HHS Public Access

Author manuscript

*Schizophr Res*. Author manuscript; available in PMC 2019 June 23.

Published in final edited form as:

*Schizophr Res*. 2018 July ; 197: 522–530. doi:10.1016/j.schres.2017.12.001.

# Latent Class Cluster Analysis of Symptom Ratings Identifies Distinct Subgroups Within the Clinical High Risk for Psychosis Syndrome

**Arthur T. Ryan**[a], **Jean Addington**[b], **Carrie E. Bearden**[c], **Kristin S. Cadenhead**[d], **Barbara A. Cornblatt**[e], **Daniel H. Mathalon**[f], **Thomas H. McGlashan**[g], **Diana O. Perkins**[h], **Larry J. Seidman**[i,j], **Ming T. Tsuang**[e], **Scott W. Woods**[f], **Tyrone D. Cannon**[k], **Elaine F. Walker**[a,l]

[a]Department of Psychology, Emory University, Atlanta, GA, United States

[b]Department of Psychiatry, University of Calgary, Calgary, Alberta, Canada

[c]Semel Institute for Neuroscience and Human Behavior & Department of Psychology, University of California, Los Angeles, Los Angeles, CA, United States

[d]Department of Psychiatry, UCSD, San Diego, CA, United States

[e]Department of Psychiatry, Zucker Hillside Hospital, Long Island, NY, United States

[f]Department of Psychiatry, UCSF, and SFVA Medical Center, San Francisco, CA, United States

[g]Department of Psychiatry, Yale University, New Haven, CT, United States

[h]Department of Psychiatry, University of North Carolina, Chapel Hill, NC, United States

[i]Department of Psychiatry, Harvard Medical School at Beth Israel Deaconess Medical Center, Boston, MA, United States

[j]General Hospital, Boston, MA, United States

[k]Department of Psychology, Yale University, New Haven, CT, United States

[l]Department of Psychiatry, Emory University, Atlanta, GA, United States

## Abstract

The clinical-high-risk for psychosis (CHR-P) syndrome is heterogeneous in terms of clinical presentation and outcomes. Identifying more homogenous subtypes of the syndrome may help

Corresponding Author: Arthur T. Ryan, 36 Eagle Row, #270, Atlanta GA 30307, Tel: 404-727-7547, Fax: 404-727-1284, arthur.t.ryan@gmail.com.

**Conflict of Interest**

There are no conflicts of interest for any of the authors with respect to the data in this paper or for the study.

**Contributors**

Dr. Arthur Ryan undertook the statistical analysis and wrote the first draft of the manuscript. Dr. Elaine Walker was involved in the writing of subsequent drafts of the manuscript. All of the authors listed were involved in study design and have contributed to and approved the final manuscript.

clarify its etiology and improve the prediction of psychotic illness. This study applied latent class cluster analysis (LCCA) to symptom ratings from the North American Prodrome Longitudinal Studies 1 and 2 (NAPLS 1 and 2). These analyses produced evidence for three to five subgroups within the CHR-P syndrome. Differences in negative and disorganized symptoms distinguished amongst the subgroups. Subgroup membership was found to predict conversion to psychosis. The authors contrast the methods employed within this study with previous attempts to identify more homogenous subgroups of CHR-P individuals and discuss how these results could be tested in future samples of CHR-P individuals.

**Keywords**

Schizophrenia; Prodrome; Finite Mixture Models; Disorganization Symptoms; Heterogeneity

## 1 Introduction

Individuals with the clinical high risk for psychosis (CHR-P) syndrome (also known as the psychosis prodrome, schizophrenia prodrome, and ultra-high-risk syndrome) have a 17–25% chance of developing a psychotic illness within two years (Fusar-Poli et al., 2016). However, symptoms and outcomes among CHR-P individuals are highly heterogeneous (Fusar-Poli, 2017). Identifying more homogenous phenotypic subgroups within the CHR-P syndrome may aid in clarifying prognosis, etiology, and response to treatment (Compton et al., 2014).

Valmaggia et al. (2013) applied a latent class cluster analysis (LCCA) to Comprehensive Assessment of At-Risk Mental State (CAARMS) symptom ratings (Yung et al., 2005) of CHR-P participants to identify more homogenous subgroups of CHR-P individuals on the basis of symptom configurations. Their analysis identified four subgroups that varied primarily in terms of symptom severity. Subgroup membership predicted important clinical outcomes, such as rates of conversion to psychotic illness.

In the current study, we apply LCCA to identify subgroups based on symptom ratings from the Structured Interview of Prodromal Symptoms (SIPS) and its companion rating scale, the Scale of Prodromal Symptoms (Miller et al., 2003). Conducting an analysis similar to the one conducted by Valmaggia et al. has several important functions. Such an analysis can determine differences between the SIPS and the CAARMS result in different clustering solutions. While similar, both the SIPS and CAARMS assess content areas not measured by the other. The SIPS and CAARMS also divide up symptomatology differently amongst their respective symptom rating scales. See Table 1 for a comparison of the symptoms assessed by the CAARMS and SIPS. If a similar cluster structure emerges from the current analysis, this would suggest that the overlapping content of the SIPS and CAARMS is sufficient to identify the same CHR-P subgroups. Contrastingly, if a different subgroup structure emerges, this would suggest that differences between the SIPS and CAARMS may prevent the identification of one or more of the subgroups identified by the other instrument. A failure of our LCCA to replicate Valmaggia's results might also suggest important differences regarding subject recruitment and other extraneous factors between our sample

and Valmaggia's: some authors have cited such factors as a pervasive challenge to developing reliable subtyping strategies generally (Cornblatt et al., 2015).

The goals of this study were: (1) employ LCCA to attempt to identify subtypes/subgroups within the CHR-P syndrome on the basis of symptom ratings and (2) determine if the LCCA-derived subgroups differed in terms of their demographics, clinical symptoms, and rates of conversion to psychotic illness.

## 2 Methods

### 2.1 Sample Description

Data were collected as part of the first and second iteration of the North American Prodrome Longitudinal Study: NAPLS 1 and NAPLS 2 (Addington et al., 2012, 2007). Detailed information regarding the samples can be found in the referenced papers. Both studies admitted individuals who met criteria for any of three risk syndromes: attenuated positive symptoms (APS), genetic risk and deterioration (GRD), and brief intermittent psychotic symptoms (BIPS). Analyses for this study were restricted to the 356 NAPLS 1 and 737 NAPLS 2 CHR-P subjects who had complete baseline symptom data. One difference between the NAPLS 1 and 2 recruitment criteria was that NAPLS 2 added an additional CHR-P syndrome: being younger than 18-years-old and having a diagnosis of schizotypal personality disorder (YSPD). Nine percent of the NAPLS 2 sample met criteria for YSPD, but only 18 individuals (2.4% of the NAPLS 2 CHR-P sample) met criteria solely for YSPD. The demographics of the NAPLS 1 and 2 samples are shown in Table 2. All procedures were approved by the Institutional Review Board (IRB) at each site. Written informed consent (with assent from participants younger than 18) was obtained from all participants.

### 2.2 Clinical Measures

CHR-P symptoms were assessed using the Structured Interview for Prodromal Syndromes (SIPS) and its companion scale, the Scale of Prodromal Symptoms (Miller et al., 2003). Nineteen SIPS symptom items are rated 0–6 based on their severity and those items are categorized into four domains (positive, negative, disorganized, and general). These domains were modeled after the ones set out by Yung et al. in the CAARMS (Fusar-Poli et al., 2017). Medication history was assessed with a lifetime medication history interview. Individual medications had only been coded into distinct classes and divided between lifetime and current use for the NAPLS 2 dataset, so psychotropic medication history analyses were restricted to the NAPLS 2 dataset. Demographic data were collected using a demographics interview.

### 2.3 Statistical Analysis

Statistical analyses were performed using R version 3.3.1 (R Core Team, 2016) supplemented with the *mclust* package (Fraley et al., 2012; Fraley and Raftery, 2002). The mclust package implements latent class cluster analysis (LCCA) by attempting to identify a best fitting Gaussian finite mixture model—i.e., the one with the lowest Bayesian information criterion (BIC) value—using an expectation-maximization (EM) algorithm. Separate LCCAs were computed for the NAPLS 1 and 2 samples. ANOVA tests, $\chi^2$ tests,

and Kaplan-Meir survival analyses were conducted to compare the LCCA-derived subgroups on relevant variables and any significant tests were followed up with pairwise comparisons. SPSS 17 was used for ANOVA and $\chi^2$ analyses.

## 3 Results

### 3.1 NAPLS 1 and 2 Sample Comparisons

Demographic and SIPS syndrome information for the NAPLS 1 and 2 samples are shown in Table 2. The samples differed significantly in race ($\chi^2 = 50.916$, df =6, $p < .001$): pairwise comparisons are shown in Table 2. NAPLS 1 had a greater proportion of individuals with APS ($\chi^2 = 7.032$, df =1, $p < .01$), although this difference was not large in absolute terms (96% APS prevalence in NAPLS 1 vs. 92% in NAPLS 2).

### 3.2 NAPLS 1 LCAA

A LCCA analysis of the NAPLS 1 baseline SIPS data identified an ellipsoidal, equal volume and orientation (EVE) Gaussian distribution with three classes as the best fitting model (log. likelihood = −10753, n = 356, df = 285, BIC = 23181, clustering table = 40/108/208). The normalized entropy value was 0.93, suggesting the model was a good fit for the data. The first subgroup was distinguished by its large size (58% of the total sample), the presence of perceptual abnormalities, and low levels of negative, disorganized, and general symptoms. This subgroup was labeled the Perceptual Abnormalities Subgroup (PAS). The second subgroup (30% of the total sample) was distinguished by elevated negative symptoms, disorganized speech, other disorganization symptoms, and general symptoms. We labeled this group the Disorganized Speech Subgroup (DSS). The third subgroup (11% of the total sample) was distinguished by hygiene impairment without marked disorganized speech. We labeled this group the Impaired Hygiene Subgroup (IHS).

Demographic and clinical measures of the NAPLS 1 subgroups are summarized in Table 3. ANOVAs and $\chi^2$ tests showed that the subgroups did not differ significantly on any of the demographic variables with the exception of gender ($\chi^2 = 6.749$, df = 2, $p < .05$). Pairwise comparisons showed that PAS individuals were more likely to be female than IHS individuals ($p < .05$).$\chi^2$ tests of the SIPS syndrome categories showed a significant difference between the subgroups in the prevalence of GRD ($\chi^2 = 29.567$, df = 2, $p < .001$). Pairwise comparisons showed that GRD prevalence was higher in the DSS ($p < .05$). The time between CHR-P participants baseline assessment and final assessment was then examined. A minority of the NAPLS 1 sample (21.6%) had final follow-up times as far out as four years, far beyond the two-year timeline of the study. In order to ensure these outliers were not having an outsized influence on the analysis, the Kaplan-Meir survival analysis was run with and without a 2.5 year follow-up cut-off. The results were similar in both cases and the 2.5 year cut-off results are reported here for parity with the NAPLS 2 analysis below. The Kaplan-Meir survival curve comparing conversion rates across the subgroups was significant (Mantel-Cox $\chi^2 = 8.104$, df = 2, $p < .05$). A graph of the survival curves is shown in Figure 1. Mantel-Cox pairwise comparisons showed that the DSS had a higher conversion rate than the PAS ($\chi^2 = 7.290$, df = 1, $p < .01$). There was a trend towards the DSS having a higher conversion rate than the IHS ($\chi^2 = 2.759$, df = 1, $p = .097$). The results of ANOVAs

comparing the subgroups' symptom ratings are shown in Table 4. The subgroup symptom means are also shown as a line graph in Figure 2.

### 3.3 NAPLS 2 LCCA

A LCCA analysis of the NAPLS 2 baseline SIPS data identified an EVE Gaussian distribution with five classes as the best fitting model (log. likelihood = −22103.56, N = 737, df = 361, BIC = 46590, clustering table = 120/134/40/30/413). The normalized entropy value was 0.82, suggesting that the model was a good fit for the data. When the five subgroup solution was examined, it was found that the three largest subgroups cumulatively comprised 91% of the sample, with the two additional subgroups making up 5% and 4% of the sample respectively. The symptom ratings and relative frequencies of the three largest subgroups resembled those found in the NAPLS 1 model. The larger of the two new subgroups was labeled the Odd and Euthymic Subgroup (OES) because of its low ratings on scales of distress and psychopathology, along with elevated ratings of odd behavior and thought. The second additional subgroup was labeled the Distressed and Avolitional Subgroup (DAS) due to its high ratings of distress and impairment, which contrasted with its uniquely low positive symptom ratings.

Demographic and clinical variables for the NAPLS 2 subgroups are summarized in Table 5. ANOVA and $\chi^2$ tests showed that the subgroups did not differ significantly on any of the demographic variables with the exception of mother's education [$F(4, 719) = 4.521$, $p = .01$]. The mothers of DSS and OES individuals were more likely to have completed college than the mothers of PAS participants ($p < .05$). $\chi^2$ tests of the SIPS syndrome categories showed a significant difference between the subgroups in the prevalence of APS ($\chi^2 = 15.612$, df = 4, $p < .01$), GRD ($\chi^2 = 21.397$, df = 4, $p < .001$), and YSPD ($\chi^2 = 25.924$, df = 4, $p < .001$). Significant pairwise comparisons are shown in Table 5. Similarly to the NAPLS 1 cohort, when the time between baseline assessment and final assessment was examined, a minority of the NAPLS 2 cohort had follow-up dates up to four years after their baseline assessment. However, only 3% (n = 18) of the NAPLS 2 sample had final assessments greater than 2.5 years after their baseline, so this small group of outliers was simply excluded before running the analysis. A Kaplan-Meir survival curve comparing conversion rates across the subgroups was significant (Mantel-Cox $\chi^2 = 11.062$, df = 4, $p < .05$). Mantel-Cox pairwise comparisons showed that DSS individuals were more likely to convert over time than PAS individuals ($\chi^2 = 9.105$, df = 1, $p < .01$). There was a trend towards DSS individuals being more likely to convert over time than IHS individuals ($\chi^2 = 3.620$, df = 1, $p = .057$) and OES individuals ($\chi^2 = 3.104$, df = 1, $p = .078$). The survival curve is shown in Figure 3. The results of ANOVAs comparing the subgroups' symptom ratings are shown in Table 6. The subgroup symptom means are also shown as a line graph in Figure 4.

The NAPLS 2 subgroups were additionally compared in terms of lifetime and baseline medication treatment with the following medication classes: antidepressant, mood stabilizer, antipsychotic, stimulant, benzodiazepine, and any psychotropic medication. The only significant $\chi^2$ test was for lifetime mood stabilizer use ($\chi^2 = 17.795$, df =4, $p < .001$). Pairwise comparisons are shown in Table 5.

## 4 Discussion

This paper describes the results of latent class cluster analyses (LCCAs) to identify subgroups of the clinical high risk for psychosis (CHR-P) syndrome based on SIPS symptom ratings. These analyses produced statistical support for the existence of three subgroups within the NAPLS 1 and NAPLS 2 samples: a Perceptual Abnormalities Subgroup (PAS), Disorganized Speech Subgroup (DSS), and Impaired Hygiene Subgroup (IHS). Two additional low-frequency subgroups were found in the NAPLS 2 sample, the Odd and Euthymic Subgroup (OES) and the Distressed and Avolitional Subgroup (DAS).

### 4.1 Comparisons Between This Study and Previous Attempts to Identify Homogenous CHR-P Subgroups

The LCCAs described here differ in some significant ways from previous attempts to identify homogenous subgroups of CHR-P individuals. Fusar-Poli and colleagues (2016) conducted a meta-analysis to determine whether different CHR-P syndromes—i.e., Attenuated Positive Symptoms (APS), Genetic Risk and Deterioration (GRD), and Brief Intermittent Psychotic Symptoms (BIPS) syndromes—have different rates of conversion to psychosis. Fusar-Poli et al. concluded that BIPS individuals did indeed appear to have an increased rate of conversion to psychosis and that GRD individuals did not appear to differ from individuals who did not meet criteria for CHR-P. In contrast to Fusar-Poli et al.'s findings, we found that LCCA-derived subgroups were not differentiated by the prevalence of BIPS. Also, the NAPLS 1 and 2 LCCA subgroup with the highest rate of conversion, the DSS, had the highest prevalence of GRD. This suggests that previously defined CHR-P syndromes may themselves be heterogenous categories and not represent natural divisions within the CHR-P syndrome.

In another attempt to identify homogenous subgroups of CHR-P individuals, Cornblatt and colleagues (2015) restricted their CHR-P sample to only those who met criteria for APS. While they were able to identify a model that predicted conversion accurately in their study's sample, their model was not predictive when applied to a similarly selected sample (Addington et al., 2017). Cornblatt et al.'s results suggest that more restrictive recruitment criteria may not be sufficient for identifying more homogenous subgroups of CHR-P individuals.

Several of the previous attempts to identify more homogenous CHR-P individuals have adopted a "staging model" framework, meaning they assumed that different subgroups represent points along a risk continuum for psychosis (Carrión et al., 2017). In contrast to this, the models described here do not make the assumption that the LCCA derived subgroups represent different points along a risk continuum. In this way, the current project more closely resembles work like that of the Bipolar-Schizophrenia Network for Intermediate Phenotypes (B-SNIPS) (Tamminga et al., 2014). Tamminga and colleagues purposefully collected their psychotic illness probands sample without regard to DSM-5 (American Psychiatric Association, 2013) distinctions between bipolar, schizophrenia, and schizoaffective disorders. Instead, they employed machine learning techniques to identify homogenous subgroups on the basis of neurobiological markers (e.g., EEG measures and neuropsych testing). Their analyses yielded three "biotypes" that cut across the diagnostic

boundaries of the DSM. Most interestingly, these biotypes proved to have novel explanatory power that samples restricted to DSM-5 diagnoses would otherwise miss. Clementz and colleagues (2016) demonstrated that two of the biotypes had similar levels of impairment on tasks of attention despite having diametrically opposed deviations from normative EEG signatures. Clementz et al. proposed that such previously unidentified biotypes might underlie the contradictory findings that have bedeviled EEG research on psychotic disorders.

We should note that our findings were congruent with those of Cornblatt et al. (2015) and Addington et al. (2017) in several respects. Each study found that disorganization symptoms are an important marker for outcomes among CHR-P individuals. Our results are also consistent with previous studies that have shown that negative symptoms, (e.g., alogia and anhedonia) are important in predicting transition to psychosis (Valmaggia et al., 2013).

It is also important to compare the results of our analyses directly with those of previous LCCAs of CHR-P individuals. Table 7 shows the CHR-P symptoms that best distinguished LCCA-derived subgroups in the NAPLS I, NAPLS II, Healey et al. (2017), and Valmaggia et al. (2013) analyses. The top seven most discriminative symptoms have been emphasized for each analysis. As can be seen, five of the top seven most informative symptoms were shared across the NAPLS 1 and 2 lists, while only one of those symptoms appeared on the Healey et al. list and two appeared on the Valmaggia et al. list. We believe that several factors lead to the greater similarity between the NAPLS 1 and 2 analyses. First, both analyses had larger subjects-to-parameters-ratio, which increases reliability. Second, the NAPLS samples were collected by several of the same sites (with some additional sites added in NAPLS 2) using similar protocols. Third, Healey et al. incorporated additional data sources (depression scores and neurocognitive testing) in their LCCA model. Finally, the Valmaggia et al. sample differed in that symptoms were measured using the CAARMS.

Some specific differences between our analyses and those conducted by Healey et al should be noted. Healey et al. combined CHR-P individuals with help-seeking controls (who generally met sub-threshold versions of CHR-P criteria) in their LCCA. This was an understandable choice, both to increase their sample size and because of their interest in determining whether these relaxed criteria still identified individuals at risk for psychosis. In practice, however, this choice meant that only 171 CHR individuals were used to estimate 106, 133, and 160 different parameter values for their three, four, and five class LCCA's respectively (see Table 1 in Healy et al. 2017). Such low subjects-to-estimated-parameters ratios are likely to produce unstable solutions (Dolnicar, 2002). Healey et al. also included several measures other than SIPS symptom ratings into their LCCA model (i.e., depression symptom ratings and neurocognitive testing). While such rich data brings the possibility of identifying novel subgroups, additional data is not necessarily useful for identifying meaningful clinical subgroups. For example, research into atypical depression has shown that the single symptom of increased appetite provides excellent discrimination of this depression subtype (Milaneschi et al., 2016).

Comparing our results to those of Valmaggia and colleagues (2013) is of particular importance, as Valmaggia et al.'s work represents, to our knowledge, the largest sample used for a LCCA of CHR-P symptom ratings prior to our own. While Valmaggia et al.

emphasized the importance of negative and disorganization symptoms in differentiating their subgroups, the individual symptom that best differentiated their subgroups was "subjective motor functioning," whose $R^2$ effect size was a remarkable .94 (see Table 4 in Valmaggia et al. 2013 and Table 7 in this paper). This is an interesting finding for several reasons. Motor functioning deficits have been shown to be an informative biomarker for psychotic illness (Schiffman, 2017). Secondly, the SIPS does not have an item which corresponds to subjective motor functioning (Miller et al., 2003). The SIPS "motor impairment" item is rated on a combination of reported/observed motor deficits and reported/observed abnormal movements. Perhaps due to this fact, the corresponding SIPS item was not nearly as discriminative in the NAPLS 1 and 2 LCCAs. It should be noted that our analyses of the NAPLS samples yielded a similarly, though not quite as, informative symptom rating, namely "impairment in personal hygiene," whose $\eta^2$ values were .67 and .66 in the NAPLS 1 and 2 LCCAs respectively. Valmaggia's finding regarding subjective motor impairment in the context of ours regarding hygiene impairment demonstrates some of the promise of LCCA and similar techniques. It shows that seemingly minor differences in the content coverage of the SIPS and CAARMS may possibly have a significant impact on their ability to detect underlying subtypes of the CHR-P syndrome.

A few additional specific differences between our LCCAs and Valmaggia et al.'s bear mentioning: (a) Valmaggia et al.'s high-risk subtype composed only approximately 6% of their sample (17 individuals) while the DSS composed 30% and 18% of the NAPLS 1 and 2 samples respectively, (b) our subgroups were not best characterized as a stepwise pattern of symptom severity, (c) our subgroups did not vary in age, (d) Valmaggia et al.'s subgroups were not differentiated by disorganized speech, other than in so far as their mild subgroup had lower levels than their other subgroups.

While the differences between Valmaggia et al.'s findings and our own are interesting, the parallels are also important. Four of the seven most informative symptoms in Valmaggia et al.'s analysis were also among the most informative in either the NAPLS 1 or 2 LCCA (see Table 7). In line with Valmaggia et al.'s discussion of their findings, these most informative symptoms were negative and cognitive symptoms.

### 4.2 Ideas for Future Testing of Putative CHR-P Subtypes and Conclusion

Narrative descriptions of similarities and differences between clustering solutions are not a conclusive method for generalizing our findings. We propose that the analyses presented here are a first step towards attempts to validate our findings in future datasets. One way that this work could be continued is by attempting to replicate the results of our LCCA analysis in a new sample by applying the weights of our LCCA models to other samples and checking if they still accurately characterize the data. Perhaps even more interestingly, one could apply a random forest analysis (Breiman, 2001) to identify a human-interpretable set of binary questions (i.e., a decision tree algorithm) that could be used to assign CHR-P individuals to subgroups on the basis of their SIPS symptom ratings. For an example of an application of decision tree algorithms in suicide research, see Kessler et al., 2017. Once CHR-P individuals have been classified into LCCA-derived subgroups, investigators may then determine whether those subgroups have similar charecteristics to the ones found in our

NAPLS LCCA-derived subgroups (e.g., that individuals identified as DSS indeed have the highest rate of conversion, etc.).

It is important to note several important limitations of our study. We were not able to directly test existing LCCA models of CHR-P symptoms, such as the ones described by Valmaggia et al. (2013) or Healey et al. (2017). In the case of Valmaggia et al., this could not be done given their use of the CAARMS. In regards to Healey et al., we could not apply their LCCA solution to our own data as they included neuropsychological testing and depression symptom ratings in their clustering solution. In contrast, we specifically sought to see what subgroups (if any) could be identified on the basis of SIPS ratings alone. Second, our analyses of conversion rates were hampered by the high rates of loss to follow-up in the NAPLS 1 and 2 datasets. However, the use of Kaplan-Meir survival analyses to account for data-censoring due to loss to follow-up helped to address this issue.

The work presented in this paper has several noteworthy strengths. To our knowledge, our analyses included the two largest samples of CHR-P symptom ratings used with LCCA to date. Second, our LCCA analyses on two separate datasets identified recognizably similar subgroup solutions. A third strength of our analysis is that the subgroups identified by our LCCAs did not differ primarily in terms of global symptom severity, but rather primarily on the basis of specific configurations of elevated symptoms (e.g., disorganized speech in the DSS). Fourth, our subgroups were not explicable terms of existing classification schemes of CHR-P (e.g., APS vs. BIPS).

This paper presents the results of a series of LCCAs to identify subgroups of CHR-P individuals using the NAPLS 1 and 2 datasets. These analyses evinced the existence of three to five subgroups which appear to differ meaningfully on clinical outcomes and other variables.

## Acknowledgments

We thank Dr. John Hanfelt for reviewing an early version of the analyses for this paper and providing helpful feedback.

**Funding Bodies**

This work was supported by a collaborative U01 award from the National Institute of Mental Health at the National Institutes of Health (MH081902 to TDC; MH081857 to BAC; MH081988 to EW; MH081928 to LJS; MH082004 to DP; MH082022 to KC; MH081984 to JA; MH082022 to SWW) and NIMH P50 MH066286 and Staglin Music Festival for Mental Health (CEB), NIMH P50 MH080272, and the Commonwealth of Massachusetts (SCDMH82101008006) to LJS. AR is supported by the Department of Veterans Affairs Office of Academic Affiliations Advanced Fellowship Program in Mental Illness Research and Treatment. His work was supported with resources and the use of facilities at the VA Capitol Health Care Network (VISN 5) MIRECC and the University of Maryland School of Medicine Department of Psychiatry. This article reflects the authors' personal views and in no way represents the official view of the Department of Veterans Affairs of the U.S. Government.

## References

Addington J, Cadenhead KS, Cannon TD, Cornblatt B, McGlashan TH, Perkins DO, Seidman LJ, Tsuang M, Walker EF, Woods SW. 2007; North American Prodrome Longitudinal Study: A collaborative multisite approach to prodromal schizophrenia research. Schizophr Bull. 33:665–672. [PubMed: 17255119]

Addington J, Cadenhead KS, Cornblatt BA, Mathalon DH, McGlashan TH, Perkins DO, Seidman LJ, Tsuang MT, Walker EF, Woods SW, Addington JA, Cannon TD. 2012; North American Prodrome Longitudinal Study (NAPLS 2): Overview and recruitment. Schizophr Res. 142:77–82. DOI: 10.1016/j.schres.2012.09.012 [PubMed: 23043872]

Addington J, Liu L, Perkins DO, Carrion RE, Keefe RSE, Woods SW. 2017; The role of cognition and social functioning as predictors in the transition to psychosis for youth with attenuated psychotic symptoms. Schizophr Bull. 43:57–63. DOI: 10.1093/schbul/sbw152 [PubMed: 27798225]

American Psychiatric Association. Diagnostic and statistical manual of mental disorders. 5. American Psychiatric Pub; Washington, DC: 2013.

Breiman L. 2001; Random Forests. Mach Learn. 45:5–32. DOI: 10.1023/A:1010933404324

Carrión RE, Correll CU, Auther AM, Cornblatt BA. 2017; A severity-based clinical staging model for the psychosis prodrome: Longitudinal findings from the New York Recognition and Prevention Program. Schizophr Bull. 43:64–74. DOI: 10.1093/schbul/sbw155 [PubMed: 28053131]

Clementz BA, Sweeney JA, Hamm JP, Ivleva EI, Ethridge LE, Pearlson GD, Keshavan MS, Tamminga CA. 2016; Identification of distinct psychosis biotypes using brain-based biomarkers. Am J Psychiatry. 173:373–384. DOI: 10.1176/appi.ajp.2015.14091200 [PubMed: 26651391]

Compton MT, Kelley ME, Ionescu DF. 2014; Subtyping first-episode non-affective psychosis using four early-course features: Potentially useful prognostic information at initial presentation. Early Interv Psychiatry. 8:50–58. DOI: 10.1111/eip.12026 [PubMed: 23343467]

Cornblatt BA, Carrión RE, Auther A, McLaughlin D, Olsen RH, John M, Correll CU. 2015; Psychosis prevention: A modified clinical high risk perspective from the Recognition and Prevention (RAP) program. Am J Psychiatry. 172:986–994. DOI: 10.1176/appi.ajp.2015.13121686 [PubMed: 26046336]

Dolnicar, S. A review of unquestioned standards in using cluster analysis for data-driven market segmentation, in: CD Conference Proceedings of the Australian and New Zealand Marketing Academy Conference 2002 (ANZMAC 2002). Presented at the Australian and New Zealand Marketing Academy Conference 2002 (ANZMAC 2002); Melbourne: Deakin University; 2002.

Fraley, C, Raftery, A, Murphy, TB, Scrucca, L. mclust Version 4 for R: Normal Mixture Modeling for Model-Based Clustering, Classification, and Density Estimation Technical Report. Department of Statistics, University of Washington; 2012.

Fraley C, Raftery AE. 2002; Model-based clustering, discriminant analysis, and density estimation. J Am Stat Assoc. 97:611–631.

Fusar-Poli P. 2017; The Clinical High-Risk State for Psychosis (CHR-P), Version II. Schizophr Bull. 43:44–47. DOI: 10.1093/schbul/sbw158 [PubMed: 28053129]

Fusar-Poli P, Cappucciati M, Borgwardt S, Woods SW, Addington J, Nelson B, Nieman DH, Stahl DR, Rutigliano G, Riecher-Rössler A, Simon AE, Mizuno M, Lee TY, Kwon JS, Lam MML, Perez J, Keri S, Amminger P, Metzler S, Kawohl W, Rössler W, Lee J, Labad J, Ziermans T, An SK, Liu CC, Woodberry KA, Braham A, Corcoran C, McGorry P, Yung AR, McGuire PK. 2016; Heterogeneity of psychosis risk within individuals at clinical high risk: A meta-analytical stratification. JAMA Psychiatry. 73:113.doi: 10.1001/jamapsychiatry.2015.2324 [PubMed: 26719911]

Fusar-Poli P, Cappucciati M, De Micheli A, Rutigliano G, Bonoldi I, Tognin S, Ramella-Cravaro V, Castagnini A, McGuire P. 2017; Diagnostic and prognostic significance of brief limited intermittent psychotic symptoms (BLIPS) in individuals at ultra high risk. Schizophr Bull. 43:48–56. DOI: 10.1093/schbul/sbw151 [PubMed: 28053130]

Healey, KM; Penn, DL; Perkins, D; Woods, SW; Keefe, RSE; Addington, J. Latent profile analysis and conversion to psychosis: Characterizing subgroups to enhance risk prediction [published online ahead of print July 14, 2017]. Schizophr Bull. 2017.

Kessler RC, Hwang I, Hoffmire CA, McCarthy JF, Petukhova MV, Rosellini AJ, Sampson NA, Schneider AL, Bradley PA, Katz IR, Thompson C, Bossarte RM. 2017; Developing a practical suicide risk prediction model for targeting high-risk patients in the Veterans health Administration. Int J Methods Psychiatr Res. 26:e1575.doi: 10.1002/mpr.1575

Milaneschi Y, Lamers F, Peyrot WJ, Abdellaoui A, Willemsen G, Hottenga JJ, Jansen R, Mbarek H, Dehghan A, Lu C, Boomsma DI, Penninx BWJH. CHARGE inflammation working group. 2016;

Polygenic dissection of major depression clinical heterogeneity. Mol Psychiatry. 21:516–522. DOI: 10.1038/mp.2015.86 [PubMed: 26122587]

Miller TJ, McGlashan TH, Rosen JL, Cadenhead K, Ventura J, McFarlane W, Perkins DO, Pearlson GD, Woods SW. 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. 29:703. [PubMed: 14989408]

R Core Team. R: A language and environment for statistical computing. R Foundation for Statistical Computing; Vienna, Austria: 2016.

Schiffman J. 2017; Motor issues in the clinical high risk phase of psychosis. Schizophr Bull. 43:937–938. DOI: 10.1093/schbul/sbx086 [PubMed: 28655205]

Tamminga CA, Pearlson G, Keshavan M, Sweeney J, Clementz B, Thaker G. 2014; Bipolar and Schizophrenia Network for Intermediate Phenotypes: Outcomes across the psychosis continuum. Schizophr Bull. 40:S131–S137. DOI: 10.1093/schbul/sbt179 [PubMed: 24562492]

Valmaggia, LR; Stahl, D; Yung, AR; Nelson, B; Fusar-Poli, P; McGorry, PD; McGuire, PK. Negative psychotic symptoms and impaired role functioning predict transition outcomes in the at-risk mental state: A latent class cluster analysis study; Psychol Med FirstView. 2013. 1–15.

Yung AR, Yung AR, Pan Yuen H, Mcgorry PD, Phillips LJ, Kelly D, Dell'olio M, Francey SM, Cosgrave EM, Killackey E, et al. 2005; Mapping the onset of psychosis: The comprehensive assessment of at-risk mental states. Aust N Z J Psychiatry. 39:964–971. [PubMed: 16343296]

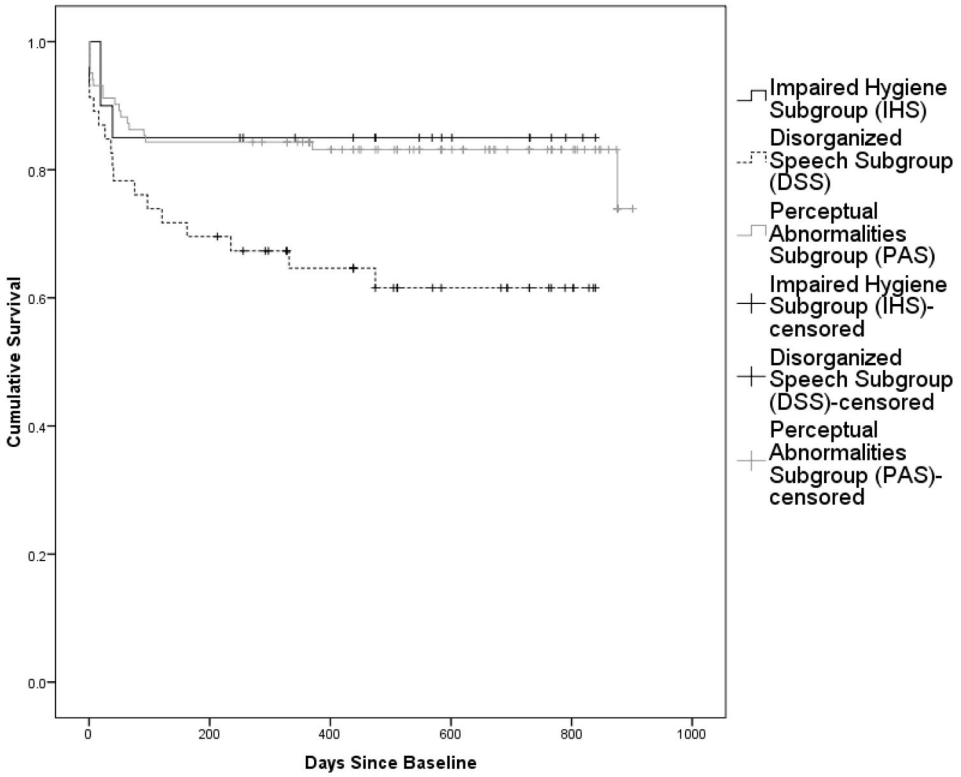


**Figure 1. NAPLS 1 Kaplan-Meir Survival Curve**

This figure graphs the Kaplan-Meir survival curve for conversion to psychosis among the subgroups identified in the NAPLS 1 latent class cluster analysis.


**Figure 2. NAPLS 1 LCCA-Derived Subgroup Symptom Means**

This line graph shows the mean symptom ratings for each of the subgroups identified in the NAPLS 1 latent class cluster analysis.

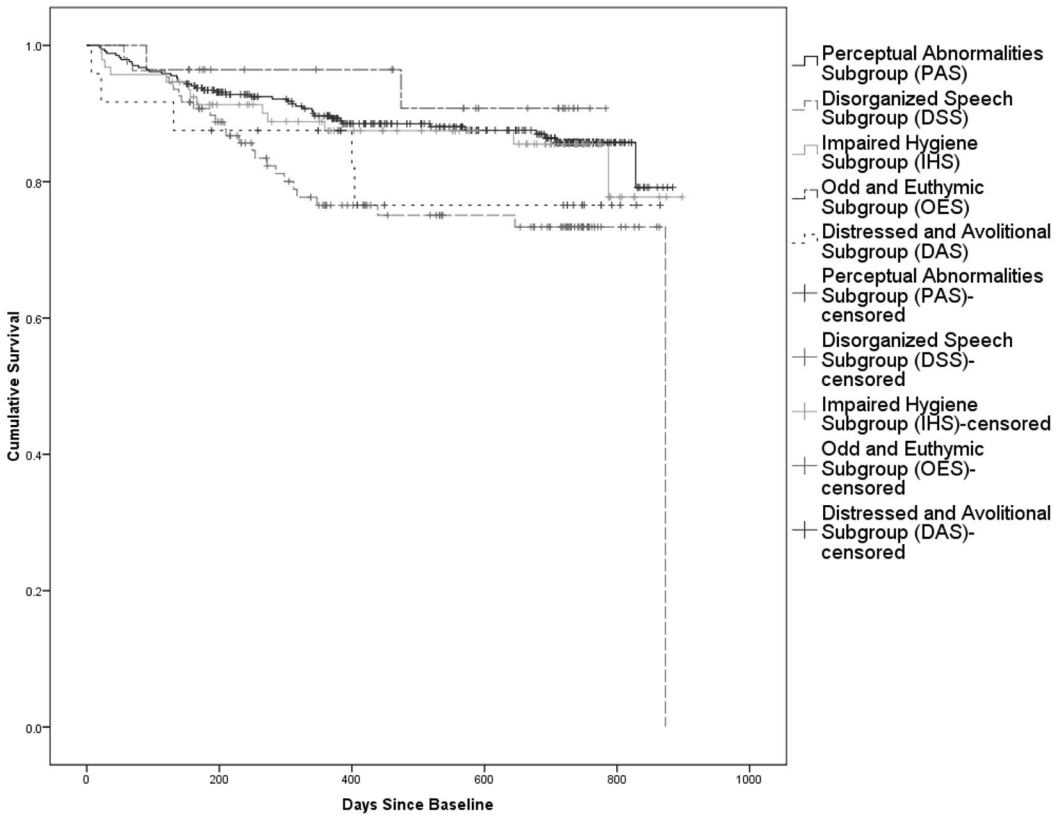


**Figure 3. NAPLS 2 Kaplan-Meir Survival Curve**

This figure graphs the Kaplan-Meir survival curve for conversion to psychosis among the subgroups identified in the NAPLS 2 latent class cluster analysis.

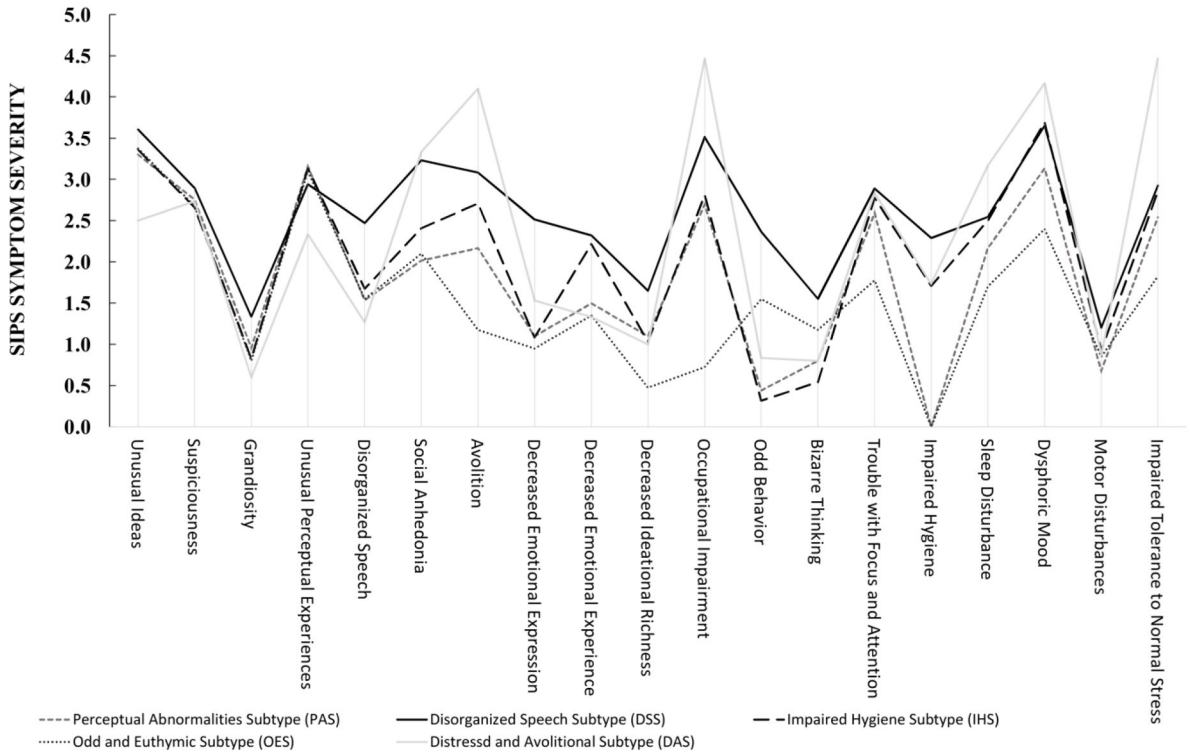


**Figure 4. NAPLS 1 LCCA-Derived Subgroup Symptom Means**

This line graph shows the mean symptom ratings for each of the subgroups identified in the NAPLS 1 latent class cluster analysis.

**Table 1**

Comparison of SIPS and CAARMS Symptom Scales

| Scales with a Close Counterpart | | Scales Whose Content is Divided Differently or With Only an Approximate Counterpart | | Scales Without a Clear Counterpart | |
|---|---|---|---|---|---|
| **SIPS** | **CAARMS** | **SIPS** | **CAARMS** | **SIPS** | **CAARMS** |
| • Perceptual Abnormalities | • Perceptual Abnormalities | • Unusual Ideas<br>• Suspiciousness | • Unusual Thought Content | • Impairment in Personal Hygiene | |
| • Disorganized Speech | • Disorganized Speech | • Trouble with Focus and Attention | • Subjective Cognitive Changes<br>• Observed Cognitive Changes | • Bizarre Thinking | |
| • Occupational Functioning | • Impaired role function | • Decreased Expression of Emotion | • Observed Blunted Affect | • Sleep Disturbance | |
| • Impaired Tolerance to Normal Stress | • Tolerance to normal stress | • Decreased Ideational Richness | • Alogia | | • Aggression |
| • Avolition | • Avolition/apathy | • Social Anhedonia | • Social Isolation | | • Subjective Bodily Sensations |
| | | • Decreased Experience of Emotion | • Anhedonia<br>• Subjective Emotional Disturbances<br>• Dissociative Symptoms | | • Subjective Autonomic Functioning |
| | | • Motor Disturbance | • Observed Motor Functioning<br>• Subjective Motor Functioning | | • Mood swings |
| | | • Grandiosity | • Mania | | |

| Scales with a Close Counterpart | | Scales Whose Content is Divided Differently or With Only an Approximate Counterpart | | Scales Without a Clear Counterpart | |
|---|---|---|---|---|---|
| **SIPS** | **CAARMS** | **SIPS** | **CAARMS** | **SIPS** | **CAARMS** |
| | | • Dysphoric Mood | • Depression<br>• Suicidality and Self-Harm<br>• Anxiety<br>• OCD | | |
| | | • Odd behavior or appearance | • Disorganized, Odd, Stigmatizing Behavior<br>• Observed inappropriate affect | | |

Note: No official method exists for linking the SIPS and CAARMS scales. This list of proposed counterparts is subjective and for illustrative purposes.

**Table 2**

Demographics of NAPLS 1 and NAPLS 2 Datasets

| | NAPLS I | NAPLS II |
|---|---|---|
| N | 356 | 737 |
| Age (SD) | 18.3 (4.7) | 18.5 (4.3) |
| Father's Education (SD) | 6.3 (1.7) | 6.2 (1.7) |
| Mother's Education (SD) | 6.4 (1.6) | 6.3 (1.6) |
| % Latino | 14% | 18% |
| % Female | 38% | 43% |
| % Native American and First Nations *** | 0% | 2% † |
| % Asian *** | 6% | 8% |
| % Black *** | 13% | 16% |
| % Other *** | NA # | 5% |
| % White *** | 71% † | 57% |
| % Native Hawaiian and Pacific Islander *** | 0% | 0% |
| % Multiracial *** | 9% | 12% † |
| Converted | 12% | 13% |
| APS ** | 96% † | 92% |
| GRD | 11% | 11% |
| BIPS | 3% | 3% |
| YSPD | NA # | 9% |

** = $p < .01$,

*** = $p < .001$,

† = value significantly larger in group contrast

Note: APS, GRD, and BIPS do not add up to 100% as an individual may belong to multiple groups.

# = "Other" was not an option for racial identity in NAPLS I and YSPD was not assessed as a diagnostic category in NAPLS 1. APS = Attenuated Psychosis Syndrome; GRD = Genetic Risk and Deterioration Syndrome; BIPS = Brief Intermittent Psychosis Syndrome; YSPD = Youth and Schizotypal Personality Disorder Syndrome

## Table 3

Demographics and Clinical Features by Subgroup for the LCCA of the NAPLS 1 Dataset

| | Perceptual Abnormalities Subgroup (A) | Disorganized Speech Subgroup (B) | Impaired Hygiene Subgroup (C) | Total |
|---|---|---|---|---|
| N (% of total) | 208 (58%) | 108 (30%) | 40 (11%) | 356 |
| % Female * | 43%C | 34% | 23% | 38% |
| Age | 18.3 | 18.3 | 18.5 | 18.3 |
| Father's Education (SD)† | 4.7 | 4.5 | 5.1 | 4.7 (2.0) |
| Mother's Education (SD)† | 4.8 | 4.6 | 4.9 | 4.8 (1.9) |
| % Latino | 14% | 14% | 15% | 14% |
| % Native American and First Nations | 0% | 0% | 0% | 0% |
| % Asian | 5% | 4% | 3% | 4% |
| % Black | 12% | 8% | 0% | 9% |
| % White | 74% | 80% | 90% | 77% |
| % Native Hawaiian and Pacific Islander | 0% | 0% | 0% | 0% |
| % Multiracial | 9% | 8% | 8% | 9% |
| APS | 97% | 95% | 95% | 96% |
| GRD *** | 5% | 26%A | 11% | 12% |
| BIPS | 3% | 3% | 5% | 3% |
| Conversion * | 11% | 17%A | 8% | 12% |

* = $p < .05$,

*** = $p < .001$

† = Absolute values of father's and mother's education reported here differ from those shown in Table 2 as the Table 2 values were converted to the education score metric used in NAPLS 2.

Note: Letters next to a value indicate the subgroups which that value was significantly larger than according to post hoc uncorrected pairwise Z or T tests. Pairwise comparisons were only conducted when the omnibus ANOVA, $\chi^2$, or Kaplan-Meir test was significant.

APS = Attenuated Psychosis Syndrome; GRD = Genetic Risk and Deterioration Syndrome; BIPS = Brief Intermittent Psychotic Symptoms Syndrome

**Table 4**

ANOVA of the LCCA-Derived NAPLS 1 Subgroups Symptom Ratings

| | PAS Mean (SD) (A) | DSS Mean (SD) (B) | IHS Mean (SD) (C) | Total Mean (SD) | P | $\eta^2$ |
|---|---|---|---|---|---|---|
| Unusual Ideas | 3.1 (1.5) | 3.2 (1.6) | 3.4 (1.4) | 3.2 (1.5) | =.61 | 0.00 |
| Suspiciousness | 2.9 (1.5)[C] | 3.1 (1.4)[C] | 1.9 (1.5) | 2.8 (1.5) | <.001 | 0.05 |
| Grandiosity | 1 (1.3) | 1.2 (1.4) | 1.7 (1.6)[A] | 1.1 (1.4) | =.01 | 0.02 |
| Unusual Perceptual Experiences | 3.1 (1.7)[C] | 2.9 (1.6) | 2.4 (1.8) | 2.9 (1.7) | <.05 | 0.02 |
| Disorganized Speech | 1.5 (1.4) | 2.5 (1.4)[A C] | 1.8 (1.3) | 1.8 (1.4) | <.001 | 0.09 |
| Social Anhedonia | 1.9 (1.8) | 3.6 (1.6)[A C] | 2.1 (1.9) | 2.5 (1.9) | <.001 | 0.16 |
| Avolition | 1.8 (1.6) | 3.6 (1.3)[A C] | 1.9 (1.7) | 2.3 (1.7) | <.001 | 0.24 |
| Decreased Emotional Expression | 1.1 (1.4)[C] | 2.5 (1.6)[A C] | 0.4 (0.7) | 1.4 (1.6) | <.001 | 0.22 |
| Decreased Emotional Experience | 1.2 (1.6)[C] | 2.6 (1.6)[A C] | 0.1 (0.3) | 1.5 (1.7) | <.001 | 0.22 |
| Decreased Ideational Richness | 1 (1.4)[C] | 2 (1.5)[A C] | 0.5 (0.8) | 1.3 (1.5) | <.001 | 0.12 |
| Occupational Functioning | 2.7 (1.9) | 3.9 (1.4)[A C] | 3.1 (2) | 3.1 (1.9) | <.001 | 0.09 |
| Odd Behavior | 1 (1.3) | 2.1 (1.5)[A] | 1.8 (1.3)[A] | 1.4 (1.4) | <.001 | 0.13 |
| Bizarre Thinking | 1.5 (1.5) | 2.1 (1.6)[A] | 1.9 (1.3) | 1.7 (1.5) | =.001 | 0.04 |
| Trouble with Focus and Attention | 2.2 (1.4) | 3.1 (1.1)[A C] | 2.2 (1.3) | 2.5 (1.4) | <.001 | 0.09 |
| Impairment in Personal Hygiene | 0 (0) | 2.2 (1.1)[A C] | 1.8 (1.2)[A] | 0.9 (1.3) | <.001 | 0.67 |
| Sleep Disturbance | 1.9 (1.7) | 2.4 (1.7)[A] | 1.9 (1.6) | 2 (1.7) | =.02 | 0.02 |
| Dysphoric Mood | 2.9 (1.7) | 3.7 (1.5)[A C] | 2.5 (1.9) | 3.1 (1.7) | <.001 | 0.06 |
| Motor Disturbances | 0.5 (1) | 1.2 (1.4)[A C] | 0.6 (1.3) | 0.7 (1.2) | <.001 | 0.06 |
| Impaired Tolerance to Normal Stress | 1.9 (1.7) | 3 (1.7)[A C] | 1.5 (1.4) | 2.2 (1.7) | <.001 | 0.09 |

Note: Grey lines separate positive domain scores, negative domain scores, disorganized domain scores, and general domain scores respectively. Letters indicate the other subgroups that were lower on the indicated symptom. Pairwise comparisons were only conducted when the omnibus ANOVA was significant.

PAS = Perceptual Abnormalities Subgroup; DSS = Disorganized Speech Subgroup; IHS = Impaired Hygiene Subgroup

**Table 5**

Demographics by Subgroup for the LCCA of the NAPLS 2 Dataset

| | Perceptual Abnormalities Subgroup (A) | Disorganized Speech Subgroup (B) | Impaired Hygiene Subgroup (C) | Odd and Euthymic Subgroup (D) | Distressed and Avolitional Subgroup (E) | Total |
|---|---|---|---|---|---|---|
| N (% of total) | 413 (56%) | 134 (18%) | 120 (16%) | 40 (5%) | 30 (4%) | 737 |
| % Female | 43% | 38% | 43% | 58% | 40% | 43% |
| Age | 18.6 | 18.4 | 18.2 | 18.3 | 20.5 | 18.5 (4.3) |
| Father's Education (SD) | 6.2 | 6.3 | 6.3 | 6.4 | 6.1 | 6.2 (1.7) |
| Mother's Education (SD) * | 6.1 | 6.8$^{A}$ | 6.4 | 6.7$^{A}$ | 6.2 | 6.3 (1.6) |
| % Latino | 22% | 15% | 13% | 18% | 13% | 18% |
| % Native American and First Nations | 2% | 3% | 2% | 0% | 0% | 2% |
| % Asian | 7% | 9% | 6% | 8% | 13% | 7% |
| % Black | 17% | 13% | 13% | 13% | 13% | 15% |
| % Other | 5% | 4% | 6% | 3% | 7% | 5% |
| % White | 55% | 58% | 64% | 67% | 60% | 58% |
| % Native Hawaiian and Pacific Islander | 0% | 1% | 0% | 0% | 0% | 0% |
| % Multiracial | 14% | 13% | 10% | 10% | 7% | 13% |
| APS ** | 92%$^{E}$ | 90%$^{E}$ | 98%$^{ABDE}$ | 90% | 77% | 92% |
| GRD *** | 10% | 16%$^{C}$ | 7% | 5% | 33%$^{ABCD}$ | 11% |
| BIPS | 2% | 3% | 2% | 5% | 7% | 3% |
| YSPD *** | 6% | 20%$^{ACE}$ | 8% | 13% | 3% | 9% |
| Lifetime Mood-stabilizer *** | 6% | 12%$^{A}$ | 14%$^{A}$ | 27%$^{AB}$ | 10% | 9% |
| Conversion * | 11% | 20%$^{ACD}$ | 11% | 5% | 17% | 13% |

* = $p < .05$,

** = $p < .01$,

*** = $p < .001$

Note: Letters next to a value indicate the other subgroups which that value was significantly larger than according to post hoc uncorrected pairwise Z or T tests. Pairwise comparisons were only conducted when the omnibus ANOVA, $\chi^2$, or Kaplan-Meir tests was significant.

APS = Attenuated Psychosis Syndrome; GRD = Genetic Risk and Deterioration Syndrome; BIPS = Brief Intermittent Psychosis Syndrome; YSPD = Youth and Schizotypal Personality Disorder Syndrome

**Table 6**

ANOVA of the LCCA Derived NAPLS 2 Subgroup Symptom Ratings

| | PAS (SD) (A) | DSS (SD) (B) | IHS (SD) (C) | OES (SD) (D) | DAS (SD) (E) | Total | P | Eta 2 |
|---|---|---|---|---|---|---|---|---|
| Unusual Ideas | 3.3 (1.3)$^{E}$ | 3.6 (1.1)$^{A\ E}$ | 3.4 (1.4)$^{E}$ | 3.4 (1.2)$^{E}$ | 2.5 (2) | 3.3 (1.3) | =.001 | .02 |
| Suspiciousness | 2.8 (1.6) | 2.9 (1.4) | 2.7 (1.4) | 2.7 (1.6) | 2.7 (1.7) | 2.8 (1.5) | =.78 | .00 |
| Grandiosity | 1 (1.3) | 1.3 (1.5)$^{A\ C\ D\ E}$ | 0.8 (1.2) | 0.8 (1) | 0.6 (1.2) | 1 (1.3) | =.004 | .02 |
| Perceptual Abnormalities | 3.2 (1.5)$^{E}$ | 2.9 (1.5)$^{E}$ | 3.1 (1.5)$^{E}$ | 3.1 (1.4)$^{E}$ | 2.3 (1.6) | 3.1 (1.5) | <.05 | .01 |
| Disorganized Speech | 1.5 (1.4) | 2.5 (1.5)$^{A\ C\ D\ E}$ | 1.7 (1.3) | 1.6 (1.1) | 1.3 (1.4) | 1.7 (1.5) | <.001 | .06 |
| Social Anhedonia | 2 (1.7) | 3.2 (1.7)$^{A\ C\ D}$ | 2.4 (1.7)$^{A}$ | 2.1 (1.4) | 3.3 (1.5)$^{A\ C\ D}$ | 2.4 (1.7) | <.001 | .08 |
| Avolition | 2.2 (1.6)$^{D}$ | 3.1 (1.5)$^{A\ D}$ | 2.7 (1.5)$^{A\ D}$ | 1.2 (1.1) | 4.1 (1.1)$^{A\ B\ C\ D}$ | 2.4 (1.6) | <.001 | .12 |
| Dcr Expression of Emotion | 1.1 (1.4) | 2.5 (1.7)$^{A\ C\ D\ E}$ | 1.1 (1.3) | 1 (1.2) | 1.5 (1.4) | 1.4 (1.5) | <.001 | .13 |
| Dcr Experience of Emotion | 1.5 (1.6) | 2.3 (1.7)$^{A\ D\ E}$ | 2.2 (1.8)$^{A\ D\ E}$ | 1.4 (1.2) | 1.3 (1.4) | 1.7 (1.7) | <.001 | .05 |
| Dcr Ideational Richness | 1.1 (1.3)$^{D}$ | 1.6 (1.5)$^{A\ C\ D\ E}$ | 1 (1.2)$^{D}$ | 0.5 (0.7) | 1 (1.4) | 1.2 (1.3) | <.001 | .04 |
| Occupational Functioning | 2.7 (2)$^{D}$ | 3.5 (1.8)$^{A\ C\ D}$ | 2.8 (2)$^{D}$ | 0.7 (1) | 4.5 (1.6)$^{A\ B\ C\ D}$ | 2.8 (2) | <.001 | .11 |
| Odd behavior/appearance | 0.4 (0.8) | 2.4 (1.3)$^{A\ C\ D\ E}$ | 0.3 (0.6) | 1.6 (1.4)$^{A\ C\ E}$ | 0.8 (1.1)$^{A\ C}$ | 0.8 (1.2) | <.001 | .40 |
| Bizarre Thinking | 0.8 (1.1)$^{C}$ | 1.6 (1.4)$^{A\ C\ E}$ | 0.5 (0.9) | 1.2 (1.3)$^{C}$ | 0.8 (0.9) | 0.9 (1.2) | <.001 | .07 |
| Trouble Focus/Attention | 2.6 (1.3)$^{D}$ | 2.9 (1.4)$^{A\ D}$ | 2.8 (1.2)$^{D}$ | 1.8 (1.2) | 2.8 (1)$^{D}$ | 2.6 (1.3) | <.001 | .03 |
| Impaired Personal Hygiene | 0 (0) | 2.3 (1.3)$^{A\ C\ D\ E}$ | 1.7 (0.9)$^{A\ D}$ | 0 (0) | 1.7 (1)$^{A\ D}$ | 0.8 (1.2) | <.001 | .66 |
| Sleep Disturbance | 2.2 (1.6) | 2.5 (1.4)$^{A\ D}$ | 2.5 (1.5)$^{A\ D}$ | 1.7 (1.5) | 3.2 (1.2)$^{A\ B\ C\ D}$ | 2.3 (1.6) | <.001 | .03 |
| Dysphoric Mood | 3.1 (1.7)$^{D}$ | 3.6 (1.4)$^{A\ D}$ | 3.7 (1.4)$^{A\ D}$ | 2.4 (1.6) | 4.2 (1.3)$^{A\ D}$ | 3.3 (1.6) | <.001 | .05 |
| Motor Disturbance | 0.7 (1) | 1.2 (1.3)$^{A\ C}$ | 0.9 (1)$^{A}$ | 0.9 (1) | 0.9 (1) | 0.8 (1.1) | <.001 | .04 |
| Imp Tolr to Stress | 2.5 (1.9)$^{D}$ | 2.9 (1.9)$^{A\ D}$ | 2.9 (1.9)$^{D}$ | 1.8 (1.7) | 4.5 (1)$^{A\ B\ C\ D}$ | 2.7 (1.9) | <.001 | .06 |

Note: Grey lines separate positive domain scores, negative domain scores, disorganized domain scores, and general domain scores respectively. Letters indicate the subgroups that were lower on the respective symptom. Pairwise comparisons were only conducted when the omnibus ANOVA was significant.

PAS = Perceptual Abnormalities Subgroup; DSS = Disorganized Speech Subgroup; IHS = Impaired Hygiene Subgroup; OES = Odd and Euthymic Subgroup; DAS = Distressed and Avolitional Subgroup

**Table 7**

Clinical Symptom Ratings that were Most Discriminative of CHR-P Subgroups Across LCCA

| | NAPLS 1 | NAPLS 2 | Healey et al. (in Press) | Valmaggia et al. (2013) |
|---|---|---|---|---|
| 1 | Imp Personal Hygiene (0.67) | Imp Personal Hygiene (0.66) | Dysphoric mood (0.44) | Subjective Motor Functioning (0.94) |
| 2 | Avolition (0.24) | Odd Behavior/Appearance (0.4) | Avolition (0.35) | Avolition/Apathy (0.33) |
| 3 | Dcr Exprsn of Emotion (0.22) | Dcr Exprsn of Emotion (0.13) | Imp Tolerance Nrm Stress (0.3) | Anhedonia [Dcr Em Exprnc] (0.32) |
| 4 | Dcr Emotional Experience (0.22) | Avolition (0.12) | Dcr Ideational Richness (0.29) | Social Isolation [Social Anh] (0.32) |
| 5 | Social Anhedonia (0.16) | Occupational Functioning (0.11) | Sleep Disturbance (0.29) | Sbj Cog Chg [T w/Foc Atn] (0.28) |
| 6 | Odd Behavior/Appearance (0.13) | Social Anhedonia (0.08) | Occupational Functioning (0.25) | Subj Autonomic functioning (0.27) |
| 7 | Dcr Ideational Richness (0.12) | Bizarre Thinking (0.07) | Trbl w/Focus & Attn (0.23) | Alogia [Dcr Id Richness*] (0.24) |
| 8 | Trbl w/Focus & Attn (0.09) | Disorganized Speech (0.06) | Odd Behavior or Appearance (0.2) | Impaired Role Function (0.23) |
| 9 | Imp Tolerance Nrm Stress (0.09) | Imp Tolerance Nrm Stress (0.06) | Social Anhedonia (0.16) | Depression (0.16) |
| 10 | Disorganized Speech (0.09) | Dysphoric Mood (0.05) | Dcr Exp of Emotion (0.16) | Tolerance to Normal Stress (0.16) |
| 11 | Occupational Functioning (0.09) | Dcr Experience of Emotion (0.05) | Dcr Experience of Emotion (0.15) | Anxiety (0.15) |
| 12 | Motor Disturbances (0.06) | Dcr Ideational Richness (0.04) | Imp Personal Hygiene (0.09) | Disorganized Speech (0.12) |
| 13 | Dysphoric Mood (0.06) | Motor Disturbance (0.04) | Suspiciousness (0.08) | Subj Emotional Disturbances (0.12) |
| 14 | Suspiciousness (0.05) | Trbl w/Focus & Attn (0.03) | Motor Disturbance (0.08) | Observed Blunted Affect (0.11) |
| 15 | Bizarre Thinking (0.04) | Sleep Disturbance (0.03) | Perceptual Abnormalities (0.07) | Unusual Thought Content (0.1) |
| 16 | Grandiosity (0.02) | Unusual Ideas (0.02) | Disorganized Speech (0.06) | Aggression (0.1) |
| 17 | Sleep Disturbance (0.02) | Grandiosity (0.02) | Unusual Ideas (0.05) | Subjective Bodily Sensations (0.1) |
| 18 | Perceptual Abnormalities (0.02) | Perceptual Abnormalities (0.01) | Bizarre Thinking (0.02) | OCD (0.08) |
| 19 | Unusual Ideas (0) | Suspiciousness (0) | Grandiosity (0) | Dissociative Symptoms (0.06) |
| 20 | | | | Objective Cognitive Changes (0.05) |
| 21 | | | | Disorg, Odd, Stigm Bhvr (0.05) |
| 22 | | | | Observed Inappropriate Affect (0.04) |
| 23 | | | | Mood Swings (0.04) |
| 24 | | | | Mania (0.03) |
| 25 | | | | Suicidality and Self-harm (0.03) |
| 26 | | | | Perceptual Abnormalities (0) |

Note: Numbers after NAPLS symptoms are $\eta^2$ values. Numbers after Healey et al. and Valmaggia et al. are $R^2$. Symptom names in brackets under Valmaggia et al. (2013) are SIPS symptom ratings that were considered equivalent for this comparison. Highlighting is used to emphasize symptoms that repeated across lists