Schizophrenia Research xxx (2017) xxx–xxx

Contents lists available at ScienceDirect

Schizophrenia Research

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

The Early Psychosis Screener (EPS): Item development and qualitative validation

Brodey B.B. [a,*], Addington J. [b], First M.B. [c], Perkins D.O. [d], Woods S.W. [e], Walker E.F. [f], Walsh B. [e], Nieri J.M. [d], Nunn M.B. [g], Putz J. [h], Brodey I.S. [a,i]

[a] TeleSage, Inc., 201 East Rosemary St., Chapel Hill, NC 27514, USA
[b] Hotchkiss Brain Institute, Department of Psychiatry, University of Calgary, 3280 Hospital Drive NW, Calgary, Alberta T2N 4Z6, Canada
[c] Department of Psychiatry, Columbia University Medical Center, 1051 Riverside Drive, New York, NY 10032, USA
[d] Department of Psychiatry, School of Medicine, The University of North Carolina at Chapel Hill, 101 Manning Dr, Chapel Hill, NC 27514, USA
[e] PRIME Psychosis Prodrome Research Clinic, Connecticut Mental Health Center, B-38, 34 Park Street, New Haven, CT 06519, USA
[f] Departments of Psychology and Psychiatry, Emory University, 36 Eagle Row, Atlanta, GA 30322, USA
[g] Centerstone Tennessee, 1921 Ransom Place, Nashville, TN 37217, USA
[h] Centerstone Research Institute, 645 South Rogers Street, Bloomington, IN 47403, USA
[i] The University of North Carolina at Chapel Hill, 434 Greenlaw, Campus Box 3520, Chapel Hill, NC 27599, USA

ARTICLE INFO

Article history:
Received 27 July 2017
Received in revised form 6 November 2017
Accepted 24 November 2017
Available online xxxx

Keywords:
SIPS
PQ-B
CAARMS
NAPLS
Psychosis;
Prodromal
Schizophrenia
Screener

ABSTRACT

A panel of experts assembled and analyzed a comprehensive item bank from which a highly sensitive and specific early psychosis screener could be developed. Twenty well-established assessments relating to the prodromal stage, early psychosis, and psychosis were identified. Using DSM-5 criteria, we identified the core concepts represented by each of the items in each of the assessments. These granular core concepts were converted into a uniform set of 490 self-report items using a Likert scale and a 'past 30 days' time frame. Partial redundancy was allowed to assure adequate concept coverage. A panel of experts and TeleSage staff rated these items and eliminated 189 items, resulting in 301 items. The items were subjected to five rounds of cognitive interviewing with 16 individuals at clinically high risk for psychosis and 26 community mental health center patients. After each round, the expert panel iteratively reviewed, rated, revised, added, or deleted items to maximize clarity and centrality to the concept. As a result of the interviews, 36 items were revised, 52 items were added, and 205 items were deleted. By the last round of cognitive interviewing, all of the items were clearly understood by all participants. In future work, responses to the final set of 148 items and machine learning techniques will be used to quantitatively identify the subset of items that will best predict clinical high-risk status and conversion.

1. Introduction

Interest in identifying individuals at clinically high risk (CHR) of developing a psychotic spectrum disorder has grown over the past decade. The most widely used assessments are the Structured Interview for Psychosis-risk Syndromes (SIPS) and the Comprehensive Assessment of At-Risk Mental States (CAARMS) (Fusar-Poli et al., 2016). Both the SIPS and the CAARMS have very high sensitivity 91.6% (Webb et al., 2015; Fusar-Poli et al., 2016). Unfortunately, proper administration of these semi-structured interviews requires extensive training in order to assure high inter-rater reliability (Addington et al., 2012). Even with extensive SIPS or CAARMS training, only about 19.6% of individuals who are identified as CHR based on their SIPS score will actually go on to develop a psychotic disorder vs. 1.8% for help-seeking clinical controls. An additional 10.7% of CHR patients will develop bipolar disorder, unipolar depression, or an anxiety disorder vs. 11.8% of controls (Webb et al., 2015).

At present, the most widely used self-report screener for early psychosis is the Prodromal Questionnaire – Brief Version (PQ-B) (Loewy et al., 2005, 2011a). In general, the PQ instruments were extensively validated against the SIPS and the CAARMS (Loewy et al., 2011b; Ising et al.,

Abbreviations: CAARMS, Comprehensive Assessment of At-Risk Mental States; CHR, clinically high risk; CI, cognitive interviewing; IP, interviewer probe; NAPLS, North American Prodrome Longitudinal Study; PQ-B, Prodromal Questionnaire – Brief Version; PROMIS, Patient-Reported Outcomes Measurement Information System; SIPS, Structured Interview for Psychosis-risk Syndromes; TA, think aloud.

* Corresponding author.
E-mail addresses: bb@telesage.com (B.B. Brodey), jmadding@ucalgary.ca (J. Addington), mbf2@cumc.columbia.edu (M.B. First), diana_perkins@med.unc.edu (D.O. Perkins), scott.woods@yale.edu (S.W. Woods), psyefw@emory.edu (E.F. Walker), barbara.walsh@yale.edu (B. Walsh), jennifer_nieri@med.unc.edu (J.M. Nieri), brad.nunn@centerstone.org (M.B. Nunn), john.putz@centerstone.org (J. Putz), brodey@email.unc.edu (I.S. Brodey).

https://doi.org/10.1016/j.schres.2017.11.027
0920-9964/

2012). The PQ-B has high sensitivity, but as an outpatient screener it may lack sufficient specificity to make more widespread screening practical. There are a few additional difficulties with the PQ-B that this study aims to improve upon. First, although individual items are clearly written, they tend to focus on fairly mild symptoms that are common in the general population (e.g., seeing a fortune teller). Individual items also tend to combine several related but distinct experiences (e.g., "experiences with telepathy, psychic forces, or fortune telling") without any ability to distinguish between them. Finally, the primary response set for the PQ-B is a limited, binary 'yes/no'. A Likert scale was subsequently added, but its use is only indicated for items that are already endorsed with a 'yes'. The selective addition of a Likert scale also complicates statistical analyses.

Since the initial publication of the PQ-B in 2005, there have been several developments that can improve the creation of self-report screeners. These include cognitive interviewing (CI), which is useful in qualitative validation (DeWalt et al., 2007); Item Response Theory (IRT) and related aspects of Modern Measurement Theory (Reeve, 2002); machine learning strategies (Peng et al., 2005 & von Luxburg, 2007); and lessons learned from the National Institutes of Health Patient-Reported Outcomes Measurement Information System (PROMIS) initiative (Cella et al., 2007). While it is unusual for an item development manuscript not to conclude with a quantitative analysis defining the utility of the items, our premise has been that psychosis is one of the most difficult of human experiences to assess and that the quality of the items in an assessment naturally place an upper bound on predictive power, regardless of analytic strategy. (It is not possible to accurately and precisely interpret an item that is confusing, has multiple interpretations, or which includes several concepts e.g. in the case of depression: sad, depressed, or hopeless.) For these reasons, we have chosen to dedicate this manuscript to a detailed description of the application of these techniques to the development of a comprehensive self-report item bank that can be used to predict CHR status. The approach is primarily synthetic in nature, encompassing the theoretical frameworks for each of the assessments that form the basis of our item bank.

Our hypothesis is that we can develop a comprehensive set of simple Likert scale items that each represent a single, granular, core symptom associated with the prodromal period, including psychotic-like and psychotic experiences. Our belief is that this item bank will serve as the foundation for creating a self-report screener for early psychosis that could be used to predict SIPS CHR status and ultimately predict conversion with high specificity.

2. Methods

2.1. Stage I: item pool development

The first step was to gather widely used prodromal, early psychosis, and psychosis measures that have been described in the peer-reviewed literature. These measures are presented in a recent review of self-report and clinician-administered early psychosis screeners (Kline and Schiffman, 2014). Using these screeners and DSM-5 criteria, we identified the core concepts represented by each of the items in each of the assessments. These core concepts covered all of the criteria for schizophrenia spectrum and other psychotic disorders described in the DSM-5. Under the supervision of Dr. Brodey, who used similar techniques to develop the Perinatal Depression Inventory (Brodey et al., 2016), TeleSage staff rewrote items in a simplified self-report format. They based the items on a fifth-grade reading level, with one concept per item so that minimal interpretation of each item was required. Each item was intended to elicit a simple direct report of the individual's experiences and feelings. Wherever possible, items were written in a non-judgmental, non-pathologizing format. We avoided words and phrases with pejorative, multiple, or abstract connotations. (For an example, see the revised item 'I felt anxious.' in Results section, Table 1). Foreign words and words known to translate poorly into other languages were avoided.

Items were written to match a 5-point Likert scale (Never, Rarely, Sometimes, Often, Always) response set. This is the same response set that was used in the PROMIS initiative (DeWalt et al., 2007), except that we included a 'does not apply' response option as a second alternative to 'Never', for some items related to work and school experiences. This was for participants who were unemployed or not in school, therefore 'Never' could be ambiguous. We created items intended to represent different extremes of a symptom so as not to rely exclusively on the Likert scale for differentiation (Comparelli et al., 2014). Furthermore, we attempted to avoid items that might have a ceiling or floor effect. Items were written with a standard 'past thirty days' time frame. The panel of experts discussed using a scale of severity or distress instead of frequency; however, no single scale appeared to work perfectly to assess the prodromal period. Frequency appeared to act as an adequate proxy to capture intermittent prodromal episodes as well as attenuated symptoms. Although the PROMIS initiative used a 'past seven days' time frame, we reasoned that we needed a longer time frame in order to pick up the episodic symptoms that are associated with the prodromal period. Patients tend to answer consistently when asked about frequency or intensity, so the panel concluded that a 30-day time frame (typical for assessing prodromal symptoms) would be sufficient to capture the presence of intermittent episodes. In addition, we selected a uniform 'in the past 30 days' time frame rather than a 'past month' time frame to avoid confusion among people who might be thinking about the most recent named month (e.g., September) while answering questions.

In order to define and represent concepts associated with the prodromal period, early psychosis, and psychosis, we included concepts relating to positive symptoms, negative symptoms, and the exclusionary criteria listed in the DSM-5, as well as general symptoms that have been associated with conversion. We then subdivided the items into category 'bins' to assure adequate coverage of related concepts. Concepts included the symptoms listed in DSM-5, such as delusions, hallucinations, disorganized speech, gross disorganization, avolition, and a decrease in functioning; yet, for our purposes, the DSM-5 nomenclature was not sufficiently specific. The term 'delusion' alone, for example, can refer to any number of phenomena: paranoid delusions, persecutory delusions, religious delusions, grandiose delusions, delusions of control, thought insertion, telepathy, thought broadcasting, erotomania, and somatic delusions, to name a few. Drawing from the well-established instruments, we made our 'concepts' as granular as possible. We recognized that overlap in the nomenclature and categories was inevitable and that partial redundancy was, in fact, desirable. In addition, since the DSM-5 criteria for schizophrenia include exclusions relating to schizoaffective disorder, bipolar disorder, and substance abuse, we also included items on depression, anxiety, mania, and substance use for exploratory purposes.

The item pool was iteratively reviewed and rated by a panel of eight experts, including three psychiatrists and three psychologists with combined expertise in the prodromal period and early psychosis, SIPS and SCID (Structured Clinical Interview for the DSM) administration, community mental health, and biostatistics. Other members of the panel included an English professor, a linguist, and two TeleSage, Inc. interns. The panel members reviewed the items for breadth of coverage across the concepts. In addition, each item was rated independently by each panel member on a 3-point scale for clarity and centrality (i.e., 1 = neither clear, nor central to the concept; 2 = clear, but not central to concept OR central to concept, but not clear; and 3 = clear and central to concept). Experts participated in focus groups, where they were asked to describe the benefits and/or problems associated with each item and provide a rationale for each item rating. Experts were also asked to rank similarly worded items in order of their preference. We averaged the results from the expert panel ratings, and considered elaborations provided through the experts' comments and rankings.

Please cite this article as: Brodey, B.B., et al., The Early Psychosis Screener (EPS): Item development and qualitative validation, Schizophr. Res. (2017), https://doi.org/10.1016/j.schres.2017.11.027