013). The current study builds on the existing literature by suggesting that respondent age might moderate screening accuracy. In the current sample, accuracy of Prime Screen predicting CHR/EP diagnoses varied as a function of participant age. Although on average younger participants endorsed more risk items at a level of 6 ("definitely agree"), upon clinical interview, younger participants who screened high were given a CHR/EP diagnosis at lower rates compared to their middle and older aged peers, suggesting a mismatch between self-report endorsement and clinician determined risk/early psychosis in this group.

The impact of age on the validity of the Prime Screen was notably stronger when risk items were defined as those endorsed at the extreme end of the scale (e.g., 6). This finding is consistent with previous studies of self-report response patterns of younger individuals that note tendencies of younger individuals to use Likert-type scales in a yes/no fashion (Chambers and Johnston, 2002; Mellor and Moore, 2014). Developmental theory suggests that younger children are more apt to employ dichotomous thinking, and are less likely to engage in dimensional conceptualizations of abstract concepts such as mental or emotional states (Gelman and Baillargeon, 1983; Rosen, 1985). More frequent endorsement by younger

participants of extreme items in the present study may in part reflect an increased reliance on dichotomous thinking in these participants rather than implying psychosis liability.

Although previous work demonstrates that children and younger adolescents endorse greater rates of attenuated symptoms as compared to older individuals (Bartels-Velthius et al., 2011; Brandizzi et al., 2014; Kelleher et al., 2012a; Laurens et al., 2007; Rubio et al. 2012), self-reported or clinician rated symptoms of attenuated psychosis appear to predict truer risk for psychopathology as age increases (Kelleher et al., 2012b; Schimmelmann et al., 2015). At the same time, endorsement of psychotic-like experiences decreases with age for adolescents (Bourque et al., 2017; Mackie et al., 2013). Thus, in relation to older youth, younger participants may be more likely to endorse experiences on psychosis-risk screening tools, but their positive endorsement will be less likely related to the presence of true early psychosis. Younger children's experience and reporting of these symptoms may reflect normative developmental processes (Carol and Mittal, 2015).

4.1 Limitations

Potential limitations warrant discussion. First, the cross-sectional design limits our ability to evaluate the extent to which age impacts screening capability to identify individuals at greater risk for developing psychosis over time. The cross-sectional design of this study also limits our ability to determine whether age influenced the validity of the SIPS itself in terms of predicting psychosis. The extent to which younger age is related to later psychosis among those at risk remains unclear (Schimmelmann and Schultz-Lutter, 2012), which has implications for age-related strengths and limitations of screening tools. Longitudinal studies examining the influence of age on both pretest risk and later psychosis would be a valuable contribution to the field of early detection in psychosis. Additionally, the relatively small sample size and limited power to detect interactions suggests our findings should be interpreted with caution. Due to these same concerns, we did not control for diagnoses in analyses. Participants across diagnostic groups did not differ by age for anxiety, mood, substance use, or other disorders, alleviating potential

concern that differences in these diagnoses would differentially influence findings. There were age differences among those diagnosed with an externalizing disorder compared to non-externalizing peers, but exploratory analyses indicated that the pattern of results for the primary analysis held when presence of an externalizing disorder was included in the model. Similarly, sampling method and characteristics may have yielded results idiosyncratic to the current study sample. Given the known impact of recruitment strategies on risk enrichment and risk samples (Fusar-Poli et al., 2016), results should be interpreted with caution and with the present context in mind before they are replicated in a number of diverse samples.

Additionally, while the current findings are in line with previous work indicating an impact of age on psychometric validity of other psychosis risk screening measures (Brandizzi et al., 2014; Kelleher et al. 2012a; Pelizza et al., 2018a; Pelizza et al., 2018c), other findings indicating a non-effect of age on measurement validity (Pelizza et al., 2018b) indicate the need for more work clarifying the cause of discrepant findings.

Inclusion/exclusion criteria prohibited participants older than 25 years of age from inclusion in the current study, reducing the upper limit of age as compared to some other longitudinal studies of early psychosis and psychosis-risk (Addington et al., 2012). The ability of the current study to detect a trending effect of age given the limited range in the sample, however, indicates the potential strength of this effect. The use of the SIPS clinical interview as the "gold-standard" measure for psychosis-risk and early psychosis also deserves comment, given the relatively low rate of transition to psychosis shown by individuals deemed at-risk (Fusar-Poli et al., 2012). Despite the fact that most individuals at CHR do not develop a psychotic disorder, the distress and impairment experienced by CHR non-converters renders screening efforts to identify psychosis-risk worthwhile (Addington et al., 2011; Lin et al., 2015; Millman and Schiffman, 2018). Current limitations suggest need for replication in larger and diverse samples to determine the presence and strength of the impact of age on emerging psychosis screening tool validity.

4.2 Clinical Implications and Future Directions

Small sample and the need for replication notwithstanding, the current findings suggest possible benefits of age adjusted scoring procedures and/or cutoff recommendations. Current cutoffs for Prime Screen risk status include two or more item endorsements at a score of 5 or one or more 6. Exploratory analyses within our small sample suggest that optimal cutoffs varied by age, with a higher number of items required for younger participants (cutoff = 6 items), as compared to middle age (cutoff = 3 items) and older age participants (cutoff = 1 item).

The current study builds on existing literature examining psychometric properties of psychosis-risk screening tools among help-seeking children and adolescents (Addington et al., 2015a; Kline et al., 2016; Kline and Schiffman, 2014; Schiffman, 2018). Age of respondent appeared to impact validity of the screening measure, with younger participants demonstrating differential response patterns as compared to older participants on the Prime Screen. These findings, as well as previous literature demonstrating age differences in early psychosis symptoms, highlight the need for more work understanding the processes that underlie symptom endorsement in younger individuals across psychosis risk screening tools. Identification of such obfuscating factors may aid in improving the clinical utility of psychosis-risk screening tools. In reducing false positive rates and improving accuracy, self-report measures will have broadened ability to identify at-risk youth, while not inadvertently pathologizing experiences common in typical development.

Acknowledgments

The UMBC YouthFIRST research team was instrumental in supporting the work that made this research possible. Additionally, research in our field entirely depends on participants and their families, to whom a debt of gratitude is owed.

Conflict of Interest

The authors have no actual or potential conflicts of interest to disclose.

Role of Funding

This work was supported by the National Institute of Mental Health (grants R01MH112612 and R34MH110506) and the Maryland Department of Health and Mental Hygiene, Behavioral Health Administration through the Center for Excellence on Early Intervention for Serious Mental Illness (OPASS# 14-13717G/M00B4400241). The funding sources did not have a role in the writing of this manuscript.

References

Addington, J., Cadenhead, K.S., Cornblatt, B.A., Mathalon, D.H., McGlashan, T.H., Perkins, D.O., et al., 2012. North American Prodrome Longitudinal Study (NAPLS 2): Overview and recruitment. Schizophr Res 142 (1-3), 77-82.

Addington, J., Cornblatt, B.A., Cadenhead, K.S., Cannon, T.D., McGlashan, T.H., Perkins, D.O., et al., 2011. At clinical high risk for psychosis: outcome for nonconverters. Am J Psychiatry 168 (8), 800-805.

Addington, J., Liu, L., Buchy, L., Cadenhead, K. S., Cannon, T. D., Cornblatt, B. A., et al., 2015b. North American prodrome longitudinal study (NAPLS 2): the prodromal symptoms. J. Nerv. Ment. Dis., 203 (5), 328.

Addington, J., Stowkowy, J., Weiser, M., 2015a. Screening tools for clinical high risk for psychosis. Early Interv Psychiatry 9 (5), 345-356.

Aiken, L.S., West, S.G., 1991. Multiple Regression: Testing and Interpreting Interactions. Sage Publications, Newbury Park.

Bartels-Velthuis, A.A., van de Willige, G., Jenner, J.A., van Os, J., Wiersma, D., 2011. Course of auditory vocal hallucinations in childhood: 5-year follow-up study. Br J Psychiatry 199 (4), 296-302.

Bourque, J., Afzali, M.H., O'Leary-Barrett, M., Conrod, P., 2017. Cannabis use and psychotic-like experiences trajectories during early adolescence: the coevolution and potential mediators. J Child Psychol Psychiat 58 (12), 1360-1369.

Brandizzi, M., Schultze-Lutter, F., Masillo, A., Lanna, A., Curto, M., Lindau, J.F., et al., 2014. Self-reported attenuated psychotic-like experiences in help-seeking adolescents and their association with age, functioning and psychopathology. Schizophr Res 160 (1-3), 110-117.

Carol, E.E., Mittal, V.A., 2015. Normative adolescent experiences may confound assessment of positive symptoms in youth at ultra-high risk for psychosis. Schizophr Res 166 (1-3), 358– 359.

Chambers, C.T., Johnston, C., 2002. Developmental differences in children's use of rating scales. J Pediatr Psychol 27 (1), 27-36.

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.

Daneault, J.G., Stip, E., 2013. Genealogy of instruments for prodrome evaluation of psychosis. Front Psychiatry 4, 25.

Davis, E., Nicolas, C., Waters, E., Cook, K., Gibbs, L., Gosch, A., et al., 2007. Parent-proxy and child self-reported health-related quality of life: using qualitative methods to explain the discordance. Qual Life Res 16 (5), 863-871.

Ellis, J. A., Ootoova, A., Blouin, R., Rowley, B., Taylor, M., DeCourtney, C., et al., 2011. Establishing the psychometric properties and preferences for the Northern Pain Scale. Int J Circumpolar Health, 70 (3), 274-285.

Fusar-Poli, P., Bonoldi, I., Yung, A.R., Borgwardt, S., Kempton, M.J., Valmaggia, L., et al., 2012. Predicting psychosis: meta-analysis of transition outcomes in individuals at high clinical risk. Arch Gen Psychiatry 69 (3), 220-229.

Fusar-Poli, P., Schultze-Lutter, F., Cappucciati, M., Rutigliano, G., Bonoldi, I., Stahl, D., et al. 2015. The dark side of the moon: meta-analytical impact of recruitment strategies on risk enrichment in the clinical high risk state for psychosis. Schizophr Bull 42 (3), 732-743.

Gelman, R., Baillargeon, R., 1983. A Review of Some Piagetian Concepts. Wiley Press, New York.

Gilles, C. D., Arnould, C., Thonnard, J. L., Lejeune, T. M., 2008. ABILOCO-Kids: a Rasch-built 10-item questionnaire for assessing locomotion ability in children with cerebral palsy. J Rehabil Med 40 (10), 823-830.

Kaufman, J., Birmaher, B., Brent, D., Rao, U. M. A., Flynn, C., Moreci, P., et al. 1997. Schedule for affective disorders and schizophrenia for school-age children-present and lifetime version (K-SADS-PL): initial reliability and validity data. J Am Acad Child Adolec Psychiatry 36 (7), 980-988.

Kelleher, I., Connor, D., Clarke, M.C., Devlin, N., Harley, M., Cannon, M. 2012a. Prevalence of psychotic symptoms in childhood and adolescence: a systematic review and meta-analysis of population-based studies. Psychol Med 42 (9), 1857-1863.

Kelleher, I., Keeley, H., Corcoran, P., Lynch, F., Fitzpatrick, C., Devlin, et al., 2012b. Clinicopathological significance of psychotic experiences in non-psychotic young people: evidence from four population-based studies. Br J Psychiatry 201(1), 26-32.

Kline, E., Thompson, E., Schimunek, C., Reeves, G., Bussell, K., Pitts, S.C., et al., 2013. Parent–adolescent agreement on psychosis risk symptoms. Schizophr Res 147 (1), 147-152.

Kline, E., Schiffman, J., 2014. Psychosis risk screening: a systematic review. Schizophr Res 158 (1-3), 11-18.

Kline, E., Thompson, E., Demro, C., Bussell, K., Reeves, G., Schiffman, J., 2015. Longitudinal validation of psychosis risk screening tools. Schizophr Res 165 (2-3), 116-122.

Kline, E., Thompson, E., Demro, C., Bussell, K., Reeves, G., Schiffman, J., 2016. Self-report instruments for clinical monitoring of psychosis risk states. Psychiatr Serv 67 (4), 456-459.

Kline, E., Wilson, C., Ereshefsky, S., Denenny, D., Thompson, E., Pitts, S.C., et al., 2012. Psychosis risk screening in youth: a validation study of three self-report measures of attenuated psychosis symptoms. Schizophr Res 141 (1), 72-77.

Kobayashi, H., Nemoto, T., Koshikawa, H., Osono, Y., Yamazawa, R., Murakami, M., et al., 2008. A self-reported instrument for prodromal symptoms of psychosis: testing the clinical validity of the Prime screen-revised (PS-R) in a Japanese population. Schizophr Res 106 (2-3), 356-362.

Kramer, J., 2009. Presentation 2: How children with disabilities use Likert scales to report competence and importance of everyday activities. Arch Phys Med Rehabil 90 (10).

Kramer, J. M., Kielhofner, G., Smith, E. V., 2010. Validity evidence for the child occupational self assessment. Am J Occup Ther 64 (4), 621-632.

Laurens, K.R., Hodgins, S., Maughan, B., Murray, R.M., Rutter, M.L., Taylor, E.A., 2007. Community screening for psychotic-like experiences and other putative antecedents of schizophrenia in children aged 9-12 years. Schizophr Res 90 (1-3), 130-146.

Liddle, I., Carter, G. F., 2015. Emotional and psychological well-being in children: the development and validation of the Stirling Children's Well-being Scale. Educ Psychol Pract 31 (2), 174-185.

Lin, A., Wood, S.J., Nelson, B., Beavan, A., McGorry, P., Yung, A.R., 2015. Outcomes of nontransitioned cases in a sample at ultra-high risk for psychosis. Am J Psychiatry 172 (3), 249-258.

Mackie, C.J., O'Leary-Barrett, M., Al-Khudhairy, N., Castellanos-Ryan, N., Struve, M., Topper, L., et al., 2013. Adolescent bullying, cannabis use and emerging psychotic experiences: a longitudinal general population study. Psychol Med 43 (5), 1033-1044.

McGlashan, T.H., Walsh, B., Woods, S., 2010. The Psychosis-Risk Syndrome: Handbook for Diagnosis and Follow-Up. Oxford University Press, New York.

Mellor, D., Moore, K.A., 2014. The use of Likert scales with children. J Pediatr Psychol 39 (3), 369-379.

Miller, T.J., Cicchetti, D.V., Markovich, P.J., Woods, S., 2004. The SIPS Screen: a brief self-report screen to detect the schizophrenia prodrome. Schizophr Res 70, 78.

Miller, T.J., McGlashan, T.H., Rosen, J.L., Cadenhead, K., Cannon, T., Ventura, et al., 2003. Prodromal assessment with the structured interview for prodromal syndromes and the scale of prodromal symptoms: predictive validity, interrater reliability, and training to reliability. Schizophr Bull 29 (4), 703-715.

Millman, Z.B., Rakhshan, P.J., DeVylder, J.E., Edmondson Smith, M., Phalen, P.L., Woods, et al., (Submitted). Evidence for differential predictive performance of the prime screen between black and white help-seeking adolescents.

Millman, Z., Schiffman, J., 2018. False positives and clinical heterogeneity among youth at clinical high-risk for psychosis: clinical and ethical implications for assessment and treatment. J Ethics Mental Health 10.

Pelizza, L., Azzali, S., Paterlini, F., Garlassi, S., Scazza, I., Chiri, L. R., et al., 2018a. Screening for psychosis risk among help- seeking adolescents: Application of the Italian version of the 16- item prodromal questionnaire (iPQ- 16) in child and adolescent neuropsychiatry services. Early Interv Psychiatry. doi: 10.1111/eip.12554.

Pelizza, L., Azzali, S., Paterlini, F., Garlassi, S., Scazza, I., Simona, P., et al., 2018b. The Italian version of the 16-item prodromal questionnaire (iPQ-16): Field-test and psychometric features. Schizophr Res 199, 353-360.

Pelizza, L., Azzali, S., Paterlini, F., Scazza, I., Garlassi, S., Chiri, L. R., et al., 2018c. The Italian version of the brief 21-item Prodromal Questionnaire: Field test, psychometric properties and age-sensitive

Rietdijk, J., Klaassen, R., Ising, H., Dragt, S., Nieman, D.H., van de Kamp, J., et al., 2012. Detection of people at risk of developing a first psychosis: Comparison of two recruitment strategies. Acta Psychiatr Scand 126 (1), 21-30.

Rosen, H., 1985. Piagetian Dimensions of Clinical Relevance. Columbia University Press, New York.

Rubio, J.M., Sanjuan, J., Florez-Salamanca, L., Cuesta, M.J., 2012. Examining the course of hallucinatory experiences in children and adolescents: a systematic review. Schizophr Res 138 (2-3), 248-254.

Schiffman, J., 2018. Considerations for the development and implementation of brief screening tools in the identification of early psychosis. Schizophr Res. doi: 10.1016/j.schres.2018.03.002

Schimmelmann, B. G., Michel, C., Martz-Irngartinger, A., Linder, C., Schultze-Lutter, F., 2015. Age matters in the prevalence and clinical significance of ultra-high-risk for psychosis symptoms and criteria in the general population: Findings from the BEAR and BEARS-kid studies. World Psychiatry 14 (2), 189–197. http://doi.org/10.1002/wps.20216

Scott, J., Welham, J., Martin, G., Bor, W., Najman, J., O'Callaghan, M., et al., 2008. Demographic correlates of psychotic-like experiences in young Australian adults. Acta Psychiatr Scand 118 (3), 230-237.

Tomyn, A. J., Fuller-Tyszkiewicz, M. D., Cummins, R. A., Norrish, J. M., 2017. The validity of subjective wellbeing measurement for children: Evidence using the Personal Wellbeing Index—School Children. J. of Happiness Stud. 18 (6), 1859-1875.

van der Gaag, M., McGorry, P., Cuijpers, P., 2012. Re: Early detection and intervention evaluation for people at risk of psychosis: multisite randomised controlled trial. BMJ 344.

van Os, J., Linscott, R.J., Myin-Germeys, I., Delespaul, P., Krabbendam, L., 2009. A systematic review and meta-analysis of the psychosis continuum: evidence for a psychosis proneness-persistence-impairment model of psychotic disorder. Psychol Med 39 (2), 179-195.

Wei, T., Barnard-Brak, L., Stevens, T., Lan, W. Y., 2018. Item parameter drift of the Self-Description Questionnaire I. Eur J. Psychol Assess. doi: 10.1027/1015-5759/a000473

Woods, S. W., Addington, J., Cadenhead, K. S., Cannon, T. D., Cornblatt, B. A., Heinssen, R., et al., 2009. Validity of the prodromal risk syndrome for first psychosis: Findings from the North American Prodrome Longitudinal Study. Schizophr Bull 35 (5), 894-908.

Table 1.

Descriptive Statistics and Frequencies for Analysis Sample by Age Group

	Younger Age 12 – 13.99 years $N = 43$		Middle Age 14 – 15.99 years $N = 45$		Older Age 16 – 23 years $N = 46$	
Continuous measures	*M*	*SD*	*M*	*SD*	*M*	*SD*
Age	12.99	0.50	14.97	0.52	18.40	1.84
Prime Screen Risk Level Symptoms	3.02	2.81	2.82	2.99	1.98	2.23
Categorical measures	*Freq*	*%*	*Freq*	*%*	*Freq*	*%*
Prime Screen diagnosis (at-risk)	28	65%	25	56%	25	54%
CHR/EP Status (at-risk/early psychosis)	22	51%	28	60%	22	48%
Gender (female)	22	51%	31	69%	29	63%
Race[a]						
Asian	0	0%	0	0%	1	2%
Black/African American	28	65%	20	44%	22	48%
White/Caucasian	8	19%	14	31%	17	37%
Biracial	5	12%	10	22%	4	9%
Not reported	2	5%	1	2%	2	4%
Income[b]						
<$20,000	11	26%	11	24%	14	30%
$20,000 - $39,999	13	30%	9	20%	5	11%
$40,000 - $59,999	2	5%	5	11%	4	9%
$60,000 - $79,999	5	12%	4	9%	6	13%

$80,000 - $99,999	4	9%	4	9%	3	7%
>$100,000	5	12%	6	13%	8	17%
Not reported	3	7%	6	13%	6	13%
Diagnoses						
Anxiety Disorder	22	51%	30	67%	33	72%
Eating Disorder	0	0%	1	2%	3	7%
Externalizing Disorder	38	88%	23	51%	21	46%
Mood Disorder	23	54%	22	49%	28	61%
Other Disorder	3	7%	2	4%	3	7%
Substance Disorder	0	0%	2	4%	5	11%

Note.

[a] Neither Asian nor "Not reported" participants included in inferential statistics

[b] Income was recorded as total family/household income, and was treated as ordinal variable in inferential statistics

Table 2.

Moderated Binary Logistic Regression Predicting SIPS Risk Diagnosis.

Predictors	β (SE)	$\chi 2$	p	OR
Age	1.73 (0.93)	3.62	0.06	5.63
Age^2	-0.05 (0.03)	3.57	0.06	0.95
Prime Symptoms	0.77 (0.20)	21.81	<0.001	2.16
Age • Prime Symptoms	0.14 (0.08)	2.97	0.09	1.15
Age^2 • Prime Symptoms	-0.03 (0.02)	3.29	0.07	0.97

Note.

Age and Prime symptoms centered prior to creation of higher order terms and inclusion in the model. Reported χ^2 are Likelihood Ratio tests, not Wald χ^2.

Table 3.

Simple Effects of Prime Symptoms on CHR/EP Status at Levels of Age

Predictors	*b* (SE)	$\chi 2$	*p*	OR
Younger Age (at 13 years)	0.35 (0.16)	5.81	0.02	1.42
Middle Age (at 15 years)	0.77 (0.20)	21.81	<0.001	2.16
Older Age (at 18 years)	0.89 (0.28)	14.18	<0.001	2.44

Note.

Reported χ^2 are Likelihood Ratio tests, not Wald-χ^2.

Table 4.

Sensitivity, specificity, PPV, NPV, and accuracy of Prime Screen diagnosis for CHR/EP status

	N	Age Cut-offs	Sensitivity	Specificity	PPV	NPV	Accuracy
Total Sample	134	12-23	0.75	0.60	0.68	0.68	0.68
Younger Age	43	12-13.99	0.77	0.47	0.61	0.67	0.63
Middle Age	45	14-15.99	0.78	0.78	0.84	0.70	0.78
Older Age	46	16-23	0.68	0.58	0.60	0.67	0.63

Note.

Outcome is CHR/EP status

The Youden method was used to calculate sensitivity, specificity, positive predictive validity, negative predictive validity, and accuracy values.