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): Quantitative validation against the SIPS using machine learning

Brodey B.B. [a,*], Girgis R.R. [b], Favorov O.V. [c], Addington J. [d], Perkins D.O. [e], Bearden C.E. [f], Woods S.W. [g], Walker E.F. [h], Cornblatt B.A. [i], Brucato G. [b], Walsh B. [g], Elkin K.A. [a], Brodey I.S. [j]

[a] TeleSage, Inc., 201 East Rosemary St., Chapel Hill, NC 27514, USA
[b] New York State Psychiatric Institute, 1051 Riverside Drive, Unit 31, New York, NY 10032, USA
[c] Department of Biomedical Engineering, The University of North Carolina at Chapel Hill, 152 MacNider Hall, Campus Box 7575, Chapel Hill, NC 27599, USA
[d] Hotchkiss Brain Institute, Department of Psychiatry, University of Calgary, 3280 Hospital Drive NW, Calgary, Alberta T2N 4Z6, Canada
[e] Department of Psychiatry, School of Medicine, The University of North Carolina at Chapel Hill, 101 Manning Dr, Chapel Hill, NC 27514, USA
[f] Department of Psychiatry and Biobehavioral Sciences and Department of Psychology, University of California Los Angeles, 300 Medical Plaza, Rm 2265, Los Angeles, CA 90095
[g] PRIME Psychosis Prodrome Research Clinic, Connecticut Mental Health Center B-38, 34 Park Street, New Haven, CT 06519, USA
[h] Departments of Psychology and Psychiatry, Emory University, 36 Eagle Row, Atlanta, GA 30322, USA
[i] Department of Psychiatry Research, The Zucker Hillside Hospital, 75-59 263rd St., Glen Oaks, New York 11004, USA
[j] The University of North Carolina at Chapel Hill, 434 Greenlaw, Campus Box 3520, Chapel Hill, NC 27599

ARTICLE INFO

Article history:
Received 28 July 2017
Received in revised form 6 November 2017
Accepted 24 November 2017
Available online xxxx

Keywords:
SIPS
PQ-B
NAPLS
Psychosis
Prodromal
Schizophrenia
Screener
Machine learning

ABSTRACT

Machine learning techniques were used to identify highly informative early psychosis self-report items and to validate an early psychosis screener (EPS) against the Structured Interview for Psychosis-risk Syndromes (SIPS). The Prodromal Questionnaire–Brief Version (PQ-B) and 148 additional items were administered to 229 individuals being screened with the SIPS at 7 North American Prodrome Longitudinal Study sites and at Columbia University. Fifty individuals were found to have SIPS scores of 0, 1, or 2, making them clinically low risk (CLR) controls; 144 were classified as clinically high risk (CHR) (SIPS 3–5) and 35 were found to have first episode psychosis (FEP) (SIPS 6). Spectral clustering analysis, performed on 124 of the items, yielded two cohesive item groups, the first mostly related to psychosis and mania, the second mostly related to depression, anxiety, and social and general work/school functioning. Items within each group were sorted according to their usefulness in distinguishing between CLR and CHR individuals using the Minimum Redundancy Maximum Relevance procedure. A receiver operating characteristic area under the curve (AUC) analysis indicated that maximal differentiation of CLR and CHR participants was achieved with a 26-item solution (AUC = 0.899 ± 0.001). The EPS-26 outperformed the PQ-B (AUC = 0.834 ± 0.001). For screening purposes, the self-report EPS-26 appeared to differentiate individuals who are either CLR or CHR approximately as well as the clinician-administered SIPS. The EPS-26 may prove useful as a self-report screener and may lead to a decrease in the duration of untreated psychosis. A validation of the EPS-26 against actual conversion is underway.

1. Introduction

Clinicians that attempt to ameliorate the symptoms of schizophrenia and other psychoses, after the symptoms have developed, have been met with limited success. A newer approach is identifying individuals who are at increased risk of developing psychotic disorders in order to prevent progression of the illness and to decrease the duration of untreated psychosis (Kline and Schiffman, 2014). The Structured Interview for Psychosis-risk Syndromes (SIPS) was developed to identify clinically high risk (CHR) individuals in order to evaluate the natural history of the illness during the prodromal period and to identify interventions that could help prevent progression (Miller et al., 1999, 2002; McGlashan et al., 2001). The SIPS is the "gold standard" early psychosis assessment in North America, but it is also a structured interview that

Abbreviations: AUC, area under the curve; CHR, clinically high risk; CLR, clinically low risk; COPE, Center of Prevention and Evaluation; EPS, early psychosis screener; FEP, first episode psychosis; mRMR, Minimum Redundancy Maximum Relevance; NAPLS, North American Prodrome Longitudinal Study; PQ-B, Prodromal Questionnaire – Brief Version; ROC, receiver operating characteristic; SIPS, Structured Interview for Psychosis-risk Syndromes.

* Corresponding author.
E-mail addresses: bb@telesage.com (B.B. Brodey), Ragy.girgis@nyspi.columbia.edu (R.R. Girgis), favorov@email.unc.edu (O.V. Favorov), jmadding@ucalgary.ca (J. Addington), diana_perkins@med.unc.edu (D.O. Perkins), cbearden@mednet.ucla.edu (C.E. Bearden), scott.woods@yale.edu (S.W. Woods), psyefw@emory.edu (E.F. Walker), cornblat@lij.edu (B.A. Cornblatt), barbara.walsh@yale.edu (B. Walsh), kelkin@telesage.com (K.A. Elkin), brodey@email.unc.edu (I.S. Brodey).

https://doi.org/10.1016/j.schres.2017.11.030
0920-9964/

takes about 90 minutes to administer and requires extensive training to assure high inter-rater reliability (Miller et al., 2003). For these reasons, its use is often restricted to research centers. The Prodromal Questionnaire – Brief Version (PQ-B) was developed a few years later in order to simplify the process of identifying individuals who are CHR (Loewy et al., 2005, 2011a). Although other instruments have been developed for screening purposes, the PQ-B is the most researched self-report screener (Jarrett et al., 2012; Kline et al., 2012a, 2012b; Loewy et al., 2011b; Okewole et al., 2015). Despite the research behind it, the high false positive rate of the PQ-B may make it unsuitable for widespread use as a screener in many populations (Kline et al., 2012b; Xu et al., 2016). Given the low prevalence of early psychosis in the general population, it is desirable to have a more specific screener for early psychosis to promote early intervention (Cohen and Marino, 2013; Comparelli et al., 2014).

In an earlier project, TeleSage developed a self-report item bank to serve as the foundation for developing an early psychosis screener (EPS) (Brodey et al., 2017). We assembled a panel of experts and implemented a rigorous survey item development, modification, and selection process. This process included 40 participants and up to five rounds of cognitive interviewing per item (Willis, 2005). We identified a subset of 148 items that were well understood by prodromal individuals and that our expert panel believed would cover the breadth of concepts associated with the prodromal period and early psychosis. After removing items from the survey that were unnecessary for our analyses (see Section 3.1.1), we were left with 124 items for the machine learning analysis.

In initiating the present study, we wanted to validate an EPS instrument based on the rigor of the established North American Prodrome Longitudinal Study (NAPLS) clinics and the Center of Prevention and Evaluation (COPE) clinic at Columbia University. We used machine learning techniques and the response sets gathered from established prodromal sites to maximize our ability to develop a useful EPS.

Our hypothesis is that machine learning techniques can be used to select a minimal subset of the 124 self-report items that can be used to identify with high sensitivity and specificity individuals who are at clinically high risk for developing psychosis.

2. Methods

2.1. Participants

TeleSage, Inc. partnered with the Columbia University COPE Clinic and seven NAPLS research sites, located at Emory University, University of Calgary, UCLA, UCSD, UNC-Chapel Hill, Yale University, and Zucker Hillside Hospital. All of the clinical participants in this study were recruited from these eight sites. Overall, we recruited 229 participants (demographic information is presented in Table 1). The recruitment procedures for the NAPLS sites and COPE have been comprehensively described in the literature (Addington et al., 2012; Brucato et al., 2017).

IRB approval was obtained for all sites at their host institutions, and all participants provided IRB-approved informed consent. At the NAPLS sites and at the COPE clinic the CLR, CHR, and FEP groups were defined by the Criteria of Psychosis-risk Syndromes (COPS), contained in the SIPS (McGlashan et al., 2001). Exclusion criteria included attenuated positive symptoms better accounted for by another psychiatric condition, past or present full-blown psychosis, I.Q. < 70, medical conditions known to affect the central nervous system, and current serious risk of harm to self or others. Eligible participants in this study were recruited from a pool of patients who were already receiving a SIPS evaluation for a primary CHR-related study (see Miller et al., 2003 for a description of the SIPS assessment procedures). Individuals who received the SIPS were asked to participate in the EPS study. Participants who scored a 0, 1, or 2 on the all of the SIPS positive symptoms were placed in the clinically low risk (CLR) group. Participants who scored a 3, 4, or 5 on one or more of the SIPS positive symptoms were placed in the CHR group. Participants scoring 6 on any of the SIPS positive symptoms were placed in the active psychosis (FEP) group. All participants completed paper assessments including 9 demographics items, our 148 test items, and the PQ-B.

Table 1
Demographics of the studied groups.

Group	n	Age (years)	Female[a]	White[a]	Black[a]	Asian[a]	Hispanic[a]	Other[a]
CLR	50	20.1 ± 4.0	26.0	48.2	16.1	7.1	12.5	16.1
CHR	144	20.7 ± 4.8	42.4	53.2	16.7	9.0	9.6	11.5
FEP	35	22.6 ± 4.6	45.7	54.1	21.6	5.4	8.1	10.8

[a] Data reported as percentages of the assigned group.

2.2. Analytical procedures

The analyses were performed on the participants' answers to the questionnaire items. The goal of this study was to develop the most effective computational procedure for reducing the Likert scale survey answers of a tested individual to a single quantitative metric, or a score, that could be used to infer that individual's SIPS class identity. The simplest such metric is a linear sum of answers to all the items:

$$M_{LS} = \sum_{i \in Q} L_i \quad (1)$$

where Q is a set of questionnaire items and L_i is the Likert scale answer to the ith item.

The linear sum metric M_{LS} is limited in its representational power, however, since it treats all the items as contributing uniformly to SIPS class estimation. In the Supplementary information published online, we consider more versatile linear and nonlinear metrics but find that their CLR vs. CHR discriminatory performance is not superior to the performance of the linear sum metric M_{LS}. Consequently, we chose M_{LS} as the best metric suited for our screener.

The capacity of M_{LS} to accurately predict which SIPS class a tested individual belongs to based on his/her EPS questionnaire answers was evaluated using receiver operating characteristic (ROC) analyses. The classification accuracy was expressed as the area under the ROC curve (AUC). AUC values can range between 0.5 (for classifiers whose performance is completely random) and 1 (for perfectly accurate classifiers).

Two analytical approaches were used to identify those among the original list of 124 survey items that could be safely omitted from the final list. The first approach was spectral clustering, which was used to identify clusters of the questionnaire items with distinctly different patterns of answers among individuals belonging to CLR, CHR, and FEP groups (Shi and Malik, 2000; Ng et al., 2001; von Luxburg, 2007). We measured the similarity between different items by computing their correlation coefficient over all four groups of subjects. Such pairwise correlation coefficients make up a similarity matrix S. Importantly, no information about the subjects' group membership was used in computing the correlation coefficients and, therefore, in creating the similarity matrix S. This similarity matrix S is used to construct normalized graph Laplacian matrix:

$$L_{NCut} = D^{-1/2} \cdot (D - S)^{-1/2} \quad (2)$$

where D is a diagonal matrix, in which $D_{ii} = \sum_j S_{ij}$. To determine how many distinct groups are present among the items, we compute and plot "eigengaps" between consecutive eigenvalues $\lambda_1 \ldots \lambda_N$ of L_{NCut} matrix (the ith eigengap is defined as a difference $\Delta\lambda_i = \lambda_{i+1} - \lambda_i$; with the first eigengap, $\Delta\lambda_1$, set to zero). In general, if a dataset has K distinct clusters, the eigengap plot will have an outstanding eigengap in the K position ($\Delta\lambda_K$) and also likely to the left of it, but not to the right. The corresponding Kth eigenvector sorts all the items into two groups, which can be seen by plotting that eigenvector. (For an in-

Please cite this article as: Brodey, B.B., et al., The Early Psychosis Screener (EPS): Quantitative validation against the SIPS using machine learning, Schizophr. Res. (2017), https://doi.org/10.1016/j.schres.2017.11.030