

ARTICLE OPEN



Identifying individuals at clinical high risk for psychosis using a battery of tasks sensitive to symptom mechanisms

Trevor F. Williams¹✉, James M. Gold², James A. Waltz², Jason Schiffman³, Lauren M. Ellman⁴, Gregory P. Strauss⁵, Elaine F. Walker⁶, Scott W. Woods⁷, Albert R. Powers⁷, Joshua Kenney⁷, Minerva K. Pappu⁷, Philip R. Corlett⁷, Tanya Tran⁸, Steven M. Silverstein⁸, Richard E. Zinbarg⁹ and Vijay A. Mittal¹⁰

© The Author(s) 2025

The clinical high risk for psychosis (CHR-P) population is important for understanding disease progression and treatment; however, standard approaches to identifying CHR-P individuals are expensive and labor-intensive. Focusing on neurocognitive mechanisms that underlie individual psychosis symptoms (positive, negative, and disorganization) may improve screening and identification. The present study examines whether a behavioral task battery that assays symptom mechanisms can identify CHR-P individuals and predict risk severity. Participants ($N = 621$) were recruited from clinics and the community as part of the Computerized Assessment of Psychosis Risk (CAPR) consortium study. Structured clinical interviews, a dimensional risk calculator, and behavioral tasks were administered. Clinical interviews identified the following groups: (a) CHR-P ($n = 273$), (b) non-CHR-P individuals with limited psychosis like experiences (PLEs; $n = 120$), (c) participants with mental disorders and no PLEs (CLN; $n = 82$), and (d) healthy controls (HC; $n = 146$). Multinomial logistic regression indicated that the task battery differentiated groups ($p < 0.001$), with utility for identifying CHR-P individuals (Sensitivity = 0.87, PPV = 0.51, NPV = 0.77), though with high false positives that varied based on comparison group (Specificity = 0.21–0.43). Tasks also predicted psychosis risk calculator scores (Adjusted $R^2 = 0.12$), with the two unique predictors being positive symptom task variables associated with updating beliefs regarding environmental volatility. Overall, symptom mechanism tasks differentiated CHR-P individuals from control groups, suggesting their potential as novel screening tools. Using tasks to more efficiently identify CHR-P individuals (e.g., enrich samples), may lower barriers and identify individuals that may otherwise be missed.

Translational Psychiatry (2025)15:311 ; <https://doi.org/10.1038/s41398-025-03539-5>

INTRODUCTION

Psychosis spectrum disorders (e.g., schizophrenia) are a major source of disability worldwide and cause considerable personal suffering [1]. Recently, increased emphasis has been placed on early identification, prevention, and mechanism research, in hopes of better understanding and mitigating the impacts of these disorders [2]. Research on individuals at clinical high risk for psychosis (CHR-P) has been at the forefront of psychosis risk research; however, these individuals are challenging to identify. Complicating matters, other clinical populations have overlapping symptoms and some individuals report psychosis-like experiences (PLEs) of limited frequency, intensity, and impact, such that they are not considered to have “at-risk” status. The present study tests a novel approach to identifying CHR-P individuals: a behavioral task battery that assesses underlying symptom mechanisms.

CHR-P individuals experience attenuated positive symptoms of psychosis (e.g., seeing human-like shadows), brief full positive symptoms, or genetic risk for psychosis combined with functional decline [2]. A recent meta-analysis estimated that 25% of these individuals later develop a psychosis spectrum disorder [3] and CHR-P individuals who do not often still experience significant symptoms and impairment [4]. Given that these participants often are unmedicated and have not developed full symptomatology, researching and intervening on CHR-P may advance our understanding of symptom progression, decrease duration of untreated psychosis, and provide a window for preventive treatments [2].

One challenge to CHR-P research is that it is resource-intensive and costly to recruit participants. Structured clinical interviews are the gold-standard approach to CHR-P identification, but take 1.5–2 h to administer and require considerable training [5]. Self-report questionnaires are another screening tool, but are limited

¹Department of Psychology, Kent State University, Kent, OH 44242, USA. ²Maryland Psychiatric Research Center, Department of Psychiatry, University of Maryland School of Medicine, Baltimore, MD 21228, USA. ³Department of Psychological Science, 4201 Social and Behavioral Sciences Gateway, University of California, Irvine, CA 92697, USA. ⁴Department of Psychology & Neuroscience, Temple University, Philadelphia, PA 19122, USA. ⁵Departments of Psychology and Neuroscience, University of Georgia, Athens, GA 30602, USA. ⁶Department of Psychology and Program in Neuroscience, Emory University, Atlanta, GA 30322, USA. ⁷Department of Psychiatry, Yale University, New Haven, CT 06519, USA. ⁸Departments of Psychiatry, Neuroscience and Ophthalmology, University of Rochester Medical Center, Rochester, NY 14642, USA. ⁹Department of Psychology, Northwestern University, Evanston, IL 60208, USA. ¹⁰Institutes for Policy Research (IPR) and Innovations in Developmental Sciences (DevSci), Departments of Psychology, Psychiatry, Medical Social Sciences, Northwestern University, Evanston, IL 60208, USA. Previous Presentation: this work has not been presented or published.

✉email: twill253@kent.edu

Received: 5 May 2025 Revised: 17 July 2025 Accepted: 11 August 2025

Published online: 23 August 2025

in this population by poor insight and guardedness [6]. The present study builds upon advances in clinical cognitive neuroscience and computational psychiatry to consider a new approach: behavioral tasks focused on symptom-specific mechanisms.

Across the psychosis spectrum, progress has been made identifying cognitive mechanisms that underlie specific symptoms [7]. This literature is defined by its emphasis on individual symptoms, over diagnostic constructs, and integrating sophisticated behavioral tasks with neural markers. For instance, research on delusions has indicated that aberrant prediction error signaling may lead to the development of spurious beliefs or a sense of unpredictability, which form the seeds of delusions [8, 9]. Similarly, hallucination research suggests that the overweighting of prior beliefs (e.g., perceptual expectations) may influence the processing of sensory information, such that aberrant patterns are formed [10]. Negative symptoms have been connected to reward processing and response initiation, observed in how individuals with negative symptoms represent and update beliefs about expected rewards and the degree of effort required to obtain them [11, 12]. Finally, disorganization symptoms have been linked to reductions in context processing, perceptual organization, and motor coordination [13].

Symptom mechanism tasks have considerable promise for identifying CHR-P individuals, in that they capture the insights of neuroscience research, without requiring resource-intensive neural data. Despite this, these tasks have yet to impact CHR-P identification for several reasons. First, symptom-specific tasks have often been studied in isolation to each other, limiting their ability to identify CHR-P individuals, who experience varied symptoms. It is necessary to examine their collective predictive power and potential as a testing battery. A second limitation, which extends beyond behavioral task research, is the limited consideration of clinical controls [14]. Although demonstrating the ability of a measure to differentiate CHR-P from healthy individuals is useful for community screening, the reality is that much screening is done in help-seeking populations [15]. A stronger test of CHR-P identification would thus involve comparisons to groups with varied clinical symptoms and especially to individuals with PLEs that fall short of meeting CHR-P criteria.

The present study represents the first step in a broader program of research aiming to fundamentally reshape psychosis risk research, using symptom-specific mechanisms [16]. Specifically, this study examined behavioral tasks that assess symptom mechanisms as predictors of CHR-P status in a large sample. Notably, this included not only differentiating CHR-P individuals from healthy controls, but also from individuals with lower-risk PLEs and other clinical diagnoses. It was hypothesized that the task battery would significantly differentiate CHR-P individuals from all three groups; however, because these tasks are so often studied individually and without clinical controls, specific predictions were not made regarding which tasks would be predictive. Finally, relations to a conversion risk calculator were examined within CHR-P individuals, as such risk calculators are predictive of development of a psychosis spectrum disorder in this population and have recently been used to enrich CHR-P samples [17].

METHODS

Data were baseline assessments from the Computerized Assessment of Psychosis Risk (CAPR) study, a multisite study of psychosis risk (see published protocol [16]). Participants ($N = 621$) completed assessments between January 2020 and July 2023, which included clinical interviews, questionnaires, and a behavioral tasks. CHR-P ($n = 273$) participants were included based on meeting criteria for a psychosis risk syndrome. Three control groups were included, consisting of individuals experiencing psychosis like experiences (PLEs) that did not meet CHR-P criteria (PLE group; $n = 120$), individuals without PLEs but with other current or recent clinical diagnoses (e.g., Major Depressive Disorder; CLN group; $n = 82$), and

healthy controls without PLEs or psychopathology symptoms (HC group; $n = 146$). The overall sample size was considered to provide adequate power for multinomial logistic regression, based on common guidelines [18]. Exclusion criteria for all participants were severe head injury, the presence of a neurological disorder, and lifetime history of a psychotic disorder. The study was approved by the Northwestern University IRB (STU00211351) and acknowledged by the IRBs of participating sites. All adult participants provided informed consent after receiving a complete description of the study, whereas minors provided written assent and their parents or guardians provided written consent.

Interviews were conducted by formally trained postdoctoral fellows, graduate students, and research assistants. Behavioral tasks were administered online by research assistants trained to explain task instructions and monitor participant performance for data quality. Participants self-reported on demographic characteristics in a format consistent with the United States National Institutes of Health reporting requirements.

Clinical Measures

The Structured Interview of Psychosis Risk Symptoms (SIPS) is a semi-structured interview that identifies psychosis risk syndromes [5]. Positive symptoms are central in identifying psychosis risk syndromes and include unusual/delusional thoughts, suspiciousness/persecutory ideas, grandiosity, perceptual abnormalities, and disorganized communication.

The Structured Clinical Interview for DSM-5, Research Version (SCID) [19] was used to determine the presence of other mental disorders and thus inclusion status for CLN and HC groups. Interviewers administered psychotic, bipolar, depressive, anxiety, obsessive-compulsive, trauma-related, eating, and substance use disorder modules.

The North American Prodrome Longitudinal Study-Risk Calculator (NAPLS-RC) [20] was developed in a large longitudinal study of CHR-P individuals to predict the probability of transitioning to a psychosis spectrum disorder during a 2-year timespan. The NAPLS-RC uses SIPS ratings and participant age, as well as several other measures that were administered as part of the CAPR study: the Brief Assessment of Cognition in Schizophrenia [21], the Hopkins Verbal Learning Test-Revised [22], negative life events scores from the Research Interview Life Events Scales [23], and social functioning decline as measured by the Global Functioning Scale [24]. In the present study, it was used as a dimensional assessment of psychosis risk within CHR-P individuals.

Behavioral tasks

Eleven behavioral tasks were administered by trained research personnel that monitored task engagement and performance. They were presented in randomized order, with each contributing one or two predictor variables to analyses. These tasks and indices are briefly reviewed here; more information can be found in the CAPR protocol paper and the supplement to this article [16]. Descriptive statistics and correlations with symptoms are provided in the supplement to this article.

Positive symptom mechanisms. The *Kamin Blocking (KB)* task assesses aberrant causal learning and produces scores for overall causal learning (control) and inappropriate learning of “blocked” cues (blocking) [8]. The *Probabilistic Reversal Learning (PRL)* task measures prior beliefs and learning about environmental volatility, with two variables: lose-stay rate (LSR) and win-switch rates (WSR) [9]. The *Sine Wave Speech (SWS)* task assesses influence of priors on auditory perception, with the main variables being increases in sensitivity (SWS d') across the task and early bias (SWS c) [25]. The *Conditioned Hallucinations (CH)* task measures prior influence (v) on perceptual judgments about the presence of an auditory stimulus presented concurrently with a visual stimulus; the probability of the target auditory stimulus changes over time and the participant’s learning about this change is captured by the volatility belief evolution rate (ω_3) [10].

Negative symptom mechanisms. The *Gain vs. Loss Avoidance Task (GLIAT)* task measures expected value estimation through two indices that capture enhanced learning about losses relative to rewards (learning, transfer) [11]. The *Effort Expenditure for Rewards (EEFRT)* task assesses over-estimation of effort costs through the tendency to choose high effort tasks that are coupled with high reward probability (EEFRT hard) [26]. The *Delay Discounting (DD)* task measures the valuation of future rewards, with the central outcome being large reward trial performance (DD large) [12]. The *Hedonic Reactivity (HR)* [27] task measures hedonic response to pleasant visual stimuli (HR positive). The *Computerized Finger Tapping (CFT)* task

assesses initiation and persistence of volitional movements, with average number of taps in a speeded condition being the central outcome (CFT speed) [28].

Disorganization symptom mechanisms. The *Ebbinghaus Illusion* (EI) task measures susceptibility to the classic illusion, producing a context sensitivity score [29]. The *Mooney Faces* (MF) task involves integrating degraded visual information into face percepts, both when faces are upright and inverted, producing separate scores for each condition [30].

Statistical analyses

Analyses were conducted in R [31] and the *mice* package was used for multiple imputation of missing data, with results pooled across 100 imputed datasets [32]. Multinomial logistic regression was used to predict group status, with CHR-P used as a reference group contrasted with each control group. This analysis produced three separate equations and coefficients for each predictor, such that they indicated differences between the reference group (CHR-P) and one of the control groups. An intercept model was estimated as a baseline, followed by adding potential demographic covariates (age, sex, race, and household income). Following this, a model with all task indices was tested. Model comparisons were made using the Akaike Information Criterion (AIC; lower is better) and the Wald test for model differences ($p < 0.05$). All p values reflect two-tailed tests.

The classification value of models was further examined through calculating sensitivity, specificity, positive predictive value (PPV), and negative predictive value (NPV) for identifying CHR-P individuals. The task battery's ability to predict risk severity within CHR-P individuals was examined through a multiple regression analysis in which the NAPLS-RC was the dependent variable and task variables were entered as predictors.

RESULTS

Participants ($N = 621$) were on average 23.59 years old ($SD = 4.15$), and racially diverse (see Table 1). There were no large differences between groups on demographic variables, though there was a small sex difference ($p = 0.04$). Most participants (55%) had no missing data and the average missingness across tasks varied (12–22%; see Table 2 descriptives). There were few extreme outliers across tasks (± 5 SD) and these values were marked missing. Correlations between tasks were small ($M = 0.01$, $SD = 0.10$), with the exception of the correlation between the inverted and upright Mooney Face task variables ($r = 0.68$, $p < 0.001$), suggesting the tasks provide information on unique and distinct mechanisms.

Multinomial logistic regression models were fit to examine the value of behavioral tasks for differentiating groups (see Table 3). An intercept model and one with demographic control variables were compared; the Wald test ($p = 0.39$) and AIC (Intercept, 1604 vs. Demographics, 1620) did not indicate differences, suggesting demographic variables did not differentiate groups. Next, the intercept model was compared to one with all tasks entered as predictors. Both the significant Wald test ($p < 0.001$) and difference in model AIC indicated that the task battery fit better than the intercept model. Further emphasizing the predictive value of the task battery, classification accuracy statistics indicated that the model correctly identified most CHR-P individuals (Sensitivity = 0.87) and that a model prediction of “not CHR-P” was relatively accurate (NPV = 0.77); the model had somewhat low specificity (0.33) and thus a propensity toward false positives, lowering the accuracy of model predicted CHR-P status (PPV = 0.51).

Next, the individual predictive equations that contrasted CHR-P and specific control groups were examined in terms of predicted probabilities (see Fig. 1), classification accuracy, and significant predictors (see Table 4). The model performed best at separating the CLN group from CHR-P individuals (e.g., PPV = 0.84), driven by a relatively lower rate of false positives in this group (Specificity = 0.43). The MF inverted condition variable was the only predictor that significantly ($OR = 0.49$; $p < 0.001$) differentiated groups, with the effect opposite of what is found in

schizophrenia, but consistent with a recent CHR-P study [30]; a model in which only this variable was entered did not change this result ($OR = 0.57$, $p < 0.001$), suggesting the effect was not driven by suppression. The model performed nearly as well at differentiating HC from CHR-P individuals (e.g., Specificity = 0.38, PPV = 0.72); however, there were a larger number of significant predictors ($p < 0.05$), which spanned symptom domains. Notably, the contribution of MF inverted paralleled the CLN model; however, there was also an effect in the opposite direction for KB control, which did appear to be a suppression effect (single predictor model $OR = 0.86$, $p = 0.16$). Finally, although the model still was accurate in identifying CHR-P individuals relative to the PLE group (Sensitivity = 0.87, PPV = 0.71), many PLE individuals were incorrectly identified as CHR-P (i.e., elevated false-positives; Specificity = 0.21, NPV = 0.42). MF inverted was again a significant predictor and EI context was a significant predictor, but similar to MF inverted it was predictive in a direction counter to theory and remained so when examined as a single predictor (e.g., $OR = 0.77$, $p = 0.024$).

Finally, the task battery's ability to predict probabilistic risk for developing a psychotic disorder (NAPLS-RC) was examined using linear multiple regression (Zero-order correlations are presented in the supplement to the article (Supplemental Tables 2 and 3)). As is standard, the NAPLS-RC was examined only within CHR-P individuals ($n = 273$) and the adjusted R^2 for this model was 0.12 (95% CI = [0.04, 0.22]), though a similar result also emerged in the full sample (adjusted $R^2 = 0.10$; see Supplement). Overall, the task battery was substantially related to psychosis risk, with two unique predictors emerging: PRL WSR ($\beta = 0.17$, $p = 0.023$) and CH ω^3 ($\beta = 0.14$, $p = 0.045$).

DISCUSSION

One challenge to psychosis risk research is identifying at-risk individuals, especially among help-seeking populations. The present study took a novel approach to this problem, examining whether symptom mechanism tasks can differentiate CHR-P individuals from control groups that mirror real-world recruitment populations. The task battery successfully classified CHR-P individuals and predicted risk severity, suggesting that this approach may complement existing methods through its distinct advantages.

The sensitivity of the task battery for identifying CHR-P was high (i.e., 87%); however, the battery's predictive accuracy (i.e., PPV, NPV) varied. The task battery was most accurate for differentiating CHR-P individuals from individuals with psychopathology and healthy individuals, but had more false positives and less accuracy in comparisons with individuals with PLEs. The sensitivity of the task battery minimally suggests utility for screening or “enriching” samples, reducing interviewing time and allowing collection of larger samples with high proportions of CHR-P individuals. Consistent with this, the battery's relation to the NAPLS-RC within CHR-P individuals suggests the task battery predicts risk severity. Nonetheless, the false positive rates indicated that interviewing is still needed to definitively identify cases, particularly when differentiating individuals with low risk PLEs (e.g., low frequency) from individuals that are truly at-risk. Notably, these findings compare favorably to studies using neuroimaging indicators to predict psychosis risk (e.g., 51% PPV) [33, 34].

Beyond demonstrating a task battery's potential utility, the present study may suggest how to optimize task batteries, through specific predictors. First, the Mooney Faces inverted variable was the strongest predictor of CHR-P status (A follow-up analysis removing MF variables from the model indicated that the model remained significantly predictive of group status, though was less accurate (see Supplemental Results.); however, disorganization mechanisms did not drive this effect. The Mooney Faces task requires individuals to detect faces in a visually degraded

Table 1. Demographics by clinical group.

	CHR-P	PLE	CLN	HC
<i>N</i>	273	120	82	146
<i>M</i> Age (<i>SD</i>)	23.40 (4.30)	23.70 (4.35)	23.95 (3.59)	23.58 (4.06)
Sex				
Female	67.30%	66.95%	61.73%	53.79%
Male	32.70%	33.05%	38.27%	46.21%
Race				
African American	17.05%	16.95%	13.58%	13.79%
American Indian ¹	0.76%	0.69%	0.00%	1.69%
Asian	17.80%	22.03%	22.22%	31.72%
Caucasian	51.52%	50.00%	53.09%	44.83%
Multiracial	10.61%	6.78%	9.88%	8.28%
Native Hawaiian or Other Pacific Islander ¹	0.76%	0.69%	0.00%	0.00%
Unknown or not reported ^a	1.52%	0.00%	1.23%	2.54%
Hispanic	14.45%	11.02%	12.35%	7.59%
<i>Mdn</i> Household Income	\$60,000.00	\$60,000.00	\$95,500.00	\$100,000.00
Current Medication				
Antidepressant	23.00%	24.00%	17.00%	8.60%
Antipsychotic	6.50%	0.91%	0.00%	0.00%
Mood stabilizer	6.20%	1.80%	1.30%	0.00%
Stimulant	6.20%	2.70%	1.30%	0.74%
Anxiolytic	7.30%	5.50%	4.00%	1.50%
Current Substance Use Disorder ^b	30.33%	21.98%	17.91%	00.00%

CHR-P clinical high risk for psychosis, PLE psychosis like experiences (not meeting CHR-P criteria), CLN current or recent mental health diagnoses, and HC healthy controls.

^adue to low prevalence, these groups were combined for analyses. Only a small significant difference between groups was present for sex ($p = 0.04$) and, as noted below, when demographic variables were considered as a whole there were no differences between groups.

^bBased on the presence of any current SCID Substance Use Disorder.

Table 2. Descriptive statistics for continuous predictors and outcomes.

	N	M	SD	Mdn	Min	Max	Skew	Kurtosis
NAPLS-RC	205	12.15	7.52	10.29	3.19	41.09	1.65	2.51
KB Blocking	484	−0.05	0.74	−0.18	−1.00	1.00	0.18	−1.48
KB Control	483	0.59	0.57	0.84	−1.00	1.00	−1.69	1.85
PRL WSR	524	−1.25	0.55	−1.29	−2.00	−0.01	0.26	−1.01
PRL LSR	524	0.24	0.21	0.18	0.00	1.00	0.82	−0.22
SWS d'	521	−0.08	0.86	−0.04	−2.46	2.69	0.02	0.18
SWS c	521	−0.29	0.62	−0.28	−2.01	2.01	0.14	0.34
CH ν (prior)	495	0.66	0.62	0.52	0.12	3.47	2.47	6.46
CH ω_3	493	−6.01	0.47	−6.00	−8.16	−3.54	0.70	10.95
HR positive	518	3.94	0.71	4.00	1.00	5.00	−1.04	1.54
EEFRT hard	514	0.55	0.26	0.59	0.00	1.00	−0.33	−0.59
DD large	518	−1.66	0.36	−1.86	−1.99	−0.59	1.29	0.66
CFT speed	522	55.11	8.94	54.33	16.33	92.67	0.30	2.67
GLIAT learning	547	0.02	0.17	0.02	−0.83	0.88	0.28	4.87
GLIAT transfer	547	0.54	0.26	0.54	0.00	1.00	−0.11	−0.68
EI context	520	73.06	36.81	80.00	−102.00	149.00	−1.83	5.18
MF inverted	491	32.38	21.76	30.23	0.00	93.02	0.61	−0.34
MF upright	489	78.14	13.04	81.40	27.91	100.00	−0.94	0.97

The NAPLS-RC was only calculated for CHR-P participants, given that its original validation was based on use in this population only. The PRL WSR and DD large were logarithmically transformed, consistent with previous work (see supplement). Descriptive statistics for each group are provided in Supplemental Table 1.

picture and, although challenging for individuals with schizophrenia, it may be that CHR-P individuals engage compensatory processes that over-integrate stimuli. The present finding is consistent with a recent CHR-P study [30] and may parallel the predictive coding mechanisms underlying positive symptoms [8, 25]. Consistent with this, positive symptom mechanism tasks also identified CHR-P individuals and the two predictors of risk severity (NAPLS-RC) were from positive symptom mechanism tasks. In particular, the variables related to the NAPLS-RC, captured beliefs about changing learning environments, with one reflecting perceived environmental volatility relevant to paranoia (PRL WSR) [9] and the other a failure to update perceptual beliefs as contingencies change (CH ω_3 ; e.g., persisting conditioned hallucinations) [10]. Belief updating has previously distinguished patients with schizophrenia from non-clinical voice-hearing controls [10], suggesting that these processes may represent important mechanisms underlying conversion to psychosis. Despite the emphasis on positive symptoms in the present findings, one negative symptom task showed relevance for predicting CHR-P status (CFT), suggesting value in continuing to explore other symptom mechanisms for identifying CHR-P individuals.

In considering the value of a task battery, it is worth contrasting it with existing methods. Focusing on the significant predictors of CHR-P status, tasks took 8–10 min to complete [16] and administering *all* such tasks would take 56 min, though further task and battery refinement may reduce administration time. Minimally trained research assistants administered tasks online; however, these tasks have also been successfully administered without supervision, suggesting that these tasks may offer a scalable and low-burden approach to screening [9]. In contrast, the SIPS and similar interviews take 1.5–2 h to administer [35], require intensive training, and ongoing clinical supervision. Although self-report measures provide another brief screening option, they require insight and self-disclosure that may not be present in this

population. Both interviews and self-reports may also be affected by symptom associated stigma, which limits disclosure [36, 37], and cultural variation in how symptoms are described [38, 39]. In contrast, behavioral tasks do not require insight or self-disclosure, and are less directly mediated by language. As such, symptom mechanism tasks provide an intriguing alternative to screening and identifying CHR-P individuals, representing a potential practical application of computational psychiatry [40, 41].

Limitations and future directions

There are several limitations to this study worth considering. First, the specificity of the battery for identifying CHR-P individuals was relatively low in the full sample and in particular when individuals with PLEs were considered. Nonetheless, performance was stronger in controls without PLEs and the current battery may still be useful as a pre-screening tool for sample enrichment and future work may refine the battery, improving specificity. For instance, each task had additional indices and many would be amenable to computational modeling, which may ultimately enhance prediction [9, 16]. Future work should optimize individual tasks for efficiently identifying CHR-P individuals. Second, although this study demonstrates the value of symptom mechanism tasks for screening and identifying psychosis risk, it does not present a ready-to-use tool. Further work is needed to hone a task battery and develop implementation infrastructure. Finally, the present study focused on cross-sectional data and future work would do well to include longitudinal data and outcomes (e.g., conversion to psychosis).

CONCLUSIONS

This study demonstrated the potential value of symptom mechanism tasks for identifying CHR-P individuals and assessing psychosis risk severity, with comparisons to control groups of varied composition. Positive symptom mechanisms were most effective for identifying CHR-P individuals. They were particularly accurate at differentiating CHR-P from individuals with other mental disorders and healthy individuals, but had higher false-positive rates when differentiating CHR-P from individuals with low-risk PLEs (less frequent, intense, etc.). Overall, these symptom mechanism tasks may have value for screening, oversampling, and identifying psychosis risk, without the resource costs of clinical interview or subjectivity of self-reports.

Table 3. Model fit and comparisons.

Model	AIC	LL	Wald <i>p</i> -value
(a) Intercept	1604	−799	
(b) Demographic	1620	−783	(a) vs. (b), <i>p</i> = 0.39
(c) Task Battery	1578	−735	(a) vs. (c), <i>p</i> < 0.001

Model fit indices and tests are based on pooled comparisons across 100 imputed datasets.

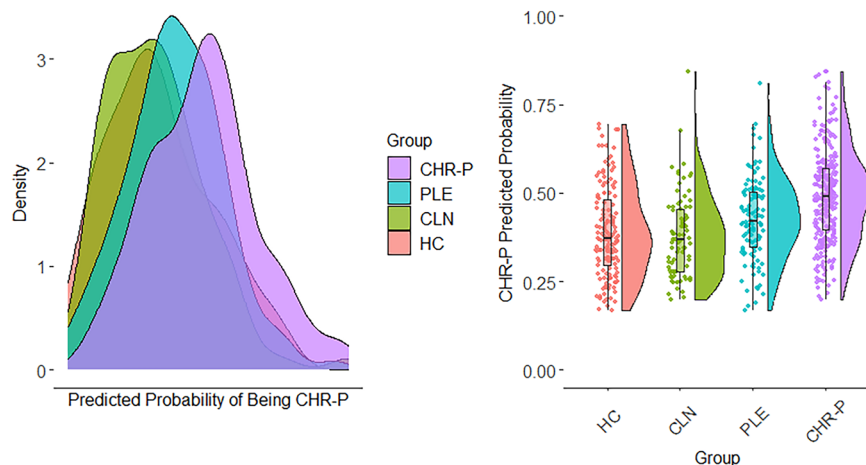


Fig. 1 Distribution of predicted probability of CHR-P status by group. CHR-P = clinical high risk for psychosis, PLE = psychosis like experiences that are not CHR-P, CLN = clinical disorders without PLEs, and HC = healthy controls.

Table 4. Significant predictors from the task battery multinomial regression.

Predictor	Estimate	SE	Z	p value	OR
<i>HC vs. CHR-P (SN = 0.87, SP = 0.38, PPV = 0.72, NPV = 0.61)</i>					
KB Control	−0.27	0.12	−2.20	0.029	0.76
PRL WSR	0.34	0.13	2.57	0.010	1.41
SWS c	0.28	0.14	2.05	0.041	1.33
CFT speed	0.25	0.13	2.00	0.046	1.29
MF inverted	−0.61	0.18	−3.33	0.001	0.54
<i>CLN vs. CHR-P (SN = 0.87, SP = 0.43, PPV = 0.84, NPV = 0.50)</i>					
MF inverted	−0.71	0.22	−3.20	0.001	0.49
<i>PLE vs. CHR-P (SN = 0.87, SP = 0.21, PPV = 0.71, NPV = 0.42)</i>					
EI context	−0.28	0.12	−2.31	0.019	0.75
MF inverted	−0.52	0.18	−2.80	0.005	0.59

All predictors were standardized and keyed in the theorized direction of pathology. Odds ratio (OR) values > 1.00 indicate a higher likelihood of being CHR-P.

SN sensitivity, SP specificity, PPV positive predictive value, NPV negative predictive value.

DATA AVAILABILITY

Deidentified data is available at the National Institute of Mental Health National Data Archive (nda.nih.gov; grant reference R01MH120088). Code for multinomial regression models is provided in the supplement.

REFERENCES

- Charlson FJ, Ferrari AJ, Santomauro DF, Diminic S, Stockings E, Scott JG, et al. Global epidemiology and burden of schizophrenia: findings from the global burden of disease study 2016. *Schizophr Bull.* 2018;44:1195–203.
- Fusar-Poli P, Borgwardt S, Bechdolf A, Addington J, Riecher-Rössler A, Schultze-Lutter F, et al. The psychosis high-risk state: a comprehensive state-of-the-art review. *JAMA Psychiatry.* 2013;70:107.
- De Pablo GS, Radua J, Pereira J, Bonoldi I, Arienti V, Besana F, et al. Probability of transition to psychosis in individuals at clinical high risk: an updated meta-analysis. *JAMA Psychiatry.* 2021;78:970–8.
- Addington J, Stowkowy J, Liu L, Cadenhead KS, Cannon TD, Cornblatt BA, et al. Clinical and functional characteristics of youth at clinical high-risk for psychosis who do not transition to psychosis. *Psychol Med.* 2019;49:1670–7.
- McGlashan T, Walsh B, Woods S. *The Psychosis-Risk Syndrome: Handbook for Diagnosis and Follow-Up.* USA: Oxford University Press; 2010.
- Subotnik KL, Nuechterlein KH, Irzhevsky V, Kitchen CM, Woo SM, Mintz J. Is unawareness of psychotic disorder a neurocognitive or psychological defensiveness problem? *Schizophr Res.* 2005;75:147–57.
- Gold JM, Corlett PR, Strauss GP, Schiffman J, Ellman LM, Walker EF, et al. Enhancing psychosis risk prediction through computational cognitive neuroscience. *Schizophr Bull.* 2020;46:1346–52.
- Corlett PR, Fletcher PC. The neurobiology of schizotypy: Fronto-striatal prediction error signal correlates with delusion-like beliefs in healthy people. *Neuropsychologia.* 2012;50:3612–20.
- Suthaharan P, Reed EJ, Leptourgos P, Kenney JG, Uddenberg S, Mathys CD, et al. Paranoia and belief updating during the COVID-19 crisis. *Nat Hum Behav.* 2021;5:1190–202.
- Powers AR, Mathys C, Corlett PR. Pavlovian conditioning-induced hallucinations result from overweighting of perceptual priors. *Science.* 2017;357:596–600.
- Gold JM, Strauss GP, Waltz JA, Robinson BM, Brown JK, Frank MJ. Negative symptoms of schizophrenia are associated with abnormal effort-cost computations. *Biol Psychiatry.* 2013;74:130–6.
- Heerey EA, Robinson BM, McMahon RP, Gold JM. Delay discounting in schizophrenia. *Cognit Neuropsychiatry.* 2007;12:213–21.
- Silverstein SM, Keane BP. Perceptual organization impairment in schizophrenia and associated brain mechanisms: review of research from 2005 to 2010. *Schizophr Bull.* 2011;37:690–9.
- Millman ZB, Gold JM, Mittal VA, Schiffman J. The critical need for help-seeking controls in clinical high-risk research. *Clin Psychol Sci.* 2019;7:1171–89.
- Schiffman J, Ellman LM, Mittal VA. Individual differences and psychosis-risk screening: practical suggestions to improve the scope and quality of early identification. *Front Psychiatry.* 2019;10:6.
- Mittal VA, Ellman LM, Strauss GP, Walker EF, Corlett PR, Schiffman J, et al. Computerized assessment of psychosis risk. *J Psychiatr Brain Sci.* 2021;6:e210011.
- Addington J, Liu L, Brummitt K, Bearden CE, Cadenhead KS, Cornblatt BA, et al. North American Prodrome Longitudinal Study (NAPLS 3): methods and baseline description. *Schizophr Res.* 2020. <https://doi.org/10.1016/j.schres.2020.04.010>.
- Peduzzi P, Concato J, Kemper E, Holford TR, Feinstein AR. A simulation study of the number of events per variable in logistic regression analysis. *J Clin Epidemiol.* 1996;49:1373–9.
- First MB, Williams JB, Karg RS, Spitzer RL. Structured clinical interview for DSM-5—Research version (SCID-5 for DSM-5, research version; SCID-5-RV). Arlington, VA: American Psychiatric Association; 2015.
- Cannon TD, Yu C, Addington J, Bearden CE, Cadenhead KS, Cornblatt BA, et al. An individualized risk calculator for research in prodromal psychosis. *AJP.* 2016;173:980–8.
- Keefe RSE, Goldberg TE, Harvey PD, Gold JM, Poe MP, Coughenour L. The Brief Assessment of Cognition in Schizophrenia: reliability, sensitivity, and comparison with a standard neurocognitive battery. *Schizophr Res.* 2004;68:283–97.
- Brandt J, Benedict RH Hopkins verbal learning test—revised: professional manual. Psychological Assess Resour; 2001.
- Dohrenwend BS, Askenasy AR, Krasnoff L, Dohrenwend BP. Exemplification of a method for scaling life events: the PERI life events scale. *J Health Soc Behav.* 1978;19:205–29.
- Carrión RE, Auther AM, McLaughlin D, Olsen R, Addington J, Bearden CE, et al. The global functioning: social and role scales—further validation in a large sample of adolescents and young adults at clinical high risk for psychosis. *Schizophr Bull.* 2019;45:763–72.
- Alderson-Day B, Lima CF, Evans S, Krishnan S, Shanmugalingam P, Fernyhough C, et al. Distinct processing of ambiguous speech in people with non-clinical auditory verbal hallucinations. *Brain.* 2017;140:2475–89.
- Treadway MT, Buckholtz JW, Schwartzman AN, Lambert WE, Zald DH. Worth the 'EFFORT'? The effort expenditure for rewards task as an objective measure of motivation and anhedonia. *PLoS ONE.* 2009;4:e6598.
- Strauss GP, Ruiz I, Visser KH, Crespo LP, Dickinson EK. Diminished Hedonic response in neuroleptic-free youth at ultra high-risk for psychosis. *Schizophr Res Cogn.* 2018;12:1–7.
- Damme KSF, Osborne KJ, Gold JM, Mittal VA. Detecting motor slowing in clinical high risk for psychosis in a computerized finger tapping model. *Eur Arch Psychiatry Clin Neurosci.* 2020;270:393–7.
- Mittal VA, Gupta T, Keane BP, Silverstein SM. Visual context processing dysfunctions in youth at high risk for psychosis: resistance to the Ebbinghaus illusion and its symptom and social and role functioning correlates. *J Abnorm Psychol.* 2015;124:953–60.
- Silverstein SM, Thompson JL, Gold JM, Schiffman J, Waltz JA, Williams TF, et al. Increased face detection responses on the mooney faces test in people at clinical high risk for psychosis. *NPJ Schizophr.* 2021;7:1–7.
- RStudio Team. RStudio. 2020.
- van Buuren S, Groothuis-Oudshoorn K. mice: multivariate imputation by chained equations in R. *J Stat Softw.* 2011;45:1–67.
- Collin G, Nieto-Castanon A, Shenton ME, Pasternak O, Kelly S, Keshavan MS, et al. Brain functional connectivity data enhance prediction of clinical outcome in youth at risk for psychosis. *Neuroimage Clin.* 2020;26:102108.
- Sanfelici R, Dwyer DB, Antonucci LA, Koutsouleris N. Individualized diagnostic and prognostic models for patients with psychosis risk syndromes: a meta-analytic view on the state of the art. *Biol Psychiatry.* 2020;88:349–60.
- Woods SW, Parker S, Kerr MJ, Walsh BC, Wijtenburg SA, Prunier N, et al. Development of the PSYCHS: Positive Symptoms and Diagnostic Criteria for the CAARMS Harmonized with the SIPS. *Early Intervention Psychiatry.* in press. <https://doi.org/10.1111/eip.13457>.
- Mueser KT, DeTore NR, Kredlow MA, Bourgeois ML, Penn DL, Hintz K. Clinical and demographic correlates of stigma in first-episode psychosis: the impact of duration of untreated psychosis. *Acta Psychiatr Scand.* 2020;141:157–66.
- Kular A, Perry BI, Brown L, Gajwani R, Jasini R, Islam Z, et al. Stigma and access to care in first-episode psychosis. *Early Interv Psychiatry.* 2019;13:1208–13.
- Vega WA, Lewis-Fernández R. Ethnicity and variability of psychotic symptoms. *Curr Psychiatry Rep.* 2008;10:223–8.
- Mirza A, Birtel MD, Pyle M, Morrison AP. Cultural differences in psychosis: the role of causal beliefs and stigma in white British and South Asians. *J Cross-Cultural Psychol.* 2019;50:441–59.
- Huys QJM, Maia TV, Frank MJ. Computational psychiatry as a bridge from neuroscience to clinical applications. *Nat Neurosci.* 2016;19:404–13.
- Corlett PR, Fletcher PC. Computational psychiatry: a Rosetta Stone linking the brain to mental illness. *Lancet Psychiatry.* 2014;1:399–402.

ACKNOWLEDGEMENTS

This study was supported by the National Institute of Mental Health (grant R01MH120090 to Drs. Gold and Waltz; grant R01MH120091 to Dr. Ellman; grant R01MH120092 to Dr. Strauss; grant R01MH120089 to Drs. Woods and Corlett; and grant F32MH133302 to Dr. Williams).

AUTHOR CONTRIBUTIONS

TFW conceptualized the present study, conducted formal analyses, wrote the original draft, and was involved in reviewing and editing the manuscript. JMG, JAW, JS, LME, GPS, EFW, SWW, ARP, PRC, SMS, and REZ were all involved with the study's funding acquisition, methodology, investigation, and manuscript review. JK, MKP, and TT were involved with the study's software, investigation, data curation, and manuscript review. VAM supervised the study, aided its conceptualization, and was involved in funding acquisition, methodology, investigation, and manuscript review.

ETHICAL APPROVAL AND CONSENT TO PARTICIPATE

This work was approved by Northwestern University as the Institutional Review Board of record (STU00211351) and acknowledged by the Institutional Review Boards of all other participating sites. Written informed consent was provided by all adult participants. Minors provided written assent and their parents or guardians provided written informed consent. All work was conducted in accordance with the Declaration of Helsinki.

COMPETING INTERESTS

The authors declare no competing interests.

ADDITIONAL INFORMATION

Supplementary information The online version contains supplementary material available at <https://doi.org/10.1038/s41398-025-03539-5>.

Correspondence and requests for materials should be addressed to Trevor F. Williams.

Reprints and permission information is available at <http://www.nature.com/reprints>

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/>.

© The Author(s) 2025