Psychological Medicine

cambridge.org/psm

Poor functional recovery is better predicted than conversion in studies of outcomes of clinical high risk of psychosis: insight from SHARP

TianHong Zhang[1], ShuWen Yang[1,*], LiHua Xu[1], XiaoChen Tang[1], YanYan Wei[1], HuiRu Cui[1], HuiJun Li[2], YingYing Tang[1], Li Hui[3], ChunBo Li[1], XingShi Chen[1,3,†] and JiJun Wang[1,4,5]

[1]Shanghai Mental Health Center, Shanghai Jiaotong University School of Medicine, Shanghai Key Laboratory of Psychotic Disorders, Shanghai 200030, PR China; [2]Department of Psychology, Florida A&M University, Tallahassee, Florida 32307, USA; [3]Institute of Mental Health, The Affiliated Guangji Hospital of Soochow University, Soochow University, Suzhou 215137, Jiangsu, PR China; [4]Bio-X Institutes, Key Laboratory for the Genetics of Developmental and Neuropsychiatric Disorders (Ministry of Education), Shanghai, PR China and [5]Brain Science and Technology Research Center, Shanghai Jiao Tong University, Shanghai, China

Original Article

*This author shares first authorship.

†Co-corresponding author.

Cite this article: Zhang T et al (2020). Poor functional recovery is better predicted than conversion in studies of outcomes of clinical high risk of psychosis: insight from SHARP. *Psychological Medicine* **50**, 1578–1584. https://doi.org/10.1017/S0033291719002174

Received: 12 April 2019
Revised: 16 July 2019
Accepted: 1 August 2019
First published online: 27 August 2019

Key words:
Follow-up; prediction; risk calculator; transition; ultra-high risk

Author for correspondence:
JiJun Wang, E-mail: jijunwang27@163.com; XingShi Chen, E-mail: chenxingshi2008@163.com

Abstract

Background. Few of the previous studies of clinical high risk of psychosis (CHR) have explored whether outcomes other than conversion, such as poor functioning or treatment responses, are better predicted when using risk calculators. To answer this question, we compared the predictive accuracy between the outcome of conversion and poor functioning by using the NAPLS-2 risk calculator.

Methods. Three hundred CHR individuals were identified using the Chinese version of the Structured Interview for Prodromal Symptoms. Of these, 228 (76.0%) completed neurocognitive assessments at baseline and 199 (66.3%) had at least a 1-year follow-up assessment. The latter group was used in the NAPLS-2 risk calculator.

Results. We divided the sample into two broad categories based on different outcome definitions, conversion ($n = 46$) *v.* non-conversion ($n = 153$) or recovery ($n = 138$) *v.* poor functioning ($n = 61$). Interestingly, the NAPLS-2 risk calculator showed moderate discrimination of subsequent conversion to psychosis in this sample with an area under the receiver operating characteristic curve (AUC) of 0.631 ($p = 0.007$). However, for discriminating poor functioning, the AUC of the model increased to 0.754 ($p < 0.001$).

Conclusions. Our results suggest that the current risk calculator was a better fit for predicting a poor functional outcome and treatment response than it was in the prediction of conversion to psychosis.

Effective prediction of clinical high risk of psychosis (CHR) is a current major challenge for the psychiatric medical community and has seen little progress over the last several years (in general, accuracy rates remain under 75%). One of the most important risk calculators for predicting conversion to psychosis in CHR individuals is available on the internet (http://riskcalc.org:3838/napls/), and based on clinical, cognitive, and demographic data from the NAPLS-2 sample (Cannon *et al.*, 2016). Our previously study cross-validated the NAPLS-2 psychosis risk calculator in a Chinese CHR sample (Zhang *et al.*, 2018*a*). However, a fundamental question remains, whether conversion is the appropriate outcome to be predicted.

The 'bad' outcome in prediction studies has generally been defined as conversion to psychosis, which is in accordance with the criteria of the 'Presence of a Psychotic Syndrome (POPS)' (McGlashan *et al.*, 2010). The operational definition of conversion was based on the presence of a six-level positive symptom scale (the rating '6' refers to severe and psychotic, i.e. conviction of psychotic experiences). However, such a specific definition of conversion based only on positive symptoms may not be the best option (van Os and Guloksuz, 2017). From a clinical point of view, currently defined endpoints may neither adequately address the characteristic presenting symptoms (such as the negative symptoms) and functioning (such as social and role functioning and cognitive deficits) during the progression of the disease nor reflect the clinicians' and patients' concerns (Zhang *et al.*, 2018*b*). Clinically it is extremely important to know whether a poor functional outcome is also predictable when using risk calculators in CHR individuals. Several non-drug interventions such as supportive psychotherapy (Rosenbaum *et al.*, 2012) are actually more about improving global function than treating psychotic symptoms. Since there are broad debates (Miller *et al.*, 2002; Liu and Demjaha, 2013) on antipsychotic treatment for CHR individuals, if

CAMBRIDGE UNIVERSITY PRESS

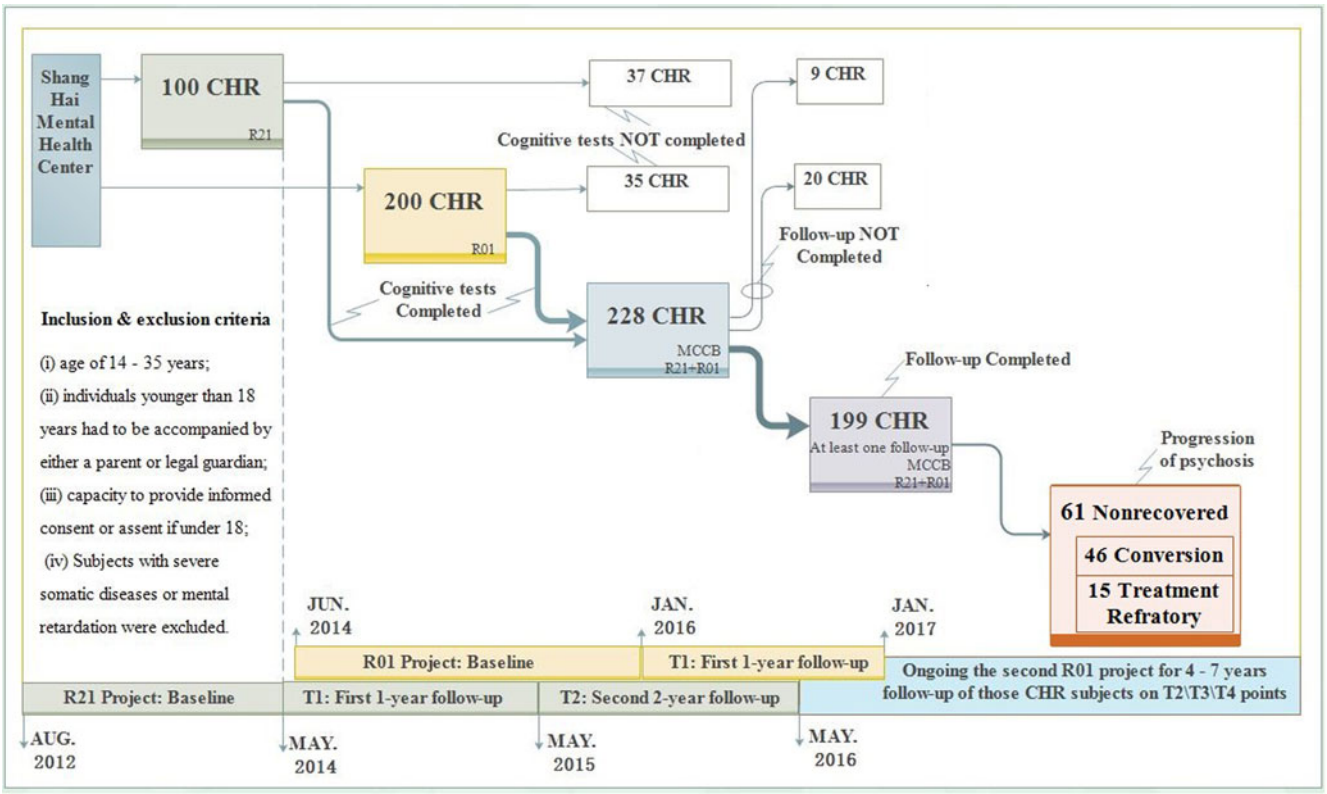

Fig. 1. Flow chart of sample recruitment and follow-up. *Note:* Overview of the Shanghai at Risk for Psychosis (SHARP) program: (1) first phase of the SHARP program recruited 100 CHR youths between August 2012 and May 2014, supported by a Fogarty and National Institutes of Mental Health (NIMH) grant (1R21 MH093294-01A1), titled, 'Broadening the Investigation of Psychosis Prodrome to Different Cultural Groups' (abbreviated to R21). (2) Second phase of the SHARP program recruited 200 CHR youths between June 2014 and January 2017, supported by an NIMH grant (1R01 MH 101052-01), titled 'Validating Biomarkers for the Prodrome and Transition to Psychosis in Shanghai' (abbreviated to R01). (3) Third phase of the SHARP program is ongoing with the second R01 NIMH grant (1 R01 MH111448-01) titled 'A Psychobiological Follow-up Study of Transition from Prodrome to Early Psychosis'.

functional outcomes can be predicted and become the target of treatment, these non-drug interventions may play a more important role in early intervention. In this context, we provide evidence based on a recent and ongoing project with a large CHR cohort, to emphasize the needs for an extended definition of CHR outcomes to improve the precision and value of predictions.

Methods

Project

The SHARP study was conducted with CHR individuals enrolled in an early identification program for psychosis, implemented at one site, the SMHC in China. The Research Ethics Committees at the SMHC approved these studies. A key element of the SHARP study is that all participants are psychotropically naïve when they enter the study and are assessed clinically. They generally have had no treatment of any kind for a psychiatric disorder. There is also no history of substance abuse or dependence. Participants are not treated in the study, but receive treatment as required, provided by their community psychiatrist, after their baseline assessment. As noted above, 300 CHR individuals were recruited and assessed during 2012–2016 (Fig. 1).

Sample

All participants gave written informed consent at the recruitment stage of the study. Subjects younger than 18 years of age had their consent forms signed by their parents and provided assent personally. A total of 300 CHR individuals were identified in a face-to-face interview using the SIPS (Miller *et al.*, 2002, 2003). Among them, 228 (76.0%) completed neurocognitive assessments at baseline. There was no difference between subjects with and without neurocognitive assessments on the NAPLS-2 key predictor variables except for age. Subjects with cognitive assessments [mean age 19.0 years (S.D. = 5.0), $n = 228$] were younger than those without cognitive assessments [mean age 20.9 (S.D. = 6.4), $n = 72$]. Of the 228 subjects with neurocognitive assessments, 199 (87.3%) CHR individuals had at least a 1-year follow-up assessment. At baseline, subjects without follow-up assessments ($n = 29$) had higher levels of unusual thought content and suspiciousness [mean score, 4.7 (S.D. = 1.3) *v.* 3.1 (S.D. = 1.5), $t = 5.424$, $p < 0.001$], and a greater decline in general functioning [3.5 (S.D. = 0.8) *v.* 3.1 (S.D. = 0.9), $t = 2.094$, $p = 0.037$] than those with follow-up assessments. Finally, the data from the 199 CHR individuals were used as the validation sample for testing the NAPLS-2 psychosis risk calculator. Details of the study procedures, study setting, implementation of the measurement and assessment are reported elsewhere (Zheng *et al.*, 2012; Zhang *et al.*, 2014; Zhang *et al.*, 2015; Zhang *et al.*, 2017).

Predictors

There were eight predictor variables included in the NAPLS-2 psychosis risk calculator in Table 1.

https://doi.org/10.1017/S0033291719002174 Published online by Cambridge University Press

Table 1. Predictor variables included in the NAPLS-2 and SHARP model

Predictor variables
Demographic variable:
(i) Age (ranged from 12 to 35 years of age)
Cognitive variables:
(ii) The BACS (Brief Assessment of Cognition in Schizophrenia) Symbol Coding Raw Score (Keefe *et al*., 2004)
(iii) The HVLT-R (Hopkins Verbal Learning Test–Revised) Total Raw Score (Shapiro *et al*., 1999)
Note: Predictor (ii) and (iii) are subtests of the MATRICS Consensus Cognitive Battery (MCCB) (Kern et al., 2008, 2011). The Chinese version of the MCCB (Shi et al., 2013) was performed in the SHARP program.
Clinical variable:
(iv) The sum of the rescaled SIPS ratings for unusual thought content (P1) and suspiciousness (P2) (range from 0 to 8). The P1 or P2 item rated 0–2 on the original scale were recoded as 0. Scores of 3–6 on the original scale were recoded as 1–4.
Functional variable:
(v) Change in Global social functioning in the year prior to baseline (range from 0 to 6), was derived from the Global Functioning: Social Functioning scale in the NAPLS project. *Note:* In the SHARP study, because the GF: social was only available on the RO1 sample, we chose to apply the global assessment of function (GAF) (Jones et al., 1995) change score which measures functional deterioration (score relative to 12 months prior) in the SIPS/SOPS interview because it was given to both R21 and RO1 samples. A GAF score that has declined to 5% or less of the previous best GAF is recoded as 0, whereas declines of 5–15%, 15–25%, 25–35%, 35–45%, 45–55%, 55–65% of the previous best GAF are recorded as 0–6.
Family history variable:
(vi) Has at least one first-degree relative with a psychotic disorder.
Life events and trauma variables:
(vii) and (viii) The undesirable life events score and number of types of trauma endorsed were not included in the SHARP study, and they were not significant in the NAPLS-2 development sample.

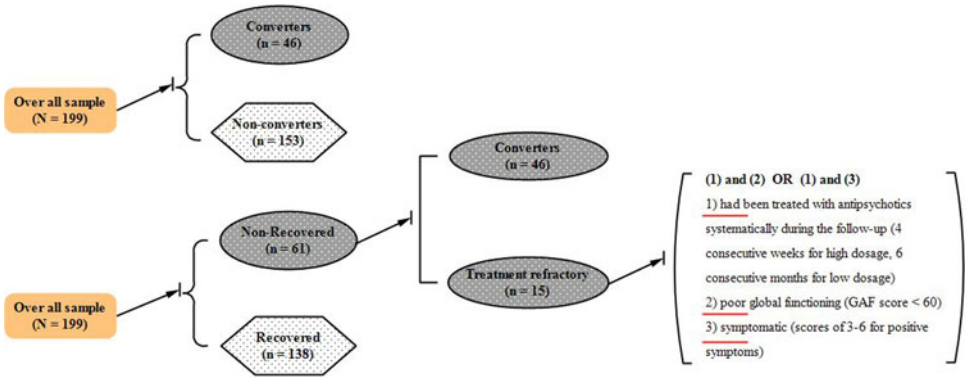

Fig. 2. Outcome classification chart.

Outcome variables

Of the total 199 CHR individuals, 46 (23.1%) converted to full psychosis over 2 years of follow-up. Conversion to psychosis was defined using the POPS criteria (McGlashan *et al*., 2010).

In addition to the outcome categories of converters and non-converters, we divided the existing sample into two broad categories: recovered and non-recovered (Fig. 2). We contend that a category of 'non-recovered' representing a poor functional outcome despite medical treatment is a helpful designation for functional outcomes to take into account the potential ameliorating effects of antipsychotics on positive symptoms. 'Non-recovered' included two subgroups: (1) converters and (2) medicated CHR individuals with either unremitting positive symptoms (but which did not reach the level of persistent psychosis) or poor global functioning at follow-up called 'treatment refractory'. The operational definition of 'treatment refractory' included treatment criteria and refractory criteria in two parts: (1) the treatment criteria (medication was almost the only treatment administered in this sample) was defined either as having taken higher therapeutic dosages of antipsychotic medicines (equivalent to ⩾3 mg/day risperidone) for at least 4 consecutive weeks, or lower therapeutic dosages (equivalent to <3 mg/day risperidone) for at least 6 consecutive months; and, (2) refractory criterion: unremitting symptoms or poor global functioning. The unremitting symptoms were defined as scores of 3–6 for positive symptoms in the SIPS (including Brief Intermittent Psychotic Symptoms/BIPS) at the

https://doi.org/10.1017/S0033291719002174 Published online by Cambridge University Press

Table 2. Baseline predictor variables, comparison between converters and non-converters, and recovered and non-recovered CHR subjects

			Conv. *v.* Non-Conv.				Non-Reco. *v.* Reco.	
Variables	Conversion	Non-conversion	$t/Z/\chi^2$ [c]	*p*	Non-recovered	Recovered	$t/Z/\chi^2$ [c]	*p*
Cases [*n* (%)]	46 (23.1)	153 (76.9)	–	–	61 (30.7)	138 (69.3)	–	–
Age (years) [mean (s.d.)]	19.7 (5.5)	18.9 (4.9)	$t = 0.957$	0.340	19.3 (5.2)	19.1 (5.0)	$t = 0.249$	0.803
Family history of psychosis [*n* (%)]	4 (8.7)	13 (8.5)	$\chi^2 = 0.002$	0.966	7 (11.6)	10 (7.2)	$\chi^2 = 0.968$	0.325
Modified P1 + P2 SIPS items	3.4 (1.6)	3.1 (1.4)	$Z = 1.364$	0.173	3.8 (1.6)	2.9 (1.3)	$Z = 4.011$	<0.001
Decline in GAF: revised score	2.7 (0.6)	2.4 (0.8)	$Z = 2.789$	0.005	2.9 (0.8)	2.3 (0.7)	$Z = 4.116$	<0.001
the BACS[a] [*Mean* (s.d.)]	54.5 (11.4)	58.9 (9.4)	$t = 2.670$	0.008	54.1 (10.4)	59.5 (9.4)	$t = 3.598$	<0.001
the HVLT-R[b] [mean (s.d.)]	22.6 (5.4)	23.7 (5.4)	$t = 1.254$	0.211	21.9 (5.2)	24.2 (5.4)	$t = 2.805$	0.006

BACS, Brief Assessment of Cognition in Schizophrenia; HVLT-R, Hopkins Verbal Learning Test–Revised; $t/Z/\chi^2$, *t* for independent *t* test, *Z* for Mann–Whitney *U* test (nonparametric test), χ^2 for κ test.

follow-up visit. Poor global functioning was defined as a current GAF score of less than 60 at the follow-up point (Austin *et al.*, 2013; Simonsen *et al.*, 2017). This subgroup represents a type of atypical conversion that CHR individuals' positive symptoms are suppressed with antipsychotic treatment, but with poor outcomes even though their condition has not progressed completely to psychosis. The remaining CHR individuals were classified as 'recovered'.

In our previous investigation (Zhang *et al.*, 2017), we observed that compared to the NAPLS sample, a substantially higher percentage of participants in the SHARP sample were prescribed antipsychotics after entering the study rather than before, i.e. after their clinical and cognitive assessments were completed. In the current study, among the final sample of 199 CHR individuals, 160 (80.4%) had taken antipsychotics, 41 (20.6%) had taken antidepressants, and only one subject had undergone four sessions of psychotherapy by the final follow-up. All of these treatments were administered by non-study psychiatrists in the community after baseline clinical and neurocognitive assessment.

Procedures

Participants who met the inclusion criteria were consecutively referred by our clinicians. The participants were informed that this part of the study involved a group of clinical, cognitive and biomarker assessments at baseline with a naturalistic follow-up every 6 months. They would follow the routine clinical treatment procedures provided at the SMHC. All participants from the first visit were followed up for at least 1 year once we attained their consent and intake evaluation information.

All CHR individuals who completed the baseline assessment were followed up every 6 months. Both the CHR individuals and their caregivers had been told that they could contact the interviewer and study clinicians anytime for questions and progress reports on the patients' medical conditions. Except for those who did not desire any further contact, the CHR participants were re-assessed by telephone at the 6th and 18th months or by face-to-face interview using the SIPS at the 12th and 24th months. The outcome determination was based mainly on the face-to-face interviews (of 199 CHR individuals, 119 had at least once face-to-face interview), partly from telephone interviews of CHR individuals or their caregivers, and on the clinical information obtained from clinician reports. The outcome measure included three major components: (1) current symptom level (SIPS interview); (2) current functional level (GAF assessment) and, (3) clinical treatment (detailed pharmacological and psychosocial therapy history).

Data analysis

SPSS version 16.0 was used for data analysis. The six key predictor variables were entered into the NAPLS-2 psychosis risk calculator by two people independently. A new variable of risk ratio for each CHR was constructed by the calculator. Receiver operating characteristic (ROC) analysis (Hanley and McNeil, 1982) was used to test whether the new risk ratio distinguished between converters and non-converters (recovered and non-recovered). The predictive value of the NAPLS-2 psychosis risk calculator was determined according to the area under the ROC curve (AUC) (Swets, 1988). The sensitivity, specificity, positive predictive value and negative predictive values of the NAPLS-2 calculator were examined across different levels of predicted risk. To evaluate the predictive value of the six key predictors from the NAPLS-2 calculator, we performed a binary logistic regression analysis to explore which predictors at baseline best predicted converters and non-converters. Similarly, we assessed recovered *v.* non-recovered subjects. The odds ratio and 95% confidence intervals were estimated for the risk of conversion to psychosis (and non-recovered) in this external validation sample. Based on this regression model, we generated the probabilities of risk for each case. The ROC methodology was used to assess the discriminative power of the probabilities.

Results

Predictor variables included in the risk calculator were compared between conversion *v.* non-conversion and recovered *v.* non-recovered CHR youths in the SHARP samples (Table 2). When comparing predictor variables between groups, the difference between non-recovered *v.* recovered was more distinct than the difference between conversion *v.* non-conversion.

We investigated whether probability risk estimates provided by the NAPLS-2 calculator for each individual in the SHARP validation sample could discriminate converters from non-converters, and recovered CHR subjects from non-recovered. Figure 3 shows that when using recovered or non-recovered as the endpoint (i.e.

https://doi.org/10.1017/S0033291719002174 Published online by Cambridge University Press

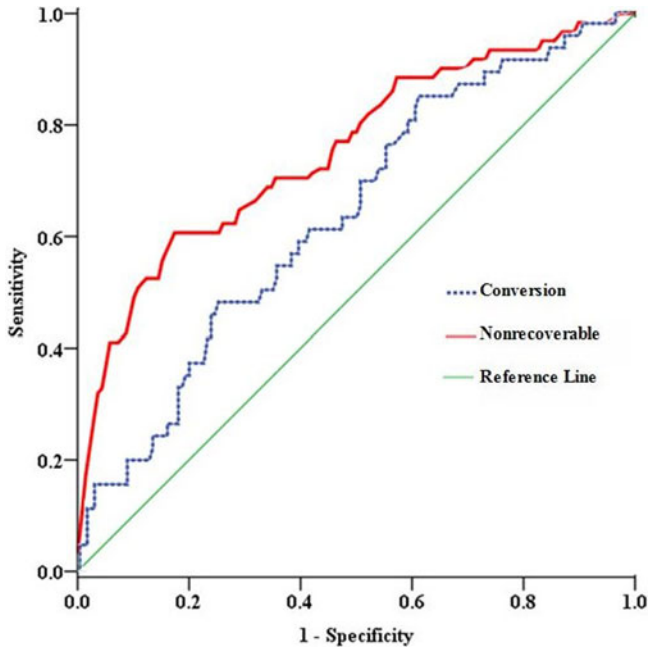

Fig. 3. Receiver operating characteristic curve for the SHARP (Shanghai At Risk for Psychosis project) validation model, classified based on conversion (conversion and non-conversion to psychosis were used as the outcome classification. Conversion was determined using the Structured Interview for Prodromal Symptoms (SIPS)] and non-recovered (converters and medicated CHR individuals with unremitting symptoms or poor global functioning were defined as non-recovered. Medication criteria were defined as either having taken high dosages of antipsychotic medicines for at least 4 weeks, or taken low dosages for at least 6 months. The unremitting symptoms criterion required subjects to have scores of 3–6 for positive symptom in the SIPS. Poor global functioning required subjects to have a current GAF score less than 60 at the follow-up point).

converters + non-converter treatment refractory = non-recovered), the ROC analysis resulted in an AUC of 0.754 ($p < 0.001$) for the probability risk estimates, which is better than the AUC for conversion as the principal endpoint (0.631, $p = 0.007$).

As shown in Table 3, a regression model was used to evaluate the effect of six key predictor variables used in the NAPLS-2 psychosis risk calculator in the prediction of poor functional outcome in the SHARP sample. The value of risk probabilities was generated using this regression model for each case and then used for ROC analysis. The overall model for predicting the outcome (either conversion or non-recovered) was not significant when all six independent variables were entered simultaneously, with an overall classification accuracy of 74.6–80.5%. In terms of individual variables, only the unusual thought content + suspiciousness/paranoia item and decline in function showed good discrimination in the model for predicting non-recovered CHR subjects.

Discussion

Through validation of the NAPLS-2 risk calculator in a Chinese CHR sample in Shanghai, the major aim of this study was to evaluate and compare the accuracy of predicting conversion *v.* poor functional outcome. To the best of our knowledge, this is the first attempt to perform such a comparison. Regarding the prediction of conversion to psychosis, the NAPLS-2 calculator did not fit our SHARP data as well as it fit the NAPLS-2 sample; probability risk estimates yielded an AUC of 0.631 with moderate predictive power, but lower than the NAPLS-2 AUC (C-index = 0.71). However, in the prediction of poor functional outcome ('non-recovered', which includes the conversion subgroup), our data show more accurate outcomes (AUC = 0.754) than for conversion alone, comparable to findings reported in NAPLS-2. To be specific, for those CHR youths with risk calculator estimates higher than 30%, these estimates had moderate sensitivity (53%) and excellent specificity (86%) for predicting a poor functional outcome. This information provides a critical first step in the development of the risk calculator for predicting a poor functional outcome, rather than predicting conversion alone.

The ultimate goal for CHR identification is to provide more precise risk prediction to guide evidence-based, personalized treatments. However, no previous studies have explored whether outcomes other than conversion, such as poor functioning or treatment responses, are better predicted. Our data obviously show that a poor functional outcome is better predicted the conversion. This may be explained by the fact that conversion only considers the progression of positive psychotic symptoms, which can be paroxysmal. However, functional performance can continue to deteriorate, which can be more predictable. Another reason is those predictors (such as cognitive and general function) included in the risk calculator are more likely to be related to functional outcomes.

Consistent with the NAPLS-2 model and our previous findings, the baseline severity level of unusual thought content + suspiciousness and global function decline were significant predictors of psychosis and poor functional outcome in the SHARP sample. Of note here, evidence has been accumulating that baseline disordered thought symptoms and functional deterioration is key risk factors for predicting the onset of psychosis in CHR syndromes (Fusar-Poli *et al.*, 2013). In addition, our data highlight the importance of a declining GAF score in the prediction of psychosis. We used this measure in place of the Global Functioning: Social scale used in the NAPLS-2 model because the latter has not been used by Chinese clinicians. Similarly, cultural differences have not been examined and may affect its validity. In addition, its predictive value in the EDIPPP model was not significant (Cornblatt *et al.*, 2015). In contrast, the GAF score can be derived from the SIPS assessment and has been widely used in China for many years. Therefore, our results further showed that the GAF could be an alternative measure for predictive models, especially for a predicting poor outcome.

Although scores on the BACS and HVLT-R neurocognitive tests were significantly different between recovered and non-recovered CHR groups, they are not significant predictors of conversion or non-recovery in the SHARP sample. One possible cause of this discrepancy is the difference in the cognitive characteristics between SHARP and NAPLS-2 samples. Mean HVLT-R scores in the SHARP sample were different from those in the NAPLS-2 sample, which cannot exclude the effect of differences in the reliability or validity of MATRICS tests across populations. Furthermore, in contrast to the NAPLS-2 studies, family history of a psychotic disorder was not a significant predictor of psychosis in the SHARP sample. As for a family history of psychosis, the lower proportion of family history-positive cases in the SHARP sample could also account for the lower predictive power. There is also the possibility that information regarding psychosis history may be easily overlooked or unreported (Milne *et al.*, 2009) in the SHARP sample due to the lack of psychiatric resources available to older generations and the stigma attached to diagnostic labels (Roy *et al.*, 1996).

https://doi.org/10.1017/S0033291719002174 Published online by Cambridge University Press

Table 3. Performance of key predictors from the NAPLS-2 psychosis risk calculator in the SHARP sample

	Conversion			Non-recovered		
Predictor	Odds ratio	95% CI	*p*	Odds ratio	95% CI	*p*
Age	0.967	0.904–1.035	0.334	1.002	0.935–1.074	0.952
Family history of psychosis	1.373	0.398–4.733	0.616	1.079	0.332–3.509	0.899
Modified P1 + P2 SIPS items[a]	0.885	0.699–1.121	0.312	0.620	0.476–0.807	<0.001
Decline in global assessment of function (GAF): revised score[b]	0.516	0.338–0.789	0.002	0.331	0.201–0.545	<0.001
Hopkins Verbal Learning Test–Revised: raw score	1.003	0.938–1.074	0.924	1.040	0.972–1.112	0.257
Brief Assessment of Cognition in Schizophrenia symbol coding test: raw score	1.026	0.988–1.065	0.183	1.032	0.994–1.112	0.104
AUC[c] for the overall model	0.746 (95% CI = 0.671–0.821, $p < 0.001$)			0.805 (95% CI = 0.742–0.867, $p < 0.001$)		

[a]'Modified P1 + P2 SIPS items' represents the severity of unusual thought content and suspiciousness (items P1 and P2 in SIPS). The P1 or P2 item rated 0–2 on the original scale are recoded as 0, 3–6 on the original scale are recoded as 1–4.
[b]'Decline in global assessment of function (GAF): revised' represents the social functioning decline, corresponding to the 'Decline in Global Functioning: Social scale score' from the NAPLS-2 calculator. The GAF score is dropped to 5% or less of previous best GAF are recoded as 0, dropped to 5–15%, 15–25%, 25–35%, 35–45%, 45–55%, 55–65% of previous best GAF are recorded as 1–6. Decline in GAF score: Change in GAF in the year prior to baseline was derived from the highest GAF score in the past year prior to baseline minus baseline GAF score.
[c]'AUC': The receiver operating characteristic curve.

In contrast to the US samples, the CHR individuals from SHARP were typically administered antipsychotic medications after their first visit to clinicians. Since the definition of conversion to psychosis in the NAPLS-2 study mainly relies on positive psychotic symptoms, the decrease in positive symptoms by antipsychotics in CHR individuals may be a new type of outcome in this Chinese cohort, which is characterized by systematic treatment of antipsychotic drugs, with a poor functional outcome or being symptomatic at the end of follow-up, but not experiencing conversion. We classified these individuals as 'treatment refractory', which represents atypical conversion. The positive symptoms among such CHR individuals are not as severe such that they should be classified as psychosis. Given the predictive value of more severe positive symptoms, which also make one more likely to be prescribed antipsychotics, it may be that a number of these individuals would have converted had they not been on medications at the time of assessment or that they may be more likely to convert over a longer follow-up period. The non-recovered CHR individuals who were treated systemically with antipsychotics in our sample manifested a lower level of positive symptoms at follow-up, but this may be due to medication effects. This possibility is supported by significantly worse negative symptoms with poorer general social and role function at follow-up compared with baseline; the functional outcome is less likely to be affected by antipsychotics. Together with conversion, prediction of non-recovery can offer clinicians more useful information than conversion alone. This underscores the importance of more frequent follow-ups of this subgroup to monitor clinical progression and treatment response.

Strengths and limitations

A major strength of our study is that we used identical measurements to the NAPLS-2 that were available (SIPS & MCCB) and had validated Chinese versions. Another strength is that in contrast to the NAPLS-2 data, which were collected from eight sites (Addington *et al.*, 2012), the current SHARP sample was recruited by one team from one catchment area, which may be more advantageous for its homogeneity. A limitation of the study is that the SHARP CHR cohort was surveyed naturalistically. There were 145 CHR individuals treated with antipsychotic for at least 2 weeks during the follow-up period. Previous studies showed that patients with schizophrenia who did not receive antipsychotics had better long-term outcomes compared to treated patients (Harrow *et al.*, 2012). Therefore, it remains unknown whether the antipsychotic can affect the functional outcome. The various medications the participants took with varying compliance may have confounded the results of functional outcome assessments, thereby limiting the generalizability to CHR subjects who have not taken any medication. Besides, the widespread use of antipsychotics in the current study may lead to the baseline positive symptoms was no longer a strong predictor of conversion (see Table 3). Moreover, although only 29 CHR individuals were lost in follow-up, they demonstrated more severe positive symptoms and poorer functioning at baseline than those who completed follow-up, which could bias our results by underestimating the clinical severity of our sample. Finally, as emphasized by Cannon (Cannon *et al.*, 2016) and Carrion (Carrion *et al.*, 2016), the risk calculator remains experimental. It should only be used in research settings with clinicians who have undergone rigorous SIPS training (SIPS scores being at the core of the model) at this point, and not yet used in general clinical settings with individuals until its clinical utility and properties are validated more firmly.

In summary, although the definition of conversion to psychosis in most CHR-based studies relies only on positive psychotic symptoms, our evidence indicates that functional recovery should be provided a far more central position with respect to the outcome prediction of CHR individuals. It is suggested that a new standardized form of clinical outcome combining functional deficits should at least be developed and applied in future studies. Together with conversion, prediction of the functional outcome can provide clinicians with more useful information than that obtained when considering conversion alone.

Acknowledgements. This study was supported by the Ministry of Science and Technology of China, National Key R&D Program of China (2016YFC1306803), National Natural Science Foundation of China (81671329, 81671332), Shanghai Key Laboratory of Psychotic Disorders (13dz2260500), Science and Technology Commission of Shanghai

https://doi.org/10.1017/S0033291719002174 Published online by Cambridge University Press

Municipality (19441907800), Shanghai Jiaotong University Foundation (ZH2018ZDB03), The Clinical Research Center at Shanghai Mental Health Center (CRC2018ZD01, CRC2018ZD04 and CRC2018YB01) and Shanghai Mental Health Center Foundation (2016-FX-01, 2017-TSXK-03). For the purpose of commemorate, Dr Larry J. Seidman passed away on 7 September 2017 and Dr Robert W. McCarley passed away on 27 May 2017. Both were founders and core members of the SHARP (ShangHai At Risk for Psychosis) project.

Conflict of interest. None.

References

Addington J, Cadenhead KS, Cornblatt BA, Mathalon DH, McGlashan TH, Perkins DO, Seidman LJ, Tsuang MT, Walker EF, Woods SW, Addington JA and Cannon TD (2012) North American Prodrome Longitudinal Study (NAPLS 2): overview and recruitment. *Schizophrenia Research* **142**, 77–82.

Austin SF, Mors O, Secher RG, Hjorthoj CR, Albert N, Bertelsen M, Jensen H, Jeppesen P, Petersen L, Randers L, Thorup A and Nordentoft M (2013) Predictors of recovery in first episode psychosis: the OPUS cohort at 10 year follow-up. *Schizophrenia Research* **150**, 163–168.

Cannon TD, Yu C, Addington J, Bearden CE, Cadenhead KS, Cornblatt BA, Heinssen R, Jeffries CD, Mathalon DH, McGlashan TH, Perkins DO, Seidman LJ, Tsuang MT, Walker EF, Woods SW and Kattan MW (2016) An individualized risk calculator for research in prodromal psychosis. *The American Journal of Psychiatry* **173**, 980–988.

Carrion RE, Cornblatt BA, Burton CZ, Tso IF, Auther AM, Adelsheim S, Calkins R, Carter CS, Niendam T, Sale TG, Taylor SF and McFarlane WR (2016) Personalized prediction of psychosis: external validation of the NAPLS-2 psychosis risk calculator with the EDIPPP project. *The American Journal of Psychiatry* **173**, 989–996.

Cornblatt BA, Carrion RE, Auther A, McLaughlin D, Olsen RH, John M and Correll CU (2015) Psychosis prevention: a modified clinical high risk perspective from the recognition and prevention (RAP) program. *The American Journal of Psychiatry* **172**, 986–994.

Fusar-Poli P, Borgwardt S, Bechdolf A, Addington J, Riecher-Rossler A, Schultze-Lutter F, Keshavan M, Wood S, Ruhrmann S, Seidman LJ, Valmaggia L, Cannon T, Velthorst E, De Haan L, Cornblatt B, Bonoldi I, Birchwood M, McGlashan T, Carpenter W, McGorry P, Klosterkotter J, McGuire P and Yung A (2013) The psychosis high-risk state: a comprehensive state-of-the-art review. *JAMA Psychiatry* **70**, 107–120.

Hanley JA and McNeil BJ (1982) The meaning and use of the area under a receiver operating characteristic (ROC) curve. *Radiology* **143**, 29–36.

Harrow M, Jobe TH and Faull RN (2012) Do all schizophrenia patients need antipsychotic treatment continuously throughout their lifetime? A 20-year longitudinal study. *Psychological Medicine* **42**, 2145–2155.

Jones SH, Thornicroft G, Coffey M and Dunn G (1995) A brief mental health outcome scale-reliability and validity of the Global Assessment of Functioning (GAF). *The British Journal of Psychiatry* **166**, 654–659.

Keefe RS, Goldberg TE, Harvey PD, Gold JM, Poe MP and Coughenour L (2004) The brief assessment of cognition in schizophrenia: reliability, sensitivity, and comparison with a standard neurocognitive battery. *Schizophrenia Research* **68**, 283–297.

Kern RS, Nuechterlein KH, Green MF, Baade LE, Fenton WS, Gold JM, Keefe RS, Mesholam-Gately R, Mintz J, Seidman LJ, Stover E and Marder SR (2008) The MATRICS consensus cognitive battery, part 2: co-norming and standardization. *The American Journal of Psychiatry* **165**, 214–220.

Kern RS, Gold JM, Dickinson D, Green MF, Nuechterlein KH, Baade LE, Keefe RS, Mesholam-Gately RI, Seidman LJ, Lee C, Sugar CA and Marder SR (2011) The MCCB impairment profile for schizophrenia outpatients: results from the MATRICS psychometric and standardization study. *Schizophrenia Research* **126**, 124–131.

Liu CC and Demjaha A (2013) Antipsychotic interventions in prodromal psychosis: safety issues. *CNS Drugs* **27**, 197–205.

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

Miller TJ, McGlashan TH, Rosen JL, Somjee L, Markovich PJ, Stein K and Woods SW (2002) Prospective diagnosis of the initial prodrome for schizophrenia based on the structured interview for prodromal syndromes: preliminary evidence of interrater reliability and predictive validity. *The American Journal of Psychiatry* **159**, 863–865.

Miller TJ, McGlashan TH, Rosen JL, Cadenhead K, Cannon T, Ventura J, McFarlane W, Perkins DO, Pearlson GD and 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. *Schizophrenia Bulletin* **29**, 703–715.

Milne BJ, Caspi A, Crump R, Poulton R, Rutter M, Sears MR and Moffitt TE (2009) The validity of the family history screen for assessing family history of mental disorders. *American Journal of Medical Genetics. Part B, Neuropsychiatric Genetics* **150**, 41–49.

Rosenbaum B, Harder S, Knudsen P, Koster A, Lindhardt A, Lajer M, Valbak K and Winther G (2012) Supportive psychodynamic psychotherapy versus treatment as usual for first-episode psychosis: two-year outcome. *Psychiatry* **75**, 331–341.

Roy MA, Walsh D and Kendler KS (1996) Accuracies and inaccuracies of the family history method: a multivariate approach. *Acta Psychiatrica Scandinavica* **93**, 224–234.

Shapiro AM, Benedict RH, Schretlen D and Brandt J (1999) Construct and concurrent validity of the Hopkins Verbal Learning Test-revised. *The Clinical Neuropsychologist* **13**, 348–358.

Shi C, He Y, Cheung EF, Yu X and Chan RC (2013) An ecologically valid performance-based social functioning assessment battery for schizophrenia. *Psychiatry Research* **210**, 787–793.

Simonsen C, Faerden A, Romm KL, Berg AO, Bjella T, Sundet K, Ueland T, Andreassen O and Melle I (2017) Early clinical recovery in first-episode psychosis: symptomatic remission and its correlates at 1-year follow-up. *Psychiatry Research* **254**, 118–125.

Swets JA (1988) Measuring the accuracy of diagnostic systems. *Science* **240**, 1285–1293.

van Os J and Guloksuz S (2017) A critique of the 'ultra-high risk' and 'transition' paradigm. *World Psychiatry* **16**, 200–206.

Zhang T, Li H, Woodberry KA, Seidman LJ, Zheng L, Li H, Zhao S, Tang Y, Guo Q, Lu X, Zhuo K, Qian Z, Chow A, Li C, Jiang K, Xiao Z and Wang J (2014) Prodromal psychosis detection in a counseling center population in China: an epidemiological and clinical study. *Schizophrenia Research* **152**, 391–399.

Zhang T, Li H, Woodberry KA, Seidman LJ, Chow A, Xiao Z and Wang J (2015) Interaction of social role functioning and coping in people with recent-onset attenuated psychotic symptoms: a case study of three Chinese women at clinical high risk for psychosis. *Neuropsychiatric Disease and Treatment* **11**, 1647–1654.

Zhang TH, Li HJ, Woodberry KA, Xu LH, Tang YY, Guo Q, Cui HR, Liu XH, Chow A, Li CB, Jiang KD, Xiao ZP, Seidman LJ and Wang JJ (2017) Two-year follow-up of a Chinese sample at clinical high risk for psychosis: timeline of symptoms, help-seeking and conversion. *Epidemiology and Psychiatric Sciences*, **26**, 287–298.

Zhang T, Li H, Tang Y, Niznikiewicz MA, Shenton ME, Keshavan MS, Stone WS, McCarley RW, Seidman LJ and Wang J (2018*a*) Validating the predictive accuracy of the NAPLS-2 psychosis risk calculator in a clinical high-risk sample from the SHARP (Shanghai At Risk for Psychosis) program. *The American Journal of Psychiatry* **175**, 906–908.

Zhang T, Xu L, Tang Y, Cui H, Li H, Wei Y, Xu Y, Jiang L, Zhu Y, Li C, Jiang K, Xiao Z and Wang J (2018*b*) Using 'WeChat' online social networking in a real-world needs analysis of family members of youths at clinical high risk of psychosis. *The Australian and New Zealand Journal of Psychiatry* **52**, 375–382.

Zheng L, Wang J, Zhang T, Li H, Li C and Jiang K (2012) The Chinese version of the SIPS/SOPS: a pilot study of reliability and validity. *Chinese Mental Health Journal* **26**, 571–576.