



HHS Public Access

Author manuscript

Am J Psychiatry. Author manuscript; available in PMC 2017 October 01.

Published in final edited form as:

Am J Psychiatry. 2016 October 1; 173(10): 989–996. doi:10.1176/appi.ajp.2016.15121565.

Personalized Prediction of Psychosis: External validation of the NAPLS2 Psychosis Risk Calculator with the EDIPPP project

Ricardo E. Carrión, Ph.D.^{1,2,3}, Barbara A. Cornblatt, Ph.D., M.B.A.^{1,2,3,4}, Cynthia Z. Burton, PhD⁵, Ivy F Tso, PhD, LP⁵, Andrea Auther, PhD^{1,3}, Steven Adelsheim, M.D.⁶, Roderick Calkins⁹, Cameron S. Carter, M.D., Ph.D.⁷, Tara Niendam, Ph.D.⁷, Stephan F. Taylor, M.D.⁵, and William R. McFarlane, M.D.^{10,11}

¹Division of Psychiatry Research, The Zucker Hillside Hospital, North Shore – Long Island Jewish Health System (NS-LIJHS), Glen Oaks, NY, USA

²Center for Psychiatric Neuroscience, The Feinstein Institute for Medical Research, North Shore – Long Island Jewish Health System, Manhasset, New York, 11030, USA

³Department of Psychiatry, Hofstra North Shore-LIJ School of Medicine, Hempstead, NY, USA

⁴Department of Molecular Medicine, Hofstra North Shore-LIJ School of Medicine, Hempstead, NY

⁵Department of Psychiatry, University of Michigan, Ann Arbor, Michigan, USA

⁶Department of Psychiatry, Stanford University, Palo Alto, California, USA

⁷Imaging Research Center, University of California Davis, Sacramento, California, USA

⁸Center for Neuroscience, University of California Davis, Sacramento, California, USA

⁹Mid-Valley Behavioral Care Network, Marion County Health Department, Salem, Oregon, USA

¹⁰Tufts University School of Medicine, Boston, MA

¹¹Maine Medical Center Research Institute, Portland, ME

Abstract

Objective—In the current issue, Cannon and colleagues, as part of the second phase of the North American Prodrome Longitudinal Study (NAPLS2), report on a risk calculator for the individualized prediction of developing a psychotic disorder in a 2-year period. The present study represents an external validation of the NAPLS2 psychosis risk calculator using an independent sample of subjects at clinical high risk for psychosis collected as part of the Early Detection, Intervention, and Prevention of Psychosis Program (EDIPPP).

Methods—176 subjects with follow-up (from the total EDIPPP sample of 210) rated as clinical high-risk (CHR) based on the Structured Interview for Prodromal Syndromes were used to

Address for correspondence: Ricardo E. Carrión, Ph.D., Zucker Hillside Hospital, Psychiatry Research, 75-59 263rd Street, Glen Oaks, New York 11004, Tel: 718-470-8878, Fax: 718-470-5815, RCarrion@NSHS.edu.

Declaration of Interest:

Drs. Carrión, Burton, Tso, Auther, Taylor, Niendam, Adelsheim, and Calkins report no financial relationships with commercial interests. Dr. Cornblatt was the original developer of the CPT-IP. Dr. McFarlane provides training on request to public and not-for-profit clinical services implementing psychosis early intervention programs.

construct a new prediction model with the 6 significant predictor variables in the NAPLS2 psychosis risk calculator (unusual thoughts, suspiciousness, Symbol Coding, verbal learning, social functioning decline, baseline age, and family history). Discrimination performance was assessed with the area under the receiver operating curve (AUC). The NAPLS2 risk calculator was then used to generate a psychosis risk estimate for each case in the external validation sample.

Results—The external validation model showed good discrimination, with an AUC of 79% (95% CI 0.644–0.937). In addition, the personalized risk generated by the NAPLS calculator provided a solid estimation of the actual conversion outcome in the validation sample.

Conclusions—In the companion papers in this issue, two independent samples of CHR subjects converge to validate the NAPLS2 psychosis risk calculator. This prediction calculator represents a meaningful step towards early intervention and personalized treatment of psychotic disorders.

Modern medicine emphasizes prevention as the optimal method of promoting health care and reducing public health costs. Over time, prevention has progressed from identifying populations at risk to personalizing risk estimates(1, 2). This movement is due to several factors. In particular, improvements in technology have made it easier to generate advanced statistical prediction models. As a result, risk calculators have been developed that provide an effective way of teasing apart individuals with the highest probability of illness who require the most aggressive intervention from those who need minimal treatment(3). A number of risk calculators are now freely available to the public and can estimate the risk, for example, of prostate, ovarian, breast, pancreatic, and colorectal cancer(4–6), type 2 diabetes(7, 8), and cardiovascular disease (9–11). Surprisingly, risk calculators for mental health conditions are almost non-existent, even though serious mental illness costs the United States of America \$193.2 billion in lost earnings per year(12).

Over the past 20 years, considerable progress has been made in formalizing and refining the prediction of psychosis. Advances include operationalizing and validating clinical criteria that identify individuals considered to be prodromal or at clinical high risk (CHR) for psychosis, approximately 29% of whom will develop a psychotic illness over two years(13). In addition, the publication of sophisticated multivariable models that predict psychosis with a wide set of risk factors(14–19) has also provided further steps towards enhancing prediction. However, progress has been less rapid in determining how to individualize prediction and provide the probability of psychosis risk on a case-by-case basis. At present, in clinical settings, a general estimate of the risk for psychosis can be roughly approximated from the presence of traditional risk factors in a patient, but solid evidence to providing a mental health professional a more precise estimate of risk for a help-seeking or referred individual during an initial clinical visit is not yet available.

In the current issue of the American Journal of Psychiatry, Cannon and colleagues, as part of the second phase of the North American Prodrome Longitudinal Study (NAPLS2)(20), report on a risk calculator for the individualized prediction of developing a psychotic disorder in a 2-year period. This prediction tool represents a potential breakthrough for early intervention in psychiatry. However, as with any predictive analytic model, its performance needs to be validated in samples of CHR subjects collected independently of NAPLS2(21). The goal of the current study was to evaluate the performance of the NAPLS2 risk calculator

in an external, independent sample of individuals at CHR for psychosis collected as part of the Early Detection, Intervention, and Prevention of Psychosis Program (EDIPPP).

The EDIPPP project was a large nationwide clinical trial designed to examine the effectiveness of Family-aided Assertive Community Treatment (FACT)(22, 23) in preventing the onset of psychosis(24). Over the span of three years (2007–2010) EDIPPP recruited a large sample (N=337) of adolescents and young-adults who were at risk or who were in the very early stages of a major psychotic disorder.

The EDIPPP sample offers a clear test of the applicability of the NAPLS2 risk calculator to an independent sample of CHR subjects, as the two projects had key differences in terms of goals, recruitment strategies, and ascertainment criteria. In addition to delivering treatment, EDIPPP developed a community education and outreach network in six urban and rural sites across the U.S. with the goals of raising awareness of the early warning signs of psychosis and demonstrating the effectiveness of community outreach, education and early referral, combined with the FACT intervention, to reduce the incidence of psychosis(25). In EDIPPP allocation to treatment was based on clinical risk (higher vs. lower) which was determined by a cut-off score (score of 7) on the total attenuated positive symptom severity score as specified by the Scale of Prodromal Symptoms (SOPS) from the Structured Interview for Prodromal Syndromes (SIPS)(24–27). As such, EDIPPP participants had a wide range of attenuated positive symptoms, from little-to-no symptoms in the prodromal range to positive symptoms that reached psychotic threshold. Therefore, since the original EDIPPP sample was categorized differently, this external validation sample was reconfigured to match the NAPLS2 intake criteria, which includes all subjects meeting the Criteria of Prodromal Syndromes (COPS) as defined by the SIPS(28–30).

The current study aimed to assess the predictive ability of the NAPLS2 psychosis risk calculator in an external, independent sample of subjects at CHR for psychosis. Thus, in this report, NAPLS2 will be referred to as the development sample and EDIPPP the external validation sample. Given that the predictors in the NAPLS2 calculator were based on theoretical considerations, we first evaluated the predictive ability of the components of the NAPLS2 model. Six key predictor variables used in the NAPLS2 psychosis risk calculator were used in the external validation sample to construct a new model predicting psychosis. Second, we assessed the performance of the NAPLS2 model by evaluating the predictive accuracy of the risk calculator when applied to the external validation sample. Evaluating the performance of the risk calculator in different CHR samples and settings than those used to initially test the model can further support its empirical validity and clinical utility prior to its widespread use(21).

Methods

The data reported here were collected as part of EDIPPP, a large multi-site clinical trial for preventing psychosis among young people funded by the Robert Wood Johnson Foundation (2007–2011)(24–26). EDIPPP consists of 6 participating sites: 1) Portland Identification and Early Referral (PIER), Portland, ME; 2) Recognition and Prevention Program (RAP), The Zucker Hillside Hospital, Glen Oaks, NY; 3) Michigan Prevents Prodromal Progression

Program (M3P), Ann Arbor, MI; 4) Early Assessment and Support Team Program (EAST), Salem, OR; 5) Early Diagnosis and Preventive Treatment (EDAPT) Clinic, Sacramento, CA; and 6) Early Assessment and Resource Linkage for Youth (EARLY), Albuquerque, NM. Details of the study design, implementation, assessment, psychosocial and pharmacological treatments, methods, and characteristics of the sample have been reported previously (24, 25). Although the Hillside site was in both projects, the present analyses did not include any overlapping subjects.

CHR subjects met one of the three COPS diagnoses based on the SIPS(28–30): 1. Attenuated Positive Symptom Syndrome (APS), presence of one or more moderate, moderately severe, or severe (scores of 3, 4, or 5) SOPS rated (scale of 0–6) attenuated positive symptoms, 2. Genetic Risk and Deterioration Syndrome (GRD), genetic risk for psychosis coupled with deterioration in functioning, and (3) Brief Intermittent Psychotic Syndrome (BIPS), intermittent psychotic symptoms that are recent, brief in duration, and not seriously disorganizing or dangerous.

Two-hundred and ten CHR subjects (92% APS, 6.3% BIPS, 1.7% GRD) were included in the current validation sample. In addition to the 210 CHR subjects, the EDIPPP sample included 32 early first-episode psychosis (EFEP) and 95 low-risk comparison subjects who were excluded from the present study.

Participants aged 12–25 years old were included in the EDIPPP study. Exclusion criteria for the study were: 1. A psychotic episode (score of 6 on any Positive Symptom Scale) 30 consecutive days, 2. A prior episode of psychosis or having received antipsychotic medication 30 days at a dosage appropriate to treat a psychotic episode, 3. IQ<70, 4. Permanent residence outside the catchment area, 5. Non-English speaking, 6. Current imprisonment in the criminal justice system, and 7. Psychotic symptoms due to an acute toxic or medical etiology.

Written informed consent was obtained from patients 18 years of age or older and from parents of patients under 18, with written assent from the patient. The research protocol was approved by the institutional review board at each site.

Baseline Assessments

Details of the baseline clinical assessment have been reported previously.(24) Prodromal symptoms were assessed by SIPS and the companion SOPS(28–30). Social and role functioning was assessed using the Global Functioning:Social (GF:S) and Global Functioning:Role (GF:R) scales(31). In addition to several other clinical measures, the baseline assessment included the Measurement and Treatment Research to Improve Cognition in Schizophrenia Consensus Cognitive Battery (MCCB)(32). The current report utilizes data from two of the tests, namely the Brief Assessment of Cognition in Schizophrenia (BACS), Symbol Coding(33) and Hopkins Verbal Learning Test (HVLT)(34).

Clinical Outcome

Of the initial 210 CHR subjects, 176 (83.8%) had at least one follow-up assessment. Of those 176, 12 (6.8%) transitioned to psychosis over 2 years.(24) Conversion to psychosis

was defined according to the Presence of Psychosis Scale (POPS) criteria on the SIPS: sustaining any 6 on the Positive Symptom Scale for at least an hour per day, at an average of 4 days per week over one month, or demonstrating seriously disorganized or dangerous behavior. The mean follow-up period (time to conversion to psychosis or last follow-up) was 99.29 weeks (SD= 21.51; median= 106.00).

Statistical Analyses

All analyses were conducted using SPSS 20.0 (SPSS Inc., Chicago, Illinois). Comparisons of demographic and clinical characteristics were performed with Student's T-tests for continuous variables, Pearson Chi-Square for categorical variables (Two-tailed, $P < .05$). Overall, 2.4% of the data (25 of 1056 values) were missing. Among those followed, missing values were imputed using mean values for Symbol Coding and HVLT-R (missing 1 value each), and modal values for Family History (missing 14 values) and Decline in GF:S (missing 9 values) prior to use in prediction analyses.

The external validation analysis was carried out in several steps. First, a multivariable Cox proportional hazards regression model was used to estimate hazard ratios (HRs) and 95% confidence intervals (CIs) for risk of conversion to psychosis in the external validation sample. We evaluated the ability of six predictor variables used in the NAPLS2 psychosis risk calculator to predict psychosis: 1. Baseline age; 2. Severity of SIPS items P1+P2 (Unusual Thought Content and Suspiciousness) recoded(18); 3. BACS Symbol Coding Test(raw score)(33); 4. The HVLT-R (sum of trials 1–3)(34); 5. Decline in social functioning in the year prior to the baseline assessment, measured using the GF:S scale (31, 35, 36); and 6. Having a first-degree relative with a psychotic disorder. Predicted probabilities of risk (based on the cumulative hazard function) were computed for each subject in the external validation sample. Trauma and life events, which were not significant in the development sample, were not included.

Discrimination performance (ability of the model to correctly distinguish between outcomes) was assessed for all models by the area under the receiver operating curve (AUC) or Harrell's C-statistic (37, 38). The NAPLS2 calculator was then used to generate risk estimates for each case in the external validation sample. The Gamma statistic examined the agreement between the predicted levels (classes) of risk as predicted by the NAPLS calculator when applied to EDIPPP cases and the predicted risk classes generated by the EDIPPP validation model (as shown in Table 3).(39) The Gamma statistic can range from -1 (perfect negative association) to $+1$ (perfect agreement) with a value of 0 indicating no association. Precision and bias of the NAPLS2 calculator were assessed with the Brier score and mean prediction error (MPE), respectively. The Brier score is the mean squared difference between the observed outcome and predicted risk score(40). MPE was calculated as the difference between the risk estimates generated by the external validation model and the NAPLS2 psychosis risk calculator, providing a metric for the tendency for overestimation/underestimation of risk(41). For both measures, a lower score indicates higher precision and less bias and 15% was considered to be an acceptable level(42). Spearman's Rho correlation analysis was also used to examine the correspondence between the risk estimates generated by both models. Finally, the diagnostic accuracy [Sensitivity,

specificity, Positive Predictive Value (PPV), Negative Predictive Value (NPV)] of the NAPLS2 calculator was examined across different levels of predicted risk.

Results

Table 1 shows baseline demographic and clinical data for the CHR sample. There was no difference between subjects with and without follow-up on any major demographic or clinical feature, including baseline age, gender, race, and education level. CHR subjects in the validation sample had a mean age of 16.6 ± 3.3 years and were mostly male (58.5%) and white (64.1%) (see Table 1). The populations in the validation (EDIPPP) and development (NAPLS2) cohorts were very similar on most of the major demographic and clinical features. However, CHR subjects in the EDIPPP sample were markedly younger than the NAPLS2 cohort (16.6 vs. 18.5 years old), most likely because cases were referred from educational sources.

Table 2 shows the regression model constructed with the validation sample that includes 6 key variables (SIPS items P1+P2, Symbol Coding, HVLT-R, social functioning decline, baseline age, and family history) selected from the NAPLS2 psychosis risk calculator. The overall model was significant when all six independent variables were entered simultaneously ($\chi^2=19.68$, $df=6$, $P=0.003$). The base model of SIPS items P1+P2 showed an acceptable discrimination performance, with an AUC of 67%. Combining the additional five variables with SIPS P1+P2 increased the AUC by 12%, resulting in an AUC of 79.0% (95% CI, 64.4–93.7, $P<0.001$), indicating good discrimination performance (see Figure 1). Scores on the neurocognitive tests and baseline age were associated with the largest increases in the C-index when added to the base model (i.e., SIPS P1+P2) (see Table 2).

As also shown in Table 2, in terms of individual variables, SIPS P1+P2 scores and baseline age bordered on statistical significance ($P=0.05$), while scores on the neurocognitive tests (Symbol Coding and HVLT) were marginally significant predictors ($P<.10$). A decline in social functioning and having a first-degree relative with psychosis were not significant predictors of psychosis in the validation sample ($P>.05$).

The NAPLS2 risk calculator was then used to provide probability estimates of conversion to psychosis for each individual in the external validation sample. Both the Mean Prediction Error (MPE) and Brier score were at acceptable levels ($<15\%$) (Brier score=7.5%, $SD=15.8$; MPE=9.5%, $SD=12.14$), suggesting that the NAPLS2 calculator provided a reasonable estimation of psychosis risk when comparing the risk prediction generated by the validation model vs. observed outcomes.

In addition, the risk estimates generated by the external validation model and the NAPLS2 psychosis risk calculator were strongly correlated ($r_s=0.66$, $p<0.001$), suggesting correspondence between the predicted risks of both models. There was also strong agreement between the predicted levels of risk generated by the NAPLS2 calculator and the EDIPPP validation model (Gamma=0.7, $p<.001$).

Table 3 shows the performance of the NAPLS2 calculator when applied to the external sample across increasing levels of model predicted risk. The sensitivity and specificity

values for each threshold were comparable to those found in the development sample. For example, 10% model predicted risk provided a sensitivity of 91% and specificity of 37% with the EDIPPP sample versus 94.1% sensitivity and 23.6% specificity in the NAPLS2 sample. A model predicted risk of 20% provides a better balance between sensitivity and specificity levels at 58.3% and 72.6, respectively, which is again similar to the NAPLS2 model (67.7% sensitivity, 72.1% specificity).

Discussion

The present study represents a critical external validation of the first major risk calculator developed in the field of mental health to estimate the probability that a given individual will develop a psychotic disorder within a 2 year period. The six major risk factors from the NAPLS2 calculator, including baseline age, unusual thought content and suspiciousness, family history of a psychotic disorder, verbal learning, processing speed performance and social decline, were able to separate individuals with and without psychosis with a good degree of accuracy in the EDIPPP validation sample. In addition, there was good agreement between the risk prediction from the NAPLS2 model and observed outcomes in the EDIPPP sample. Overall, this novel approach to constructing a psychosis prediction model using theoretical predictors has been validated in two independent CHR samples: the development sample initially used to test the theoretical model (NAPLS) and on an external, independent CHR sample (EDIPPP). To the best of our knowledge, this type of validation has not been performed on a psychosis prediction model and provides a critical first step in the introduction of the NAPLS psychosis risk calculator for widespread use.

Given the availability of risk calculators for numerous medical conditions, it is somewhat surprising that a tool for a psychiatric disorder has taken so long to be developed. Lack of replications of well-performing models and complex biological findings with limited clinical applicability may have contributed to this delay(43). Our findings highlight the importance of building a predictor model that includes a set of theoretically derived risk factors that have strong ties to the vulnerability to the disease and can easily applied in a clinical setting. The performance of the independent prediction model built with the EDIPPP sample using the risk factors included in the NAPLS2 2 calculator showed good discrimination ability comparable with that of the original development cohort with overall model accuracy rates of 79% and 72%, respectively. Moreover, the diagnostic accuracy (i.e., sensitivity, specificity) for the levels of predicted risk corresponded to that in the development model for the range of predicted risks that are adequately represented. The PPV values found in the present study were lower, however, than those found in the development sample due to the lower conversion rate (i.e., prevalence) in the validation sample.

The discrimination accuracy of the base model (SIPS P1+P2) was improved by almost 12% with the addition of the other four variables. This provides further evidence that a combination of variables can discriminate among CHR samples better than any individual predictor. It also potentially protects against a Type II error (i.e., missing a true difference with a smaller number of factors). In contrast to the development model, social functioning decline and neurocognition were not significant predictors of psychosis in the external validation sample. This may be related to sample size differences between the validation and

development cohorts rather than the inability of any single risk factor to predict psychosis. The younger age of this sample and a higher proportion of school-based ascertainment could also account for the reduced predictive power of social functional decline, since the younger participants would have had less time to deteriorate. Overall, these results suggest that the performance of the model is driven by the six risk factors working in concert to predict psychosis.

Interpreting Psychosis Risk

In continuing to establish the validity of the risk calculator, a number of additional issues and caveats must be considered. First, and perhaps foremost, the calculator should only be used by a mental health professional trained to a reliability standard in identifying COPS criteria with the SIPS/SOPS and administering neuropsychology and other clinical measures with good reliability. Second, as with any medical risk calculator, psychosis risk estimates provide a relative probability that illness will develop in the future, but not inevitability. Third, there is a risk of an inaccurate prediction. To mitigate this risk, in addition to reporting the exact estimate, the clinician should discuss the cost-benefit ratio of treatment with the patient in terms of the different levels of risk as shown in Table 3. This discussion of the risk/benefit ratio should be part of the first step in an informed decision between the clinician and the patient(44). Risk should not be overestimated, but, at the same time, these estimates convey important information, for example the difference between 5% vs. 25% risk should meaningfully impact treatment decisions(45). Fourth, and of particular relevance to intervention, treatment recommendations should take into account the possible adverse effects and: 1. The magnitude of the estimate (i.e., high vs. low risk); and 2. The particular predictor(s) that are driving the estimate. Low-risk individuals, for example, can be offered a less invasive treatment. Individuals with substantial neurocognitive difficulties could be offered, for example, cognitive training.(46) On the other hand, high-risk individuals and/or individuals with more severe positive symptoms would potentially be offered more aggressive intervention, possibly involving medications. Finally, since the field of psychosis prevention is constantly evolving, an accurate assessment of risk requires constant updating to take into account other factors not included in the prediction model that may alter risks/benefits(1).

Next Steps: Integration into to Clinical Practice

The NAPLS2 psychosis risk calculator represents a major advance toward achieving the goal of personalized medicine in psychiatry. According to the guidelines reported by McGinn et al.(21), four steps are involved in establishing a validated predictive tool and decision rules for use in clinical practice: selecting variables (level 4), validation at same site or small prospective sample (level 3), validation in different sites (level 2), and then evaluating impact on clinical practice (level 1). Our findings provide preliminary evidence for level 2 validation according to this schema. The critical last step in recommending the psychosis risk calculator (Level 1 validation) in a wide variety of settings would require an analysis of the impact of the tool in clinical practice(21). This involves demonstrating that the prediction tool can change clinician behavior as well as benefit patient outcomes.

Limitations

Our findings should be considered in the context of certain limitations. First, although we found strong evidence of the applicability of the NAPLS2 2 psychosis risk calculator in an independent CHR sample, the performance of the risk calculator needs further replications. The performance of the model should be evaluated on sub-cohorts (e.g., sociodemographically or clinically defined) and longer-term outcomes. In addition, the calculator should be continually updated and fine-tuned as new biological markers emerge(47). It should be noted that the optimal range for maximizing sensitivity and specificity lies between 15% (75.0 sensitivity, 58.5 specificity) and 25% (50.0 sensitivity, 81.7 specificity). However, additional prospective studies using the psychosis risk calculator to predict later illness are needed to validate this range as the appropriate target for intervention. Second, it is currently unclear how the NAPLS2 risk calculator would perform on CHR cohorts obtained outside of North America, such as populations recruited in the European and Australian high-risk projects. Finally, as with any risk calculator, the accuracy of the psychosis risk estimate is dependent on valid and accurate data.

Conclusions

The data reported in this study have shown that the performance of the NAPLS2 psychosis risk calculator, available on the Internet and incorporating a set of theoretically derived risk factors, can be replicated in a separate, independently collected CHR population. Although further replication is needed, at present, the risk calculator appears to have considerable potential for determining the probability that an individual will develop psychosis and may provide a foundation for the personalized treatment of CHR individuals.

Acknowledgments

Funding: Supported by grants MH61523 and MH081857 from the National Institute of Mental Health (Cornblatt), NIH 5KL2TR000434-08 (Tso), 5R01MH059883-11 (Carter), R21MH101676 (Taylor), K23MH087708 (Niendam) and the Zucker Hillside Hospital NIMH Advanced Center for Intervention and Services Research for the Study of Schizophrenia MH074543 (John M Kane, M.D.). Robert Wood Johnson Foundation (#67525) with additional institutional support from the Maine Medical Center Research Institute and the State of Maine.

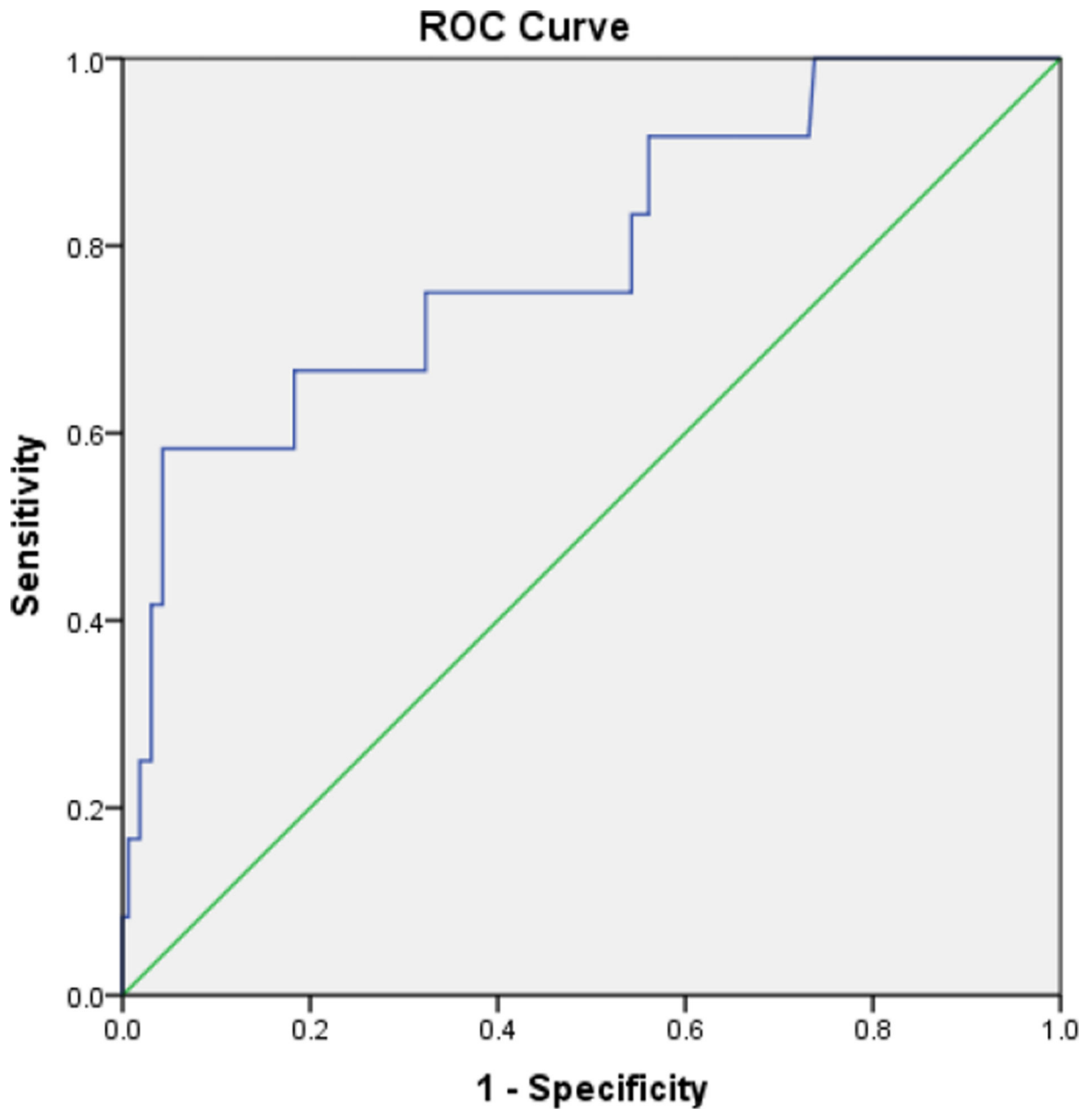
References

1. Sniderman AD, D'Agostino RB Sr, Pencina MJ. The Role of Physicians in the Era of Predictive Analytics. *JAMA*. 2015; 314:25–26. [PubMed: 26151261]
2. Moons KG, Royston P, Vergouwe Y, Grobbee DE, Altman DG. Prognosis and prognostic research: what, why, and how? *BMJ*. 2009; 338:b375. [PubMed: 19237405]
3. D'Agostino RB Sr, Pencina MJ, Massaro JM, Coady S. Cardiovascular Disease Risk Assessment: Insights from Framingham. *Glob Heart*. 2013; 8:11–23. [PubMed: 23750335]
4. Hippisley-Cox J, Coupland C. Development and validation of risk prediction algorithms to estimate future risk of common cancers in men and women: prospective cohort study. *BMJ Open*. 2015; 5:e007825.
5. Nam RK, Toi A, Klotz LH, Trachtenberg J, Jewett MA, Appu S, Loblaw DA, Sugar L, Narod SA, Kattan MW. Assessing individual risk for prostate cancer. *J Clin Oncol*. 2007; 25:3582–3588. [PubMed: 17704405]
6. Wells BJ, Kattan MW, Cooper GS, Jackson L, Koroukian S. Colorectal cancer predicted risk online (CRC-PRO) calculator using data from the multi-ethnic cohort study. *J Am Board Fam Med*. 2014; 27:42–55. [PubMed: 24390885]

7. Hippisley-Cox J, Coupland C, Robson J, Sheikh A, Brindle P. Predicting risk of type 2 diabetes in England and Wales: prospective derivation and validation of QDScore. *BMJ*. 2009; 338:b880. [PubMed: 19297312]
8. Lindstrom J, Tuomilehto J. The diabetes risk score: a practical tool to predict type 2 diabetes risk. *Diabetes Care*. 2003; 26:725–731. [PubMed: 12610029]
9. Ridker PM, Buring JE, Rifai N, Cook NR. Development and validation of improved algorithms for the assessment of global cardiovascular risk in women: the Reynolds Risk Score. *JAMA*. 2007; 297:611–619. [PubMed: 17299196]
10. Ridker PM, Paynter NP, Rifai N, Gaziano JM, Cook NR. C-reactive protein and parental history improve global cardiovascular risk prediction: the Reynolds Risk Score for men. *Circulation*. 2008; 118:2243–2251. 2244p following 2251. [PubMed: 18997194]
11. Pencina MJ, D'Agostino RB Sr, Larson MG, Massaro JM, Vasan RS. Predicting the 30-year risk of cardiovascular disease: the framingham heart study. *Circulation*. 2009; 119:3078–3084. [PubMed: 19506114]
12. Insel TR. Assessing the economic costs of serious mental illness. *Am J Psychiatry*. 2008; 165:663–665. [PubMed: 18519528]
13. 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, Klosterkötter J, McGuire P, Yung A. The psychosis high-risk state: a comprehensive state-of-the-art review. *JAMA Psychiatry*. 2013; 70:107–120. [PubMed: 23165428]
14. Cannon TD, Cadenhead K, Cornblatt B, Woods SW, Addington J, Walker E, Seidman LJ, Perkins D, Tsuang M, McGlashan T, Heinssen R. Prediction of psychosis in youth at high clinical risk: a multisite longitudinal study in North America. *Arch Gen Psychiatry*. 2008; 65:28–37. [PubMed: 18180426]
15. Nelson B, Yuen HP, Wood SJ, Lin A, Spiliotacopoulos D, Bruxner A, Broussard C, Simmons M, Foley DL, Brewer WJ, Francey SM, Amminger GP, Thompson A, McGorry PD, Yung AR. Long-term follow-up of a group at ultra high risk ("prodromal") for psychosis: the PACE 400 study. *JAMA Psychiatry*. 2013; 70:793–802. [PubMed: 23739772]
16. Ruhrmann S, Schultze-Lutter F, Salokangas RK, Heinimaa M, Linszen D, Dingemans P, Birchwood M, Patterson P, Juckel G, Heinz A, Morrison A, Lewis S, von Reventlow HG, Klosterkötter J. Prediction of psychosis in adolescents and young adults at high risk: results from the prospective European prediction of psychosis study. *Arch Gen Psychiatry*. 2010; 67:241–251. [PubMed: 20194824]
17. Cornblatt BA, Carrion RE, Auther A, McLaughlin D, Olsen RH, John M, Correll CU. Psychosis Prevention: A Modified Clinical High Risk Perspective From the Recognition and Prevention (RAP) Program. *Am J Psychiatry*. 2015; 172:986–994. [PubMed: 26046336]
18. Perkins DO, Jeffries CD, Cornblatt BA, Woods SW, Addington J, Bearden CE, Cadenhead KS, Cannon TD, Heinssen R, Mathalon DH, Seidman LJ, Tsuang MT, Walker EF, McGlashan TH. Severity of thought disorder predicts psychosis in persons at clinical high-risk. *Schizophr Res*. 2015
19. Michel C, Ruhrmann S, Schimmelmann BG, Klosterkötter J, Schultze-Lutter F. A stratified model for psychosis prediction in clinical practice. *Schizophr Bull*. 2014; 40:1533–1542. [PubMed: 24609300]
20. Addington J, Cadenhead KS, Cornblatt BA, Mathalon DH, McGlashan TH, Perkins DO, Seidman LJ, Tsuang MT, Walker EF, Woods SW, Addington JA, Cannon TD. North American Prodrome Longitudinal Study (NAPLS 2): overview and recruitment. *Schizophr Res*. 2012; 142:77–82. [PubMed: 23043872]
21. McGinn TG, Guyatt GH, Wyer PC, Naylor CD, Stiell IG, Richardson WS. Users' guides to the medical literature: XXII: how to use articles about clinical decision rules. Evidence-Based Medicine Working Group. *JAMA*. 2000; 284:79–84. [PubMed: 10872017]
22. McFarlane WR, Stastny P, Deakins S. Family-aided assertive community treatment: a comprehensive rehabilitation and intensive case management approach for persons with schizophrenic disorders. *New Dir Ment Health Serv*. 1992:43–54. [PubMed: 1579118]

23. Fjell A, Bloch Thorsen GR, Friis S, Johannessen JO, Larsen TK, Lie K, Lyse HG, Melle I, Simonsen E, Smeby NA, Oxnevad AL, McFarlane WR, Vaglum P, McGlashan T. Multifamily group treatment in a program for patients with first-episode psychosis: experiences from the TIPS project. *Psychiatr Serv.* 2007; 58:171–173. [PubMed: 17287370]
24. McFarlane WR, Levin B, Travis L, Lucas FL, Lynch S, Verdi M, Williams D, Adelsheim S, Calkins R, Carter CS, Cornblatt B, Taylor SF, Auther AM, McFarland B, Melton R, Migliorati M, Niendam T, Ragland JD, Sale T, Salvador M, Spring E. Clinical and functional outcomes after 2 years in the early detection and intervention for the prevention of psychosis multisite effectiveness trial. *Schizophr Bull.* 2015; 41:30–43. [PubMed: 25065017]
25. McFarlane WR, Cook WL, Downing D, Ruff A, Lynch S, Adelsheim S, Calkins R, Carter CS, Cornblatt BA, Milner K. Early Detection, Intervention, and Prevention of Psychosis Program: Rationale, Design, and Sample Description. *Adolescent Psychiatry.* 2012; 2:112–124.
26. Lynch S, McFarlane WR, Joly B, Adelsheim S, Auther AM, Cornblatt B, Downing D, Migliorati M, Ragland JD, Sale T, Spring E, Calkins R, Carter CS, Jaynes R, Taylor SF. Early Detection for the Prevention of Psychosis Program: Community Outreach and Early Identification at Six Sites across the United States. *Psychiatric Services.* *Psychiatric Services.* In Press.
27. McFarlane WR, Cook WL, Downing D, Verdi MB, Woodberry KA, Ruff A. Portland identification and early referral: a community-based system for identifying and treating youths at high risk of psychosis. *Psychiatr Serv.* 2010; 61:512–515. [PubMed: 20439374]
28. Miller TJ, McGlashan TH, Rosen JL, Cadenhead K, Cannon T, Ventura J, McFarlane W, Perkins DO, Pearlson GD, Woods SW. 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.* 2003; 29:703–715. [PubMed: 14989408]
29. Miller TJ, McGlashan TH, Rosen JL, Somjee L, Markovich PJ, Stein K, Woods SW. Prospective diagnosis of the initial prodrome for schizophrenia based on the Structured Interview for Prodromal Syndromes: preliminary evidence of interrater reliability and predictive validity. *Am J Psychiatry.* 2002; 159:863–865. [PubMed: 11986145]
30. Miller TJ, McGlashan TH, Woods SW, Stein K, Driesen N, Corcoran CM, Hoffman R, Davidson L. Symptom assessment in schizophrenic prodromal states. *Psychiatr Q.* 1999; 70:273–287. [PubMed: 10587984]
31. Cornblatt BA, Auther AM, Niendam T, Smith CW, Zinberg J, Bearden CE, Cannon TD. Preliminary findings for two new measures of social and role functioning in the prodromal phase of schizophrenia. *Schizophr Bull.* 2007; 33:688–702. [PubMed: 17440198]
32. Green MF, Nuechterlein KH, Gold JM, Barch DM, Cohen J, Essock S, Fenton WS, Frese F, Goldberg TE, Heaton RK, Keefe RSE, Kern RS, Kraemer H, Stover E, Weinberger DR, Zalcman S, Marder SR. Approaching a consensus cognitive battery for clinical trials in schizophrenia: The NIMH-MATRICES conference to select cognitive domains and test criteria. *Biological Psychiatry.* 2004:301–307. [PubMed: 15336511]
33. Keefe, R. Brief Assessment of Cognition in Schizophrenia (BACS) Manual—A: Version 2.1. Durham, NC: Duke University Medical Center; 1999.
34. Benedict RHB, Schretlen D, Groninger L, Brandt J. Hopkins Verbal Learning Test – Revised: Normative Data and Analysis of Inter-Form and Test-Retest Reliability. *The Clinical Neuropsychologist.* 1998; 12:43–55.
35. Carrión RE, McLaughlin D, Goldberg TE, Auther AM, Olsen RH, Olvet DM, Correll CU, Cornblatt Ba. Prediction of functional outcome in individuals at clinical high risk for psychosis. *JAMA psychiatry.* 2013; 70:1133–1142. [PubMed: 24006090]
36. Cornblatt BA, Carrión RE, Addington J, Seidman L, Walker EF, Cannon TD, Cadenhead KS, McGlashan TH, Perkins DO, Tsuang MT, Woods SW, Heinssen R, Lencz T. Risk Factors for Psychosis: Impaired Social and Role Functioning. *Schizophr Bull.* 2011
37. Hanley JA, McNeil BJ. The meaning and use of the area under a receiver operating characteristic (ROC) curve. *Radiology.* 1982; 143:29–36. [PubMed: 7063747]
38. Harrell FE Jr, Lee KL, Mark DB. Multivariable prognostic models: issues in developing models, evaluating assumptions and adequacy, and measuring and reducing errors. *Stat Med.* 1996; 15:361–387. [PubMed: 8668867]

39. Siegel, S.; Castellan, NJJ. *Nonparametric statistics for the behavioral sciences*. 2nd. New York, NY: McGraw-Hill; 1988.
40. Brier GW. Verification of forecasts expressed in terms of probability. *Monthly weather review*. 1950; 78:1–3.
41. Sheiner LB, Beal SL. Some suggestions for measuring predictive performance. *J Pharmacokinetic Biopharm*. 1981; 9:503–512. [PubMed: 7310648]
42. Steyerberg, E. *Clinical Prediction Models: A Practical Approach to Development, Validation, and Updating*. New York: Springer; 2008.
43. Kapur S, Phillips AG, Insel TR. Why has it taken so long for biological psychiatry to develop clinical tests and what to do about it? *Mol Psychiatry*. 2012; 17:1174–1179. [PubMed: 22869033]
44. Stone NJ, Robinson JG, Lichtenstein AH, Bairey Merz CN, Blum CB, Eckel RH, Goldberg AC, Gordon D, Levy D, Lloyd-Jones DM, McBride P, Schwartz JS, Shero ST, Smith SC Jr, Watson K, Wilson PW, Eddleman KM, Jarrett NM, LaBresh K, Nevo L, Wnek J, Anderson JL, Halperin JL, Albert NM, Bozkurt B, Brindis RG, Curtis LH, DeMets D, Hochman JS, Kovacs RJ, Ohman EM, Pressler SJ, Sellke FW, Shen WK, Smith SC Jr, Tomaselli GF. American College of Cardiology/American Heart Association Task Force on Practice G: 2013 ACC/AHA guideline on the treatment of blood cholesterol to reduce atherosclerotic cardiovascular risk in adults: a report of the American College of Cardiology/American Heart Association Task Force on Practice Guidelines. *Circulation*. 2014; 129:S1–S45. [PubMed: 24222016]
45. Van Calster B, Steyerberg EW, Harrell FH. Risk Prediction for Individuals. *JAMA*. 2015; 314:1875. [PubMed: 26529169]
46. Loewy R, Fisher M, Schlosser DA, Biagianni B, Stuart B, Mathalon DH, Vinogradov S. Intensive Auditory Cognitive Training Improves Verbal Memory in Adolescents and Young Adults at Clinical High Risk for Psychosis. *Schizophr Bull*. 2016
47. DeFilippis AP, Young R, Carrubba CJ, McEvoy JW, Budoff MJ, Blumenthal RS, Kronmal RA, McClelland RL, Nasir K, Blaha MJ. An analysis of calibration and discrimination among multiple cardiovascular risk scores in a modern multiethnic cohort. *Ann Intern Med*. 2015; 162:266–275. [PubMed: 25686167]



Diagonal segments are produced by ties.

Figure 1.

EDIPPP Model Discrimination with ROC Curve . The ROC curve plots the true positive rate (sensitivity) against the false-positive rate (1-specificity) for different cut-points. The more closely the curve follows the top and left-hand border of the ROC space, the more accurate the test. The area under the curve (AUC) with 95% confidence intervals was used as indicator of the probability that a randomly chosen respondent would be correctly distinguished based on the prediction model.

Table 1

Baseline demographics and clinical characteristics of EDIPPP CHR Sample

Characteristic	Followed (N=176)		Not Followed (N=34)		Statistics
	Mean	SD	Mean	SD	
Age (Years)	16.6	3.3	16.0	3.1	t=-0.86, p=0.39
Education (Years)	9.8	2.6	9.6	2.7	t=0.52, p=0.61
Modified SIPS P1+P2 ^a	2.4	1.9	2.4	2.1	t=-0.10, p=0.92
BACS Symbol Coding raw score correct	54.8	12.5	50.9	8.7	t=-2.11, p=0.04
HVLT Trials 1-3 summed	25.8	5.5	25.0	4.7	t=-0.78, p=0.44
	N	%	N	%	
Family History of Psychosis	37	22.8	5	17.2	$\chi^2=0.45$, p=0.50
Decline in GF:S score > 0	94	53.4	14	41.2	$\chi^2=1.71$, p=0.19
Male	103	58.5	19	55.9	$\chi^2=0.82$, p=0.78
Race					
White	109	64.1	20	66.7	$\chi^2=8.97$, p=0.19
Black	14	8.2	4	13.3	
Other	12	7.1	2	6.7	
Mixed	24	14.1	3	10.0	
Hispanic ethnicity	26	15.3	5	15.6	$\chi^2=0.00$, p=0.96

^aModified such that all levels in the non-prodromal range (0-2 on original scale) now defined as 0, levels in the prodromal range (3-5 on original scale) now defined as 1-3, and psychotic intensity (6 on original scale) now defined as 4.

Performance of key predictors from the NAPLS2 Psychosis Risk Calculator in the EDIPPP CHR sample

Table 2

Predictor	Multivariable Model			C-index ^a	
	Hazard Ratio	95% CI	p	Decrement if removed	Increase if added
Modified SIPS P1+P2	1.4	1.0–1.9	0.05	0.035	N/A ^b
Decline in GF:S score	0.9	0.5–1.7	0.73	0.000	0.005
HVLT Trials 1–3 summed	0.9	0.8–1.0	0.10	0.013	0.092
Symbol Coding raw score correct	1.0	0.9–1.0	0.08	–0.004	0.064
Age	1.2	1.0–1.4	0.05	0.027	0.050
Family History of Psychosis	1.7	0.4–6.6	0.50	–0.006	0.015

^aHarrell's C-index (equivalent to the area under the receiver operating characteristic curve or AUC) was used to quantify the discrimination ability for separating converters and non-converters. The C-index for the overall model was .790.

^bThe base model included only the modified SIPS P1+P2 scores; the C-index for the base model was 0.670.

Prediction statistics for conversion to psychosis across various levels of NAPLS2 model-predicted risk

Table 3

Predicted Risk % ^a	Base rate of Predicted Risk ^b	Positive Predictive Value ^c	Negative Predictive Value ^d	Sensitivity	Specificity
5	96.6	7.1	100.0	100.0	3.7
10	64.8	9.6	98.4	91.7	37.2
15	43.8	11.7	97.0	75.0	58.5
20	29.5	13.5	96.0	58.3	72.6
25	20.5	16.7	95.7	50.0	81.7
30	12.5	13.6	94.2	25.0	88.4
35	9.1	18.8	94.4	25.0	92.1
40	6.3	18.2	93.9	16.7	94.5
45	5.7	9.9	93.4	8.3	94.5
50	2.3	25.2	93.6	8.3	98.2
55	0.6	100.0	93.7	8.3	100.0
60	0.6	100.0	93.7	8.3	100.0

^a Each row and corresponding predicted risk represents an individualized prediction within a level or risk class.

^b Percentage of individuals with the predicted risk score at the specified level or higher.

^c $PPV = \text{Sensitivity} * \text{Prevalence} / (\text{Sensitivity} * \text{Prevalence} + (1 - \text{Specificity}) * (1 - \text{Prevalence}))$

^d $NPV = \text{Specificity} * (1 - \text{Prevalence}) / ((1 - \text{Sensitivity}) * \text{Prevalence} + \text{Specificity} * (1 - \text{Prevalence}))$