BQP Qualification Program Cover Letter

Date: March 20, 2020

Subject: DDT QUALIFICATION SUBMISSION

DDT Type: Biomarker Qualification

ATTN: CDER-Biomarker Qualification Program

C/O CDER Document Room: Upon receipt notify:
CDER-BiomarkerQualificationProgram@fda.hhs.gov

Biomarker DDT Tracking Number: (in bold print), if previously assigned

Check Here	Submission Type
✓	Letter of Intent
	Qualification Plan
	Full Qualification Package
	Update of Above (Check two, this box and one above)
	Other (please specify):

Biomarker Name(s): Individualized Risk Calculator for Psychosis (IRC-P)

Context of Use: Describe the intended drug development use for the biomarker named above (1 to 2 sentences, see the graphic below for how to write the context of use.)

Prognostic biomarker intended for use in clinical trials. It will be used in conjunction with clinical high-risk for psychosis (CHR-P) or Attenuated Psychosis Syndrome (APS) diagnosis in young people aged 15-35 years of age, to enrich for individuals most likely to progress to full psychosis and poor long-term functional outcomes.

Contact Information: Complete contact information including name(s), affiliation, mailing address, email address, phone and fax numbers.

1. **Tyrone Cannon, Principle Investigator, North American Prodrome Longitudinal Study (NAPLS) Research Consortium**
   Clark L. Hull Professor of Psychology and Professor of Psychiatry
   Yale University
   P.O. Box 208205, 2 Hillhouse Avenue, New Haven, CT 06520
   203-436-1545 tyrone.cannon@yale.edu

2. **Linda Brady, Co-chair, Biomarkers Consortium Neuroscience Steering Committee**
   Director, Division of Neuroscience and Basic Behavioral Science (DNBBS)
   6001 Executive Boulevard, Room 7204/MSC 9645, Bethesda, MD 20892

BQP Qualification Program Cover Letter

(301) 443-3563 lbrady@mail.nih.gov

Purpose Statement: Describe the purpose of the submission in 3-5 sentences.

To obtain formal feedback from the FDA regarding an individualized risk calculator for psychosis (IRC-P) as a prognostic biomarker. The IRC-P is proposed for use in future clinical trials and treatment studies to identify a subgroup of individuals meeting criteria for a CHR-P or APS diagnosis, that are most likely to progress to full psychosis and poor long-term functional outcomes (i.e., those most in need of treatment).

The replicability of the IRC-P as a prognostic biomarker is being assessed by a multi-consortia international effort called HARMONY (Harmonization of At Risk Multisite Observational Networks for Youth). NAPLS is collaborating with the PRONIA (https://www.pronia.eu) and PSYSCAN (http://psyscan.eu) consortia studies based in Europe that have collected similar information on large cohorts of CHR-P individuals. The purpose of this submission is to solicit feedback from the FDA regarding the viability of the IRC-P as a prognostic biomarker for use in clinical trials, the current plan for conducting confirmatory studies, and next steps for testing the robustness of particular thresholds of predicted risk as cut-points for maximizing sensitivity and specificity of psychosis prediction.

In addition, the applicant intends to explore re-analysis of prior clinical trial data on CHR-P samples incorporating individualized predicted risk as a factor in the statistical analysis, to determine whether treatment outcomes are moderated by level of predicted risk. If so, the resulting evidence may help to qualify the IRC-P for an expanded context of use in relation to treatment selection in a stepped care algorithm whereby the particular intervention (e.g., psychological, drug) is targeted/scaled to the individual's level of predicted risk.

Submission Statement: Include a statement in the cover letter that: "The physical media submission is virus free with a description of the software (name, version and company) used to check the files for viruses."

The physical media submission is virus free and has been checked for viruses with ESET Endpoint Antivirus software.

Additional Instructions for LOI/QP/FQP[1] submissions: For every electronic submission, a comprehensive table of contents should be submitted containing three or four levels of detail, with the appropriate bookmarks to key referenced sections in the document.

[1] LOI: Letter of Intent; QP: Qualification Plan; FQP: Full Qualification Plan

Biomarker Qualification Letter of Intent (LOI) Content Elements

NOTE TO REQUESTORS: FDA is currently developing its policies for submissions under the 21 Century Cures Act (section 507)[1] and expects to issue guidance to aid in the development of submission based on a decade of reviews, input from public meetings, comments to the docket and collaborative public partnerships. In the interim the Agency has assembled this resource to help requestors. Given the changes to the process as defined in section 507, we expect to see further development of this content over time, with more experience and your input. For additional resources on submission content please see prior Biomarker Qualification Program submissions that we have accepted under section 507 **HERE**. Please also note that certain information contained in submissions will be made publicly available as per section 507, as described in greater detail **HERE**.

Should you have any questions or want to provide feedback on this or other BQP resources, including the content and format of submissions and the transparency provisions under section 507, please contact us at CDER-BiomarkerQualificationProgram@fda.hhs.gov

COMMENTS: The following information will be made publicly available as per section 507, described in greater detail HERE

1. Section 3011 of the 21st Century Cures Act established section 507 of the Federal Food, Drug, and Cosmetic Act (FD&C Act).

Table of Contents

Administrative Information

1. **Submission Title**: **Individualized Risk Calculator for Psychosis (IRC-P)**
   One sentence description of your project. See Abbreviated Biomarker Descriptions in List of FDA Qualified Biomarkers. EXAMPLES:
   - Urinary nephrotoxicity biomarkers as assessed by immunoassays
   - Total Kidney Volume (TKV) as assessed by computerized tomography (CT) scan.

2. **Requesting Organization:**
   North American Prodrome Longitudinal Study (NAPLS) Research Consortium

   1. Primary contact
      Tyrone Cannon, Principle Investigator, North American Prodrome Longitudinal Study (NAPLS) Research Consortium
      Clark L. Hull Professor of Psychology and Professor of Psychiatry

Yale University
P.O. Box 208205, 2 Hillhouse Avenue, New Haven, CT 06520
203-436-1545
tyrone.cannon@yale.edu

2. Alternate contact
Linda Brady, Co-chair, Biomarkers Consortium Neuroscience Steering Committee
Director, Division of Neuroscience and Basic Behavioral Science (DNBBS)
6001 Executive Boulevard, Room 7204/MSC 9645, Bethesda, MD 20892
301-443-3563
lbrady@mail.nih.gov

3. **Submission Dates:**
LOI Submission Date: March 20, 2020

Drug Development Need Statement

Describe the drug development need that the biomarker is intended to address, including (if applicable) the proposed benefit over currently used biomarkers for similar context of uses (COUs).

Schizophrenia affects 1% of the population and is among the top 15 leading causes of disability worldwide.[1] The illness is associated with a reduction of 28.5 years in average potential life expectancy, with suicide and co-morbid medical conditions such as heart disease and diabetes contributing to premature mortality.[2] Currently available drug treatments for schizophrenia and related disorders (i.e., antipsychotics) are limited in efficacy and poorly tolerated, with most patients showing substantial residual symptomatology and continuing functional disability as well as significant side effect burdens.[3] Most in the psychosis field now believe that drugs antagonizing the dopamine system (as all current antipsychotics do) are not capable of treating the full array of symptoms, cognitive deficits, and functional impariments associated with these illnesses.[4] Accordingly, recent efforts in drug development for schizophrenia and related disorders have focused on targeting mechanisms thought to be involved in the onset and/or progression of psychosis, including NMDA receptor function, GABA interneuron function, oxidative stress, and complement-related pathways.[5] In the framework of a secondary prevention strategy, if such treatments could be applied before full illness takes hold, it may be possible to alter the trajectories of at risk individuals toward less symptomatic and more independent functional outcomes.[6] However, a key rate-limiting step to realizing the potential benefits of such early interventions is the availability of reliable and valid strategies for ascertaining individuals at greatest risk.

For the majority of patients with schizophrenia, onset of fully psychotic symptoms is preceded by the emergence of subtler changes in belief, thought, and perception that appear to represent attenuated forms of delusions, formal thought disorder, and hallucinations, respectively.[7,8] The Structured Interview for Prodromal Risk Syndromes (SIPS)[9] is a well validated instrument for ascertaining individuals with a clinical high-risk for psychosis (CHR-P) syndrome, defined by the recent emergence of attenuated or sub-threshold psychotic symptoms. These criteria are also encapsulated in the DSM-5's provisional diagnostic category *Attenuated Psychosis Syndrome*, or APS. Individuals meeting CHR-P or APS criteria are distressed and seeking treatment[10]. Although by definition their positive symptoms

(i.e., delusions, hallucinations, thought disorder) are at sub-psychotic intensity, these symptoms are nevertheless disruptive and rate-limiting for social and role functioning[11], on average at about the level associated with major depressive disorder with comorbid alcohol abuse.[12] About 15% of individuals meeting CHR-P/APS criteria transition to a fully psychotic form of mental illness within 2-years of initial ascertainment.[10,13,14] Importantly, among the cases who do not convert, about one-fourth remit the symptoms that indexed their initial risk status during the course of follow-up, while the remaining three-fourths show continuing levels of attenuated psychotic-like symptoms and functional impairments.[15-17] Thus, application of effective interventions in this population would be expected to result in significant reductions in contemporaneous disease burden as well as potentially preventing progression to more severe, chronic, and debilitating forms of illness.

Given that meeting criteria for a CHR-P or APS diagnosis implies a 15% risk of developing fully psychotic symptoms over a 2-year period,[18] enrolling such cases in clinical trials with novel treatments targeting mechanisms involved in psychosis onset/progression represents an emerging strategy for drug developers, with at least one pharmaceutical company in the midst of an active Phase II trial and several others in planning stages. However, this 15% predicted risk applies at the group rather than individual patient level. Given heterogeity in the risk factors and outcomes among cases with a CHR-P or APS diagnosis, clinical trials would be greatly facilitated by the availability of more precise methods of prediction that can scale the degree of risk at the individual case level. In the context the North American Prodrome Longitudinal Study (NAPLS),[19] we have developed such an individualized risk calculator that predicts the likehihood that a given patient meeting CHR-P or APS criteria will transition to full psychosis within a given time interval.[10] Use of this calculator is expected to improve the efficiency of clinical trials by enriching samples for CHR-P/APS cases with higher predicted risks of progression to more severe, chronic, and debilitating forms of illness.

Biomarker Information and Interpretation

Please provide high level descriptions here and more detailed descriptions in the analytical and the clinical considerations sections.

1. **Biomarker name:** abbreviated short name for biomarker, or names if multiple, AND identify each biomarker type (molecular, histologic, radiographic, or physiologic characteristics according to BEST Glossary). For molecular biomarkers, please provide a unique molecular ID e.g. from UniProt (http://uniprot.org/), HUGO Gene Nomenclature Committee (http://genenames.org), Protein Data Bank (http://rcsb.org/pdb/home/home.do), or Enzyme Commission (http://enzyme.expasy.org).

   EXAMPLES: 25 mRNA gene expression profile/signature; cardiac Troponins T (cTnT) and I (cTnI); Total Kidney Volume (TKV) (please note detection method or algorithm is not a part of the biomarker name). For more examples see the "Qualified Biomarker" column on the FDA List of Qualified Biomarkers website.

   Individualized Risk Calculator for Psychosis (IRC-P)

2. **Analytical methods:** name and briefly describe analytical methods used in raw measurement(s) of the biomarker(s). EXAMPLE: enzyme-linked immunosorbent assay (ELISA) with chromogenic reporters,

volumetric analysis of brain magnetic resonance images (MRIs). Include all elements counted/measured/identified and indicate whether measurement is a manual read or a component of the analytic.

Interview-based ratings of symptom severity, social functioning, family history of psychosis, childhood traumas, and stressful life events.
Raw scores on two standardized paper-and-pencil neurocognitive tests.

3. **Measurement units and limit(s) of detection:** describe if any.
Cox proportional hazard model regression/Probability value. The algorithm computes a 1- or 2-year probability of converting to full psychosis, with the resulting predicted probability potentially varying between 0.1% to 99.0%.

4. **Biomarker interpretation and utility**
Describe the application/conversion of the raw biomarker measurement in order for the biomarker outcomes to be used for the COU and provide the description and derivation of clinical interpretative criteria used to include:
    a. Post-analytical application/conversion of biomarker raw measure to the applied measure: briefly describe how the raw biomarker measurement is used/applied. Describe if the raw measure is used directly or if there is further processing of the raw measurement into a multi-component panel, a scoring system, or alternatively; further manipulation or transformation of the raw biomarker measurement using modeling, simulation, application of formula(e), other algorithms, or combination with other clinical information. Describe how the process is designed, including software. List the elements, inputs and output(s) of the conversion, including a description of units, if applicable.
    b. Describe rationale for post-analytical elements used as inputs in application or conversion of the raw biomarker measurement.
    c. Clinical Interpretive Criteria: describe the cut-off values, cut-points/thresholds, boundaries/limits or other comparators used in the interpretation of the biomarker measurement or its applied/converted form to draw an actionable conclusion based on the biomarker result.

Biomarker measurement

The IRC-P produces an individualized predicted risk of conversion to a fully psychotic form of mental illness over a 1- or 2-year period among individuals meeting criteria for a CHR-P or APS diagnosis, based on their scores on a set of clinical and neurocognitive measures. Examples of individual predictor variables include clinician-rated measures of attenuated psychotic symptom severity, neuropsychological test scores indexing memory functioning and neurocognitive processing speed, and interview-based determinations of family history of psychosis and stress exposures. Formulae (derived from Cox proportional hazards regression models on a large reference dataset) are used to weight the individual predictor variables and combine them into probability values reflecting the likelihood of a newly ascertained case progressing to full psychosis within 1 or 2 years from the point of ascertainment. Probabilities for two reference periods (i.e., 1 or 2 years) are provided to allow for use in clinical trials of varying lengths. Based on the distribution of predicted risks in the reference sample, selecting new CHR-P/APS cases with predicted risks of 20% or higher will result in sample with an overall conversion rate approximately double that associated with CHR-P/APS status alone,[18] thereby

doubling the efficiency of clinical trials and focusing drug development efforts on cases most in need.

Context of Use Statement (500 characters)

The proposed context of use (COU) statement is complementary to the drug development need statement. Please note that we qualify biomarkers as tools to aid in drug development. While biomarkers may be used for other purposes (e.g., to aid in clinical decision making), COUs that do not address a specified drug development use are outside the scope of the program.

The COU statement may evolve over time based on the information presented in submissions supporting the biomarker's COU and the recommendations made by FDA. However, it should be consistent and worded identically throughout the given version of the submission document. Describing the COU statement early defines the type of information needed in support of qualification for the proposed approach. Although the eventual scope of the project may span over multiple COUs, only a single COU should be initially articulated for a given biomarker qualification submission. Recommended structures of the COU statement are provide below:

BEST biomarker category to **drug development use.**

Or

BEST biomarker category that action, i.e., selects or enriches or indicates or identifies purpose of intervention, e.g., severity, toxicity, susceptibility, disease progression or pharmacodynamic response of target populations, e.g., disease name/stage, patients responsive to treatment in type of study, e.g., early phase trials

EXAMPLES:

A. PD/response biomarker that measures Crohn's Disease (CD) activity used as a co-primary endpoint in CD clinical trials in conjunction with an accepted assessment of patient reported symptoms.

B. Susceptibility/risk biomarker that indicates the potential for individuals to develop symptomatic Type 1 Diabetes (T1D) to study interventions intended to prevent the onset of T1D.

Additional examples of COU statements are available on the Biomarker Qualification Submissions and the Qualified Biomarkers web pages. If assistance in identification of the most appropriate biomarker category is needed, a requestor may contact the Biomarker Qualification Program at CDER-BiomarkerQualificationProgram@fda.hhs.gov.

Prognostic biomarker intended for use in clinical trials. It will be used in conjunction with clinical high-risk for psychosis (CHR-P) or Attenuated Psychosis Syndrome (APS) diagnosis in young people aged 15-35 years of age, to enrich for individuals most likely to progress to full psychosis and poor long-term functional outcomes.

Note: This proposed COU is in keeping with the designation of CHR-P/APS as a clinical syndrome within the psychosis spectrum that may or may not progress to a more severe form of affection status.

Analytical Considerations

Please provide the following information (if applicable or available):

- General description of what aspect of the biomarker is being measured and by what method (e.g., lesion number or specific measure of organ size by imaging, serum level of an analyte, change in the biomarker level relative to a reference such as baseline).
- If this biomarker involves an index/scoring system, please provide information about the elements and weighting of the elements. Include a rational for how the index/scoring system was developed.
- Brief description of sample source, matrix (base material and any additives), stability and composition of biomarker.
- Description of pre-analytical factors and quality assurance/quality control (QA/QC) plans to preserve specimen integrity: a standard operating procedure (SOP) for sample collection including timing and location that sample will be collected from, storage *and* test/assay methodology; reference or control samples.
- Analytical validation plan: description of measurement tool and device calibrations, validation study design with statistical analysis plan (SAP) or performance data (e.g. sensitivity, specificity, accuracy, and/or precision of the assay or method).
- Once the SOP and analytical validation plan is finalized, describe how you will use this process to validate the final version of the measurement tool.
- Additional considerations for imaging biomarkers:
  - How has the method for image acquisition, analysis, and integration of the data been optimized?
  - Does data currently exist to support the proposed cut-off point(s), if imaging results are not reported as a continuous variable?
  - Provide the name and version of the software package to be used for image acquisition and analysis.
  - Description of any software or algorithm used to delineate or segment any physiological structure (i.e. a volume of an organ, a sub-section of an organ, or a size of a vein or opening etc.)
  - Describe any interpretation or transformation of the image data that will be conducted to measure, define, or represent the biomarker in question
  - Provide information on inter-operator and intra-operator variability.

Biomarker description

Variables considered as candidates for inclusion in the risk calculator were limited to those for which at least two prior studies had established an empirical association with psychosis prediction in CHR-P samples and had to be readily obtainable in standard clinical settings. Variables were chosen blindly with respect to their performance within the reference dataset to avoid overfitting and enable the resulting model to function as an unbiased estimator of predicted risks for future cases. The raw input variables (with corresponding VARIABLE LABELS) are:

1. **age at ascertainment;**
2. **processing speed as indexed by the score on the BACS Symbol-Coding Test (BACS);**
3. **verbal learning and memory functioning as indexed by the sum of learning trials 1, 2 and 3 on the Hopkins Verbal Learning Test-Revised (HVLT);**

4. **decline in social functioning in the year prior to ascertainment assessed using the Global Scale of Functioning-Social (GFS);**
5. **family history of psychotic disorder in a first-degree relative, yes (=1) or no (=0);**
6. **sum of severity ratings (rescaled to reflect severity only in the prodromal to fully psychotic range) for item P1 Unusual Thought Content and item P2 Suspiciousness on the Scale for the Assessment of Prodromal Symptoms;**
7. **stressful life events, as represented by the sum of 35 possible life events designated as negative and potentially relevant to subjects aged 12-35 years from the Psychiatric Epidemiology Research Interview – Life Events Scale; and**
8. **childhood traumas, as represented by the score on the Childhood Trauma and Abuse Scale.**

These raw input variables are weighted according to their strength of association with psychosis conversion in a reference sample (NAPLS2, N=596 CHR-P cases, including 84 converters) based on a multivariate proportional hazards model.[10] The resulting 1- or 2-year predicted risk is expressed as a probability value, according the following formulae:

- **1-year probability of conversion to psychosis = 1 – (0.9022808^exp(lp))**
- **2-year probability of conversion to psychosis = 1 – (0.8706555^exp(lp))**

Where lp = 1.5251292 – (0.043941454 * AGE) – (0.013755692 * BACS) – (0.03796384 * HVLT) + (0.026273078 * LIFE_EVENTS) – (0.00064476875 * TRAUMA) + (0.20021491 *GFS_DECLINE) + (0.14633391 * FAMILY_HISTORY) + (0.31427037 * P1P2)

A standard operating procedure (SOP) for sample collection, storage and test/assay methodology

A clinical evaluation including an interview and brief neurocognitive testing is required to qualify the patient as CHR-P and to generate the scores needed as input for the prediction algorithm. The patient is first interviewed using SIPS[9] or comparable instrument validated for use in ascertaining individuals meeting CHR-P or APS criteria. The input variables for the IRC-P derive from information gleaned from the SIPS,[9] from brief interview-based assessments of social functioning (Global Scale of Functioning-Social[20]), stressful life events (Psychiatric Epidemiology Research Interview – Life Events Scale[21]), and trauma (Childhood Trauma and Abuse Scale[22]), and from scores on two brief paper-and-pencil neurocognitive tests (Behavior Assessment of Cognition in Schizophrenia – Symbol Coding[23] and Hopkins Verbal Learning Test – Revised[24]). These instruments are widely used in clinical practice as well as clinical trials in psychiatry. The individual performing these assessments requires training in their administration, but training on these or similar instruments is typical for mental health professionals with masters degrees or higher, as would be customary in psychiatric clinical trials.

Analytical validation plan

Based on a bootstrap internal validation with 1000 resamples, the IRC-P achieved a Concordance Index[25] (analogous to area-under-receiver-operating-curve statistic) of 0.71 in the NAPLS2 sample and has been subsequently validated in several independent cohorts.[26,27] In the NAPLS2 sample, the calibration plot revealed a high degree of consistency between observed probabilities and model-predicted probabilities of conversion to psychosis within the range of 0.0-0.4, within which 95% of the cases fell (mean = 0.18, SD = 0.11; median = 0.16). Based on the NAPLS2 sample, a predicted risk

threshold of 20% or higher yields a sensitivity of 67% and a specificity of 72% in relation to psychosis conversion.

Clinical Considerations

Please provide the following information (if applicable or available):

- Describe how the biomarker measurement is used to inform drug development. Please provide a decision tree to guide how the biomarker information would be used in drug development or a clinical trial.
- Describe patient population or drug development setting in which the biomarker will be used.
- Clinical validation: provide information to support biological and clinical relevance of the biomarker as applied in the COU:
    - Describe how normal or other reference values are established, provide study design(s), analytical plan, etc.
- Benefits and Risks of applying the biomarker in drug development or a clinical trial.
- Describe any current knowledge gaps, limitations and assumptions in applying the biomarker in drug development or a clinical trial.

It is anticipated that the IRC-P will be employed in clinical trials of CHR-P/APS samples to test the efficacy of interventions for treating symptoms, cognitive deficits, and/or functional impairments in the prodromal phase of illness and for preventing progression to full psychosis and long-term functional disability. We anticipate that trial designers will use the IRC-P to enrich samples for cases at highest risk for progression to psychosis and other target outcomes. By selecting CHR-P cases with a predicted risk 20% or greater, the resulting sample would be expected to have a conversion rate of 28%, representing a doubling of the rate associated with CHR-P status alone,[18] which in turn would double the efficiency of the clinical trial and concentrate drug development efforts on cases most in need. However, other thresholds could be selected depending on the study context. For example, treatments with low toxicity profiles could use a lower threshold of predicted risk, enabling more rapid enrollment (lower screen-fail rate), while those with greater anticipated side effect burdens could use a higher threshold of predicted risk to reduce enrollment of false positive cases (higher specificity).

Our objective is to utilize FDA written feedback solicited from this LOI and discussion to inform this process.

Clinical Validation

Internal validation within the reference sample (NAPLS2) achieved a Concordance Index[25] (analogous to AUC) of 0.71. The IRC-P has subsequently been validated in two independent samples of CHR-P/APS cases.[26,27] One of these independent replication studies was conducted in a North American cohort of 176 CHR-P cases; the IRC-P achieved an AUC of 0.79 in that cohort.[26] The other independent replication study was conducted in a Chinese cohort of 199 CHR-P cases; this study did not include all of the individual predictors that contribute to the IRC-P prediction of psychosis risk, but nevertheless, the remaining items, when weighted according to the IRC-P formulaa, achieved an AUC of 0.63 in predicting conversion to psychosis.[27]

Benefits and Risks of Applying Clinical Decision Tool

The benefits of applying the IRC-P in a clinical trial are expected to be improved efficiency and targeting of treatments to individuals who are at greatest risk for psychosis and poor functional outcomes. The

risks of applying the IRC-P in a clinical trial are expected to be low. Assessment of the items used in the scoring algorithm is standard clinical practice for this population and carries no greater risk than is typical for clinical diagnostic interviewing in psychiatry.

Describe Knowledge Gaps, Limitations and Assumptions

The IRC-P is currently based on clinical, demographic, and neurocognitive input variables. As summarized below (see Biological Rationale), these clinical, demographic, and neurocognitive predictors are all correlated highly with biological indicators of putative underlying pathophysiologic mechanisms for psychosis, including common genetic risk variants for schizophrenia, reduced cortical thickness in prefrontal and other regions, and distrupted functional connectivity within and between these regions. It is not yet known whether genomic, structural imaging, and functional imaging measures could be used in place of the clinical, demographic, and neurocognitive predictors in individual-level prediction of psychosis, but this possibility is being examined by work in progress.

Supporting Information

For example (if applicable or available):

- Provide underlying biological process or supporting evidence of association of the biological process with the biomarker.
- Summary of existing preclinical or clinical data to support the biomarker in its COU (e.g., summaries of literature findings, previously conducted studies).
- Summary of any planned studies to support the biomarker and COU. How will these studies address any current knowledge gaps?
- Please describe alternative comparator, current standard(s), or approaches.

Biological rationale

The present form of the IRC-P is based on clinical and neurocognitive input variables. Nevertheless, because the IRC-P is a prediction model of conversion to psychosis, biological measures found to be associated with conversion to psychosis are likely to be reflected at least in part in variation in the clinical and neurocognitive variables included in the IRC-P and also thereby in the predicted risk scores themselves. Certain of the individual predictors, such as family history, almost certainly primarily reflect underlying biological-level mechanisms, such as inherited risk variants.[28] A similar analysis applies to the neurocognitive measures in the IRC-P – memory and processing speed[29-31] – which depend on the activity and coordination of brain networks including prefrontal cortex, medial temporal lobe, cerebellum, and thalamus, networks that are known to be impacted in CHR-P/APS individuals and particularly in those who ultimately convert to psychosis.[32-34] In addition, our work on structural and functional imaging predictors of psychosis has observed correlations between these imaging biomarkers and prodromal symptom severity, particularly in the domains of unusual thought content and disorganized thinking.[33,35]

Although the specific biological mechanisms underlying conversion to psychosis have not yet been definitively isolated, a number of biological correlates of risk for and progression to psychosis have been uncovered. The most prominent and consistent findings in this regard include progressive thinning of prefrontal cortex and other regions,[32,36] altered activation and connectivity of brain networks involved in memory and other cognitive functions,[33,34] abnormal electrophysiological signals

associated with novelty detection and sensory memory,[37] increased levels of stress hormones,[38-40] and altered expression of proteins and genes involved in synaptic plasticity, immune signaling, and oxidative metabolism.[41,42] An emerging theoretical view of this nexus of findings points to the role of altered signaling in pathways that regulate synaptic density and function, including key components of the glutamate cascade[43] and long-term potentiation as well as of the complement system and other immune parameters involved in synaptic elimination.[44,45] Indeed, initial findings show that markers of elevated inflammatory signaling precede and predict the accelerated cortical thinning associated with onset of psychosis.[32,46] Despite these promising leads, most of the findings on biological correlates of psychosis risk and progression published to date are based on group-average differences rather than individual-level outcome prediction. Those biological variables that have been examined in relation to individual-level prediction (including leukocyte microRNA expression[42] and cerebellar-thalamo-cortical hyperconnectivity[33]) show promising results in the initial discovery samples, with efforts to test their generalizability in independent samples now underway.

Summary of existing preclinical or clinical data

The purpose of this LOI submission is to solicit feedback from the FDA regarding the viability of the IRC-P as a prognostic biomarker for use in clinical trials, the current plan for conducting confirmatory studies, and next steps for testing the robustness of particular thresholds of predicted risk as cut-points for maximizing sensitivity and specificity of psychosis prediction.

In addition to the replication tests already published,[26,27] the replicability of the IRC-P as a prognostic biomarker is being assessed by a multi-consortia international effort called HARMONY (*Harmonization of At Risk Multisite Observational Networks for Youth*). NAPLS is collaborating with the PRONIA (https://www.pronia.eu) and PSYSCAN (http://psyscan.eu) consortia studies based in Europe, which have collected similar information on large cohorts of CHR-P individuals. In a collaborative analysis with the PRONIA consortium, using leave-site-out cross-validation methods, we have recently determined that the IRC-P generalizes to psychosis prediction in Europe and is robust across the 8 North American and 6 European sites included in NAPLS and PRONIA, respectively. Work in progress is testing the robustness of particular thresholds of predicted risk as cut-points for maximizing sensitivity and specificity of psychosis prediction using leave-site-out cross-validation approaches.

Summary of any planned studies to support the biomarker and COU

Work in progress is evaluating whether adding additional measures (including polygenic risk scores,[47,48] measured hormone levels,[38] meaures of brain activity from electroecephalography[37] or functional magnetic resonance imaging[33]) improves performance of the risk calculator. Thus far, we have observed only modest increases in predictive accuracy when adding such measures to the risk calculator,[48] reflecting the fact that these biological measures covary considerably with the clinical, demographic and neurocognitive variables already in the calculator. Nevertheless, such analyses provide a basis for understanding the meaning of particular predictor variables. For example, the fact that younger age at ascertainment of CHR-P status was associated with a higher conversion risk in the NAPLS2 sample has recently been determined to reflect an accelerated neuromaturational biomarker signal ascertained from machine learning of structural MRI data to predict chrononological age.[49,50]

We are planning to conduct a re-analysis of prior clinical trial data on CHR-P samples incorporating individualized predicted risk as a factor in the stastistical analysis, to determine whether treatment

outcomes are moderated by level of predicted risk. Our first attempt at this kind of analysis, in a clinical trial involving family focused therapy, demonstrated greater responsiveness to the intervention among CHR-P subjects with higher levels of predicted risks based on the IRC-P.[51] The resulting evidence may help to qualify the IRC-P for an expanded context of use such as treatment selection in a stepped care algorithm whereby the particular intervention (e.g., psychological, drug) is targeted/scaled to the individual's level of predicted risk. In particular, we anticipate that CHR-P/APS cases with lower predicted risk (e.g., < 20%) may improve with less intensive or invasive intervention approaches.

We are also considering integrating the multivariate proportional hazards model - that is focused on co-linear effects between variables - with sophisticated analytical approaches that can explore non-linear relations and multi-modal data fusion to deliver even more accurate predictions of risk and possibly treatment assignment based on retrospective data.

Alternative/comparator/current standard(s) approaches

To our knowledge, there is currently no alternative approach to the IRC-P for selecting samples for enrichment on psychosis risk beyond that associated with CHR-P/APS status.

Previous Qualification Interactions and Other Approvals (if applicable)

For example:

- Letter of Support (LOS) issued for this biomarker
- Discussion in a Critical Path Innovation Meeting (CPIM)
- Previous FDA Qualification given to this biomarker with DDT Tracking Record Number
- Qualification submissions to any other regulatory agencies with submission number
- Prior or current regulatory submissions to Center for Biologics Evaluation and Research (CBER), Center for Drug Evaluation and Research (CDER), and Center for Devices and Radiological Health (CDRH). Provide 510(k)/PMA Numbers

Letter of Support (LOS) issued for this biomarker on date:

N/A

Discussed in a Critical Path Innovation Meeting (CPIM) on date:

N/A

Previous FDA Qualification given to this biomarker with DDT Tracking Record Number

N/A

Qualification submissions to any other agencies with submission number

N/A

Prior or current Regulatory submissions to Center for Biologics Evaluation and Research (CBER), Center for Drug Evaluation and Research (CDER), and Center for Devices and Radiological Health (CDRH)

N/A

Attachments

This section may contain:

- Please provide a list of publications most relevant to this biomarker development proposal.

Attachment: Publications Relevant to the IRC-P Biomarker Development Proposal

- Optional: If this biomarker development effort is part of a longer-term goal, please summarize your long-term objectives.

- Optional: If you have other supporting information you would like to provide, please submit as attachment(s).

  We are attaching pdf reprints of the following papers:

  Cannon TD, Yu C, Addington J, et al. An Individualized Risk Calculator for Research in Prodromal Psychosis. Am J Psychiatry 2016;173:980-8.

  Carrion RE, Cornblatt BA, Burton CZ, et al. Personalized Prediction of Psychosis: External Validation of the NAPLS-2 Psychosis Risk Calculator With the EDIPPP Project. Am J Psychiatry 2016;173:989-96.

  Zhang T, Li H, Tang Y, et al. 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. Am J Psychiatry 2018;175:906-8.

Please note that any information provided as optional attachments will not be publicly posted.

Additional Information & Submission Information:

Please refer to the Resources for Biomarker Requestors for the mailing address and other important submission-related instructions. For more about Biomarker Qualification see our program's Home Page. If you have any questions about submission procedures, please contact via email; CDER-BiomarkerQualificationProgram@fda.hhs.gov.

Publications Relevant to the IRC-P Biomarker Development Proposal

1. Disease GBD, Injury I, Prevalence C. Global, regional, and national incidence, prevalence, and years lived with disability for 328 diseases and injuries for 195 countries, 1990-2016: a systematic analysis for the Global Burden of Disease Study 2016. Lancet 2017;390:1211-59.

2. Olfson M, Gerhard T, Huang C, Crystal S, Stroup TS. Premature Mortality Among Adults With Schizophrenia in the United States. JAMA psychiatry 2015;72:1172-81.

3. Leucht S, Cipriani A, Spineli L, et al. Comparative efficacy and tolerability of 15 antipsychotic drugs in schizophrenia: a multiple-treatments meta-analysis. Lancet 2013;382:951-62.

4. Steeds H, Carhart-Harris RL, Stone JM. Drug models of schizophrenia. Ther Adv Psychopharmacol 2015;5:43-58.

5. Lin CH, Lane HY. Early Identification and Intervention of Schizophrenia: Insight From Hypotheses of Glutamate Dysfunction and Oxidative Stress. Front Psychiatry 2019;10:93.

6. Insel TR. The arrival of preemptive psychiatry. Early intervention in psychiatry 2007;1:5-6.

7. Yung AR, McGorry PD. The prodromal phase of first-episode psychosis: past and current conceptualizations. Schizophrenia bulletin 1996;22:353-70.

8. Fusar-Poli P, Borgwardt S, Bechdolf A, et al. The psychosis high-risk state: a comprehensive state-of-the-art review. JAMA psychiatry 2013;70:107-20.

9. Miller TJ, McGlashan TH, Rosen JL, et al. 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 2003;29:703-15.

10. Cannon TD, Yu C, Addington J, et al. An Individualized Risk Calculator for Research in Prodromal Psychosis. Am J Psychiatry 2016;173:980-8.

11. Olvet DM, Carrion RE, Auther AM, Cornblatt BA. Self-awareness of functional impairment in individuals at clinical high-risk for psychosis. Early Interv Psychiatry 2015;9:100-7.

12. Baker AL, Kavanagh DJ, Kay-Lambkin FJ, et al. Randomized controlled trial of MICBT for co-existing alcohol misuse and depression: outcomes to 36-months. Journal of Sustance Abuse Treatment 2014;46:281-90.

13. Fusar-Poli P, Bonoldi I, Yung AR, et al. Predicting psychosis: meta-analysis of transition outcomes in individuals at high clinical risk. Archives of general psychiatry 2012;69:220-9.

14. Cannon TD, Cadenhead K, Cornblatt B, et al. Prediction of psychosis in youth at high clinical risk: a multisite longitudinal study in North America. ArchGenPsychiatry 2008;65:28-37.

15. Addington J, Cornblatt B, Cadenhead K, et al. At clinical high risk for psychosis: Outcome for non-converters. American Journal of Psychiatry 2011.

16. Schlosser DA, Jacobson S, Chen Q, et al. Recovery from an at-risk state: clinical and functional outcomes of putatively prodromal youth who do not develop psychosis. Schizophrenia bulletin 2012;38:1225-33.

17. Allswede DM, Addington J, Bearden CE, et al. Characterizing Covariant Trajectories of Individuals at Clinical High Risk for Psychosis Across Symptomatic and Functional Domains. Am J Psychiatry 2019:appiajp201918111290.

18. Conrad AM, Lewin TJ, Sly KA, et al. Utility of risk-status for predicting psychosis and related outcomes: evaluation of a 10-year cohort of presenters to a specialised early psychosis community mental health service. Psychiatry Res 2017;247:336-44.

19. Addington J, Cadenhead KS, Cornblatt BA, et al. North American Prodrome Longitudinal Study (NAPLS 2): overview and recruitment. Schizophrenia research 2012;142:77-82.

20. Cornblatt BA, Auther AM, Niendam T, et al. Preliminary findings for two new measures of social and role functioning in the prodromal phase of schizophrenia. Schizophrenia bulletin 2007;33:688-702.

21. Dohrenwend BS, Krasnoff L, Askenasy AR, Dohrenwend BP. Exemplification of a method for scaling life events: the Peri Life Events Scale. Journal of health and social behavior 1978;19:205-29.

22. Janssen I, Krabbendam L, Bak M, et al. Childhood abuse as a risk factor for psychotic experiences. Acta psychiatrica Scandinavica 2004;109:38-45.

23. Keefe RS, 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. Schizophrenia research 2004;68:283-97.

24. Brandt J, Benedict RHB. Hopkins Verbal Learning Test-Revised (HVLT-R). Odessa, FL: Psychological Assessment Resources, Inc.; 1998.

25. Uno H, Cai T, Pencina MJ, D'Agostino RB, Wei LJ. On the C-statistics for evaluating overall adequacy of risk prediction procedures with censored survival data. Statistics in medicine 2011;30:1105-17.

26. Carrion RE, Cornblatt BA, Burton CZ, et al. Personalized Prediction of Psychosis: External Validation of the NAPLS-2 Psychosis Risk Calculator With the EDIPPP Project. Am J Psychiatry 2016;173:989-96.

27. Zhang T, Li H, Tang Y, et al. 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. Am J Psychiatry 2018;175:906-8.

28. Sullivan PF. The genetics of schizophrenia. PLoS medicine 2005;2:e212.

29. Giuliano AJ, Li H, Mesholam-Gately RI, Sorenson SM, Woodberry KA, Seidman LJ. Neurocognition in the psychosis risk syndrome: a quantitative and qualitative review. Current pharmaceutical design 2012;18:399-415.

30. Riecher-Rossler A, Pflueger MO, Aston J, et al. Efficacy of using cognitive status in predicting psychosis: a 7-year follow-up. Biological psychiatry 2009;66:1023-30.

31. Seidman LJ, Giuliano AJ, Meyer EC, et al. Neuropsychology of the prodrome to psychosis in the NAPLS consortium: relationship to family history and conversion to psychosis. Archives of general psychiatry 2010;67:578-88.

32. Cannon TD, Chung Y, He G, et al. Progressive reduction in cortical thickness as psychosis develops: a multisite longitudinal neuroimaging study of youth at elevated clinical risk. Biological psychiatry 2015;77:147-57.

33. Cao H, Chen OY, Chung Y, et al. Cerebello-thalamo-cortical hyperconnectivity as a state-independent functional neural signature for psychosis prediction and characterization. Nat Commun 2018;9:3836.

34. Cao H, McEwen SC, Chung Y, et al. Altered Brain Activation During Memory Retrieval Precedes and Predicts Conversion to Psychosis in Individuals at Clinical High Risk. Schizophrenia bulletin 2019;45:924-33.

35. Chung Y, Jacobson A, He G, et al. Prodromal Symptom Severity Predicts Accelerated Gray Matter Reduction and Third Ventricle Expansion Among Clinically High Risk Youth Developing Psychotic Disorders. Mol Neuropsychiatry 2015;1:13-22.

36. Chung Y, Allswede D, Addington J, et al. Cortical abnormalities in youth at clinical high-risk for psychosis: Findings from the NAPLS2 cohort. Neuroimage Clin 2019;23:101862.

37. Bodatsch M, Ruhrmann S, Wagner M, et al. Prediction of psychosis by mismatch negativity. Biological psychiatry 2011;69:959-66.

38. Walker EF, Trotman HD, Pearce BD, et al. Cortisol levels and risk for psychosis: initial findings from the

North American prodrome longitudinal study. Biological psychiatry 2013;74:410-7.

39. Holtzman CW, Shapiro DI, Trotman HD, Walker EF. Stress and the prodromal phase of psychosis. Current pharmaceutical design 2012;18:527-33.

40. Walker EF, Walder DJ, Reynolds F. Developmental changes in cortisol secretion in normal and at-risk youth. Development and psychopathology 2001;13:721-32.

41. Jeffries CD, Perkins DO, Fournier M, et al. Networks of blood proteins in the neuroimmunology of schizophrenia. Transl Psychiatry 2018;8:112.

42. Jeffries CD, Perkins DO, Chandler SD, et al. Insights into psychosis risk from leukocyte microRNA expression. Transl Psychiatry 2016;6:e981.

43. Bossong MG, Antoniades M, Azis M, et al. Association of Hippocampal Glutamate Levels With Adverse Outcomes in Individuals at Clinical High Risk for Psychosis. JAMA psychiatry 2018.

44. Cannon TD. How Schizophrenia Develops: Cognitive and Brain Mechanisms Underlying Onset of Psychosis. Trends Cogn Sci 2015;19:744-56.

45. McGlashan TH, Hoffman RE. Schizophrenia as a disorder of developmentally reduced synaptic connectivity. Archives of general psychiatry 2000;57:637-48.

46. Zheutlin AB, Jeffries CD, Perkins DO, et al. The Role of microRNA Expression in Cortical Development During Conversion to Psychosis. Neuropsychopharmacology 2017;42:2188-95.

47. Calafato MS, Thygesen JH, Ranlund S, et al. Use of schizophrenia and bipolar disorder polygenic risk scores to identify psychotic disorders. Br J Psychiatry 2018;213:535-41.

48. Perkins DO, Olde Loohuis L, Barbee J, et al. Polygenic Risk Score Contribution to Psychosis Prediction in a Target Population of Persons at Clinical High Risk. Am J Psychiatry 2020;177:155-63.

49. Chung Y, Addington J, Bearden CE, et al. Use of Machine Learning to Determine Deviance in Neuroanatomical Maturity Associated With Future Psychosis in Youths at Clinically High Risk. JAMA psychiatry 2018;75:960-8.

50. Chung Y, Addington J, Bearden CE, et al. Adding a neuroanatomical biomarker to an individualized risk calculator for psychosis: A proof-of-concept study. Schizophrenia research 2019.

51. Worthington MA, Miklowitz DJ, O'Brien M, et al. Selection for psychosocial treatment for youth at clinical high risk for psychosis based on the North American Prodrome Longitudinal Study individualized risk calculator. Early intervention in psychiatry 2020.

An Individualized Risk Calculator for Research in Prodromal Psychosis

Tyrone D. Cannon, Ph.D., Changhong Yu, M.S., Jean Addington, Ph.D., Carrie E. Bearden, Ph.D., Kristin S. Cadenhead, M.D., Barbara A. Cornblatt, Ph.D., Robert Heinssen, Ph.D., Clark D. Jeffries, Ph.D., Daniel H. Mathalon, Ph.D., M.D., Thomas H. McGlashan, M.D., Diana O. Perkins, M.D., M.P.H., Larry J. Seidman, Ph.D., Ming T. Tsuang, M.D., Ph.D., Elaine F. Walker, Ph.D., Scott W. Woods, M.D., Michael W. Kattan, Ph.D.

Objective: Approximately 20%–35% of individuals 12–35 years old who meet criteria for a prodromal risk syndrome convert to psychosis within 2 years. However, this estimate ignores the fact that clinical high-risk cases vary considerably in risk. The authors sought to create a risk calculator, based on profiles of risk indicators, that can ascertain the probability of conversion to psychosis in individual patients.

Method: The study subjects were 596 clinical high-risk participants from the second phase of the North American Prodrome Longitudinal Study who were followed up to the time of conversion to psychosis or last contact (up to 2 years). The predictors examined were limited to those that are supported by previous studies and are readily obtainable in general clinical settings. Time-to-event regression was used to build a multivariate model predicting conversion, with internal validation using 1,000 bootstrap resamples.

Results: The 2-year probability of conversion to psychosis was 16%. Higher levels of unusual thought content and suspiciousness, greater decline in social functioning, lower verbal learning and memory performance, slower speed of processing, and younger age at baseline each contributed to individual risk for psychosis. Stressful life events, trauma, and family history of schizophrenia were not significant predictors. The multivariate model achieved a concordance index of 0.71 and, as reported in an article by Carrión et al., published concurrently with this one, was validated in an independent external data set. The results are instantiated in a web-based risk prediction tool envisioned to be most useful in research protocols involving the psychosis prodrome.

Conclusions: A risk calculator comparable in accuracy to those for cardiovascular disease and cancer is available to predict individualized conversion risks in newly ascertained clinical high-risk cases. Given that the risk calculator can be validly applied only for patients who screen positive on the Structured Clinical Interview for Psychosis Risk Syndromes, which requires training to administer, its most immediate uses will be in research on psychosis risk factors and in research-driven clinical (prevention) trials.

AJP in Advance (doi: 10.1176/appi.ajp.2016.15070890)

Given limitations of available treatments for schizophrenia, with most patients showing substantial deficits in social and occupational functioning throughout life, there is considerable interest in developing preventive approaches to psychotic disorders (1). Ascertainment of individuals at greatest risk is crucial to these efforts. For the majority of patients, onset of fully psychotic symptoms is preceded by the emergence of subtler changes in belief, thought, and perception that appear to represent attenuated forms of delusions, formal thought disorder, and hallucinations, respectively. Among individuals 12–35 years old with a recent onset of such symptoms (termed clinical high-risk cases), approximately 20%−35% develop fully psychotic symptoms over a 2-year period, an incidence rate that is more than 100 times larger than in the same age band in the general population (2). Furthermore, it appears that the clinical high-risk criteria are sensitive to an imminent risk for onset, as most of the conversions occur during the first year after ascertainment, with a decelerating conversion rate thereafter (3).

Although clinical high-risk criteria have been validated in epidemiological studies as sensitive to conversion risk, their utility in individual decision making is currently limited, given that 65%–80% of cases ascertained by these methods do not convert to psychosis within a 2-year time frame. About a dozen studies have examined combinations of clinical and demographic variables to determine whether prediction of psychosis can be enhanced beyond the 20%−35% risk associated with clinical high-risk status (4). Multivariate algorithms requiring particular combinations of symptoms and demographic factors achieve relatively high positive predictive values and specificity (e.g., in the 50%−70% range) but low sensitivity (e.g., in the 10%−30% range) (3). There is

consistency among studies in showing (unsurprisingly) that greater severity of the psychosis-risk symptoms at baseline is the best predictor of conversion; nevertheless, the most predictive multivariate profiles vary across studies (as summarized in reference 4). Although it should be noted that few studies have attempted direct replication of one another's risk algorithms, this pattern suggests heterogeneity among profiles of clinical and demographic risk indicators among patients who convert to psychosis.

To maximize clinical utility, we require an approach that can be applied to scale the risk in an individual patient at the initial clinical contact. Such individualized risk calculation is possible when a large data set on a reference population is available from which risks can be calculated based on one or more predictor variables. Well-performing risk calculators have been developed in numerous somatic disease contexts, including cardiovascular disease and cancer (5–9), where they provide a rationale for clinicians to pursue more or less invasive intervention strategies, based on the level of risk implied by an individual's profile across a set of risk factors. They also inform patients and their family members, thus helping them make complex treatment decisions.

Here we present such an individualized risk calculator for psychosis, using data from the second phase of the North American Prodrome Longitudinal Study (NAPLS-2). Predictors were chosen a priori based on a review of the literature on psychosis risk prediction in clinical high-risk samples, blindly with respect to the empirical relationships between any of the nominated variables and psychosis outcome within the NAPLS-2 data set. We limited our scope to clinical, cognitive, and demographic measures that are readily obtainable in standard clinical settings. Using time-to-event proportional hazards regression, a risk calculator was generated that calculates risk according to an individual's values on the included variables. We evaluated the performance of the risk calculator using the concordance index (the Harrell C-index, a measure of overall accuracy, analogous to area under the receiver operating characteristic curve) and assessed the relative importance of each of the included predictor variables.

METHOD

Study Subjects and Clinical Characterization

The study protocol and consent form were reviewed and approved by the institutional review boards of the eight data collection sites (UCLA, Emory, Beth Israel Deaconess Medical Center, Zucker Hillside Hospital, University of North Carolina, UCSD, Calgary, and Yale). We have previously reported on the methods for evaluation of subjects and data collection (10). Participants were evaluated using the Structured Interview for Prodromal Syndromes (SIPS) (11) and the Structured Clinical Interview for DSM-IV Axis I Disorders (12) by trained interviewers who met high reliability standards (intraclass correlation coefficients, 0.92–0.96) (10). Patients who had ever met DSM-IV criteria for a psychotic disorder, had a history of substance dependence, had a neurological disorder, or had an estimated IQ <70 were excluded. Participants met SIPS criteria (13) for the presence of one or more clinical high-risk syndromes: attenuated psychotic symptoms syndrome, brief intermittent psychotic symptom syndrome, and familial risk and deterioration syndrome.

Follow-up clinical evaluations were scheduled every 6 months after study entry through 2 years. Conversion to psychosis was determined by SIPS criteria, which are designed to operationalize the threshold of delusional ideation or hallucination severity required for a DSM-IV diagnosis of a psychotic disorder. Participants were followed up to the time of conversion to psychosis or the last contact (up to 2 years), the dates for which were recorded, permitting calculation of length of time in the study until conversion or censoring (loss to follow-up).

A total of 743 clinical high-risk patients who met SIPS criteria were enrolled in NAPLS-2 from 2008 to 2013. In the present analysis, we excluded subjects who dropped out of the study before any clinical follow-up was conducted (N=147). The final study cohort consisted of 596 clinical high-risk participants who had at least one follow-up evaluation.

Selection of Predictor Variables

To meet the objective of developing a practical tool for risk prediction, our focus was on demographic, clinical, neurocognitive, and functioning measures that are easily administered in general clinical settings. The number of predictors was limited a priori to eight to ensure that there were a minimum of 10 converters per predictor in the model, which helps mitigate model instability due to overfitting. For the same reason, we avoided including terms for the interactions among the predictors. The NAPLS-2 data set itself was not used to select predictors; doing so would have invalidated the logic of using the underlying data to inform prediction for new cases (i.e., the predictive logic would then be circular). Rather, we evaluated the published literature on psychosis prediction in clinical high-risk samples. Our selection of indicators was based on empirical links to psychosis prediction in two or more prior studies of clinical high-risk cases; there was no attempt to select predictors based on a theoretical model of causes of psychosis or clinical knowledge or intuition. Based on this process, eight variables were chosen for inclusion.

Age at ascertainment was included to help account for variation in age at onset of psychosis (14) and in processes that undergo developmental modification during the age range of our sample (15–17). Greater severity of SIPS items P1 and P2 (unusual thought content and suspiciousness) are strongly predictive of psychosis in clinical high-risk samples (3, 4). Given that the meanings of gradations below the prodromal threshold are likely to be different from those at or above this threshold, these items were modified such that all levels in the nonprodromal range (0–2 on the original scale) were redefined as 0, levels in the prodromal range (3–5 on the original scale) were redefined as 1–3, and psychotic intensity (6 on the original scale) was redefined as 4, and these scores were

TABLE 1. Characteristics of Clinical High-Risk Subjects Who Were and Were Not Followed Up in NAPLS-2[a]

Variable	Followed (N=596)		Not Followed (N=147)		Statistical Analysis		Missing Values[b]	
	Mean	SD	Mean	SD	t	p	N	%
Age (years)	18.5	4.3	18.8	4.2	−0.88	0.38	0	0.0
Modified[c] SIPS items P1 + P2, summed score	2.6	1.6	2.6	1.6	0.07	0.94	0	0.0
BACS symbol coding, raw score (number completed)	56.8	13.1	57.9	11.6	−0.81	0.42	22	3.7
Hopkins Verbal Learning Test–Revised, trials 1–3 summed	25.6	5.2	25.1	5.4	0.85	0.39	21	3.5
Stressful life events	10.5	5.5	10.0	5.6	0.90	0.37	69	11.6
	N	%	N	%	χ^2	p	N	%
Family history of psychosis	96	16.1	18	12.2	1.38	0.24	2	0.3
Decline in functioning >0 on the Global Functioning: Social scale	270	45.4	75	53.5	3.05	0.08	1	0.2
Traumas >1	289	56.2	48	48.5	2.00	0.16	82	13.7
Male	344	57.7	77	52.4	1.36	0.24	0	0.0

[a] BACS=Brief Assessment of Cognition in Schizophrenia; NAPLS-2=second phase of the North American Prodrome Longitudinal Study; SIPS=Structured Interview for Prodromal Syndromes.
[b] Among those followed. Missing values were multiply imputed with the multivariate imputation by chained equations method prior to use in prediction analyses.
[c] Modified such that all levels in the nonprodromal range (0–2 on the original scale) are recoded as 0, levels in the prodromal range (3–5 on the original scale) are recoded as 1–3, and psychotic intensity (6 on the original scale) is recoded as 4.

summed. Several studies have found that slower processing speed and lower verbal learning and memory functioning are predictive of psychosis (18, 19) and in meta-analyses have among the largest effect sizes among converters to psychosis (20). These constructs were represented by scores on the Brief Assessment of Cognition in Schizophrenia symbol coding test (21) and the Hopkins Verbal Learning Test–Revised (sum of trials 1–3) (22), respectively. Many clinical high-risk cases who convert to psychosis show a pronounced decline in social functioning in the year prior to ascertainment (23), measured here using the Global Functioning: Social scale (24). Stressful life events, along with childhood traumas, have been shown to be predictive of psychosis in studies of clinical high-risk samples (25). To represent the former, we aggregated 31 life events designated as negative and potentially relevant to individuals 12–35 years old from the Research Interview Life Events Scale (26), and for the latter, we used the Childhood Trauma and Abuse Scale (27). Family history of psychotic disorder in a first-degree relative is by itself not a robust predictor of psychosis in studies of clinical high-risk samples (3, 4), but it was nevertheless included because it elevates risk by almost 10-fold compared with the general population (28).

Statistical Methods

Risk calculators have been developed to assist health care professionals for a variety of illnesses (5–9) (see http://www.lerner.ccf.org/qhs/risk_calculator/ for examples). Calculators can derive a risk prediction for a particular person from a given set of indicators by querying a multivariate model based on a large sample of similar cases. Through imputation, calculators can accommodate incomplete information on the panel of risk indicators. However, the more complete the information available on a given case, the more powerful they become, with a tighter range of certainty.

We built a multivariate proportional hazards model to predict the likelihood of conversion to psychosis based on each participant's demographic, cognitive, and clinical characteristics, as defined above. We tested restricted cubic splines in relation to continuous variables; as none were significant, no adjustments were made. As shown in Table 1, there were little or no missing data for age, symptom severity, family history, and social functioning. Cognitive test data were missing in less than 4% of cases, and data regarding stressful life events or traumas were missing in 12%−14% of cases. In order to reduce selection bias and maximize the sample size, missing predictors were multiply imputed with the multivariate imputation by chained equations method before the multivariate regression.

The statistical model was internally validated using 1,000 bootstrap resamples, where the discrimination and calibration performance were evaluated. Harrell's C-index was used to quantify the discrimination ability for separating psychosis converters and nonconverters, which is analogous to the area under the receiver operating characteristic curve, with a range of 0.5 (no discrimination) to 1 (perfect discrimination), but tailored for censored data (29). A plot of the model-predicted probabilities versus the observed outcomes was used to assess calibration performance.

All statistical analyses were conducted using R, version 3.0.1 (R Core Team, 2013) including the rms and Hmisc packages.

RESULTS

Baseline patient characteristics are summarized in Table 1. There were no significant differences between those who

were and were not followed up clinically on any of the predictor variables. Of the 596 participants for whom follow-up data were available, 84 converted to psychosis within the 2-year study period. The mean age of the sample was 18.5 years. Among converters, the mean time from baseline to conversion was 7.3 months, and among nonconverters, the mean follow-up time from baseline to the last contact was 19.1 months. A total of 280 cases were followed up at 24 months without converting, and the remaining "nonconverters" were lost to follow-up at various points between 6 months and 24 months.

The 2-year probability of conversion to psychosis was 0.16 (95% CI=0.13, 0.19). Figure 1 provides frequency distributions of predicted risks for converters and nonconverters. Converters occur at a proportionally higher rate than nonconverters in each successive risk class, beginning at a predicted risk of 0.20. The output of the multivariate proportional hazards model is presented in Table 2. Prodromal symptom severity (SIPS items P1 and P2, modified and summed), decline in social functioning, and verbal learning and memory (Hopkins Verbal Learning Test–Revised scores) were significant predictors, with nonsignificant effects for age at baseline and speed of processing (symbol coding score) (p values, <0.10), although all of these variables were significant in univariate analyses (p values, <0.01). Stressful life events, traumas, and family history of schizophrenia were not significant predictors in univariate or multivariate analyses.

Table 2 provides additional diagnostics of the performance of individual predictor variables. Predictors associated with the largest decreases in the C-index when removed from the model were symptom severity, decline in global social functioning, Hopkins Verbal Learning Test–Revised score, and symbol coding score. Predictors associated with the largest increases in the C-index (i.e., above that of the base model, which included only symptom severity) were symbol coding score, Hopkins Verbal Learning Test–Revised score, decline in social functioning, and age. Family history of psychosis, stressful life events, and traumas did not alter the C-index by more than 0.5% when added to or deleted from the model.

Based on the bootstrap internal validation, the multivariate model achieved a C-index of 0.71. As shown in Figure 2, the calibration plot revealed a high degree of consistency between observed probabilities and model-predicted probabilities of conversion to psychosis within the range of 0.0–0.4, within which 95% of the cases fell (mean=0.18, SD=0.11, median=0.16). Table 3 provides statistics for prediction of actual conversion to psychosis across several thresholds of model-predicted risk. There is a trade-off between the positive predictive value (proportion of cases at the threshold of predicted risk who actually converted) and sensitivity (proportion of actual converters who had predicted risks at that threshold). The positive predictive value is maximal (48.4%) at a threshold of 0.4 of model-predicted risk, but only 17.9% of converters had model-predicted risks at this threshold. Conversely, at a model-predicted risk of 0.2 or higher, the positive predictive value is 28.1%, but with a sensitivity of 66.7%.

An online version of the risk calculator was built to facilitate numeric calculation of the predicted probability of conversion to psychosis (http://riskcalc.org:3838/napls/).

DISCUSSION

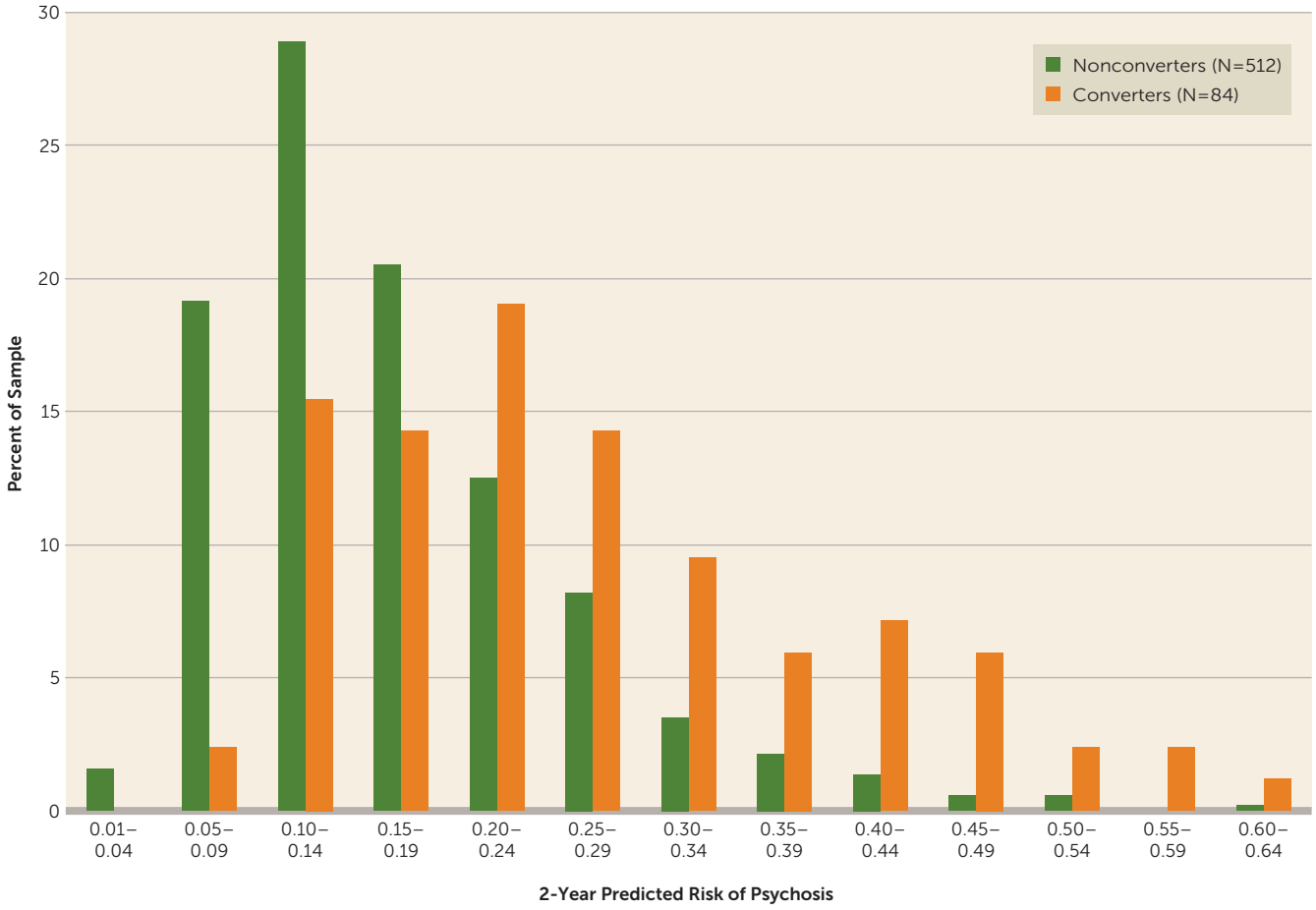
The goal of this study was to develop a practical tool for the individualized prediction of psychosis in clinical high-risk patients. A well-performing risk calculator was generated from the NAPLS-2 cohort data using a small number of demographic (age, family history of psychosis), clinical (unusual thought content and suspiciousness), neurocognitive (verbal learning and memory, speed of processing), and psychosocial (traumas, stressful life events, decline in social functioning) predictor variables. The overall model achieved a C-index of 0.71, which is in the range of values for established calculators currently in use for cardiovascular disease and cancer recurrence risk, which range from 0.58 to 0.81 (5–9).

The risk calculator generates a number representing the probability of transition to psychosis given a particular profile of input variables. Technically, this is an observed likelihood of conversion within the NAPLS-2 cohort itself, but this framework uses the logic of predictive inference to extend that observed likelihood based on past cases to the predicted probability for a newly ascertained case with the same profile. This logic rests on the assumption that the new case is ascertained from the same population and in a manner similar to those in NAPLS-2.

In particular, given that this risk calculator assumes a SIPS-based diagnosis of a prodromal risk syndrome as a starting point, the risk prediction tool would not be usable if such a risk syndrome has not been diagnosed. The risk calculator also assumes particular pathways to ascertainment, in that clinical high-risk cases in NAPLS are distressed and treatment seeking. This tool would thus be most useful to clinicians with training in psychosis risk detection using the SIPS (which, in addition to risk status, ascertains severity of unusual thought content and suspiciousness and family history of psychosis), who could then use the calculator for patients who have screened positive for a prodromal risk syndrome. Critically, risk determinations should be communicated to clients by trained clinicians who can help clients understand the meaning of the risk estimates (i.e., calibrated to the sample from which they were generated) and provide commensurate treatment recommendations. Note that within the context of NAPLS-2, with a mean predicted risk of 0.18 (SD=0.11), predicted risks of 0.3 or higher are relatively rare (12.4% prevalence among those meeting clinical high-risk criteria) and potent (39.2% positive predictive value). Proper training in the administration and scoring of the other measures included in the risk calculator (symbol coding, Hopkins Verbal Learning Test–Revised, Global Functioning: Social scale, Research Interview Life Events Scale, Childhood Trauma and Abuse Scale) is also required.

A key advantage of the risk calculator is that it inherently accommodates heterogeneity in profiles of risk factors among

FIGURE 1. Frequency Distributions of Predicted Risks Among Nonconverters and Converters[a]



[a] Beginning at a predicted risk of 0.20 or higher, there are proportionally more converters than nonconverters in each successive risk class.

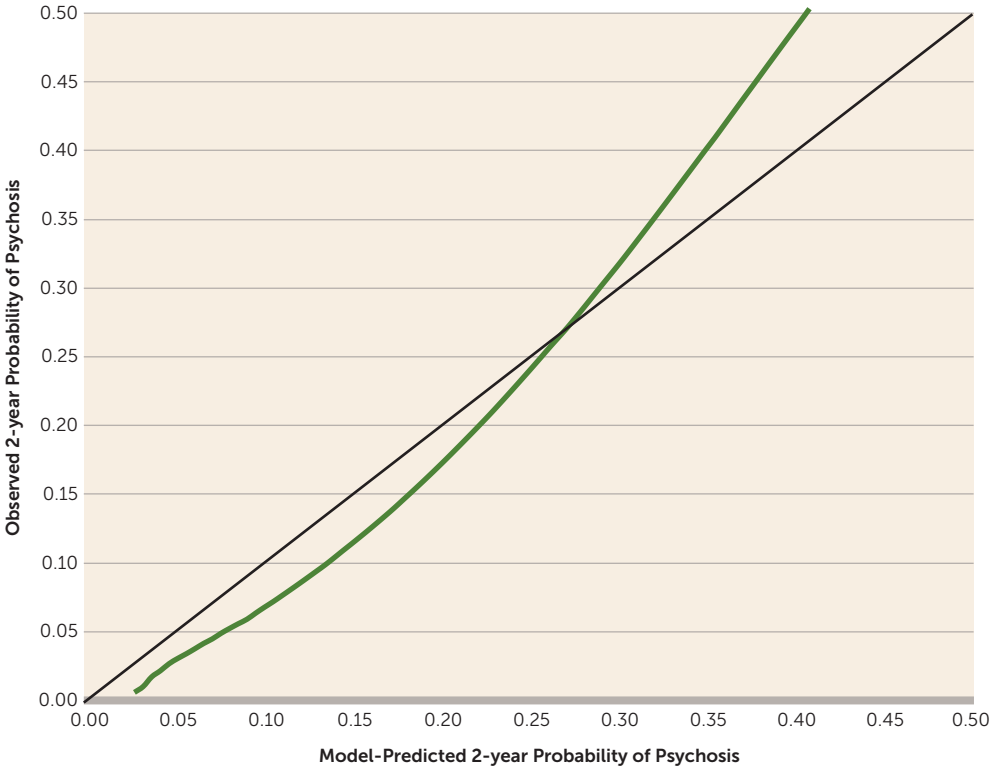
TABLE 2. Statistics for Individual Predictor Variables in the Multivariate Cox Proportional Hazards Regression Analysis of Conversion to Psychosis[a]

	Multivariate Model			C-Index[b]	
Predictor	Hazard Ratio	95% CI	p	Decrement If Removed	Increase If Added
Modified SIPS items P1 + P2	2.1	1.6–2.7	<0.001	0.092	N/A[c]
Decline in social functioning (Global Functioning: Social scale)	1.3	1.1–1.5	0.01	0.014	0.015
Hopkins Verbal Learning Test–Revised, trials 1–3 summed	0.8	0.6–0.9	0.05	0.007	0.029
BACS symbol coding, raw score (number completed)	0.8	0.5–1.1	0.10	0.006	0.033
Age	0.7	0.5–1.1	0.09	0.004	0.012
Stressful life events	1.2	0.9–1.6	0.21	0.001	−0.004
Family history of psychosis	1.2	0.7–2.1	0.55	0.000	0.001
Traumas	1.0	0.8–1.3	0.99	−0.004	0.002

[a] BACS=Brief Assessment of Cognition in Schizophrenia; SIPS=Structured Interview for Prodromal Syndromes.
[b] Harrell's C-index (equivalent to the area under the receiver operating characteristic curve) was used to quantify the discrimination ability for separating converters and nonconverters. The C-index for the overall model was 0.714.
[c] The base model included only the modified SIPS P1 + P2 score; the C-index for the base model was 0.666.

clinical high-risk cases. Examining configurations that vary across the significant predictors—greater prodromal symptom severity, lower verbal learning and memory, slower speed of processing, greater decline in social functioning, and younger age—reveals that dozens of separate permutations yield predicted conversion risks of 0.3 or higher. Stressful life events, traumas, and family history of schizophrenia have a negligible impact on their own or in combinations with other variables in the prediction of psychosis, but they were present more frequently among clinical high-risk individuals compared with

FIGURE 2. Calibration Plot of the Accuracy of Model-Predicted Probability in Relation to Observed Probability of Psychosis[a]



[a] The observed probability was estimated using proportional hazards regression evaluating the predicted 2-year probabilities in relation to the observed conversion events, taking into account time to conversion or censoring. The overfitting bias for the estimated observed probability was corrected using 1,000 bootstrap resamples. The plot shows excellent calibration across predicted probabilities of 0.0–0.4, corresponding to 95% of the NAPLS-2 sample. Predicted probabilities above 0.4 are too sparsely represented to permit adequate calibration testing.

TABLE 3. Prediction Statistics for Conversion to Psychosis Across Various Levels of Model-Predicted Risk

Individual's Predicted Risk	Base Rate of Predicted Risk Class	Positive Predictive Value	Negative Predictive Value	Sensitivity	Specificity
0.05–1.00	97.5	14.3	93.3	98.8	2.7
0.10–1.00	78.9	16.8	96.0	94.1	23.6
0.15–1.00	52.9	21.6	94.3	81.0	51.8
0.20–1.00	33.4	28.1	93.0	66.7	72.1
0.25–1.00	20.6	32.5	90.7	47.6	83.8
0.30–1.00	12.4	39.2	89.5	34.5	91.2
0.35–1.00	8.1	41.7	88.3	23.8	94.5
0.40–1.00	5.2	48.4	87.8	17.9	96.9
0.45–1.00	3.5	47.6	87.1	11.9	97.9
0.50–1.00	2.0	41.7	86.5	6.0	98.6
0.55–1.00	1.2	28.6	86.1	2.4	99.0
0.60–1.00	1.0	16.7	85.9	1.2	99.0

healthy comparison subjects. Perhaps these variables are more significant for determining presence of a clinical high-risk syndrome and thus are not as sensitive to outcomes within a group of subjects with a clinical high-risk syndrome.

A crucial test of robustness of a statistical model is validation on an independent external data set. In a companion article (30), Carrión et al. report on a replication test of the NAPLS-2 risk calculator in an independent sample from the Early Detection, Intervention, and Prevention of Psychosis Program (EDIPPP) that included 176 clinical high-risk cases diagnosed using the SIPS and followed clinically to monitor conversion. Only the stress and trauma variables—found to be negligible in predicting conversion here—were not collected and were therefore omitted from the replication testing. The remaining six NAPLS-2 risk factors yielded a significant time-to-event proportional hazards regression model predicting conversion in the EDIPPP sample ($p<0.003$), with a C-index of 0.79, which is even better than in the NAPLS-2 sample (0.71). The predictive model was well calibrated, and the NAPLS-2 calculator provided a reasonable estimation of psychosis risk when considering the risk prediction generated by the validation model and the actual observed outcomes. In addition, when applied to the external EDIPPP sample, the NAPLS-2 calculator showed sensitivity and specificity values comparable to those observed in the NAPLS-2 sample across different levels of model-predicted risk (30).

Our findings also show some degree of convergence with previous studies that reported multivariate models but that used their own samples for variable selection (i.e., model optimization) and did not present a web-based tool for extending individualized risk estimation to future patients (3, 23, 31). For example, a recent study (23) using a smaller (N=92) and nonoverlapping sample of clinical high-risk cases from one of the NAPLS sites developed a classifier that included three of the predictors included in the NAPLS-2 risk calculator (suspiciousness, verbal memory deficits, and decline in social functioning). Note that the sample in that study was less than one-fifth the size of the present sample, used a more restricted age range (ages 12–20), and was not

ascertained using SIPS criteria, factors that make risk classifications based on it much less generalizable than that of the NAPLS-2 risk calculator.

The most immediate uses of the risk calculator are likely to be in the selection of subjects for participation in clinical (prevention) trials, given the desire to avoid exposing patients with lower conversion risks to the potential adverse consequences of any interventions and given the potential to evaluate whether interventions differ in effectiveness based on initial risk levels or profiles across predictors. In terms of clinical practice outside the context of a prevention trial, at this point the most likely use of the risk calculator is for the clinician to be able to communicate to the patient and family a scaling of risk that could help in recruiting their cooperation with a monitoring and/or intervention plan.

A current limitation of the psychosis risk calculator is that risk estimates are not bounded by a confidence interval, making it unclear how well the single value output as a conversion risk represents the individual's actual likelihood of conversion. This issue is particularly problematic for computed risks of 0.5 or higher, for which there is sparse representation in the NAPLS-2 data set and for which calibration of the risk calculator could consequently not be adequately tested. Nevertheless, the use of confidence intervals is not likely to be of value in discussing risks with individual patients and their families, as the general concept of a confidence interval relates to likelihoods under future sampling rather than to an individual case, and the calculated risk is the best estimate for that individual (32).

Because the replication study (EDIPPP) included several community behavioral health centers and intergovernmental managed mental health organizations, the risk calculator appears to be generalizable beyond academic medical centers, at least within the U.S. health care system. The degree to which the risk calculator generalizes to other health care system models remains an open question.

In addition to testing the calculator's performance in independent data sets, future work could determine whether other variables, including biological tests, can improve prediction over and above the set of clinical, demographic, and cognitive measures evaluated here. Some promising leads on the use of biological assays to predict psychosis among clinical high-risk patients have emerged using empirically based discovery approaches, including machine learning algorithms for gray matter variations in structural brain images (33) and "greedy" regression algorithms for proteomic/metabolic plasma parameters (34). Studies employing discovery-oriented model-optimization methods, with parallel, independent samples, are needed to better inform future versions of this and other risk calculators. However, it is still critical to note that the data used in any risk calculator cannot be the same data that are used in the model optimization phase; as noted above, doing so would invalidate the risk predictions for new cases.

Given that in approximately one-third of clinical high-risk cases, the symptoms that determined their initial risk status remit within 6–12 months of ascertainment (35, 36), it should be possible to develop a complementary tool to predict a new case's likelihood of remission from a clinical high-risk syndrome. Such an estimate would not necessarily be merely the inverse of the conversion risk, as different predictors may be relevant.

It is also possible that risk calculators could eventually be used to select clients for different treatment regimens or to reclassify risk after completion of a particular intervention. At this stage, the knowledge base for doing so is limited, as only a small number of controlled prevention trials in clinical high-risk cases have appeared. Collectively, the results support the view that any targeted intervention, whether biological or psychological in approach, is associated with better outcomes than less targeted control conditions (37). Results of two small trials with antipsychotic drugs do not support a prophylactic effect on conversion risk beyond the period of active treatment (38, 39). In general, the use of such medications in individuals who are below the threshold of full psychosis is not recommended. Intriguing results have been obtained in an initial trial of omega-3 fatty acid supplementation (40); this finding awaits confirmation by independent studies. Psychosocial interventions such as cognitive-behavior therapy and family-focused psychoeducation may be beneficial in deflecting the course of illness severity and chronicity (41, 42); however, it remains unclear whether such approaches can prevent onset of illness. Future intervention studies are encouraged to use the risk calculator at end-stage analysis to determine whether treatment efficacy is moderated by initial risk level or profile.

Ultimately, the degree of risk estimated by the risk calculator may be useful for weighing the cost-benefit ratios of various treatment options that emerge from clinical intervention research in the clinical high-risk population. Treatments associated with greater risks to the patient (e.g., medication side effects) or greater costs to health care delivery systems (e.g., resource- and time-intensive psychotherapeutic inventions) may best be reserved for those with higher-than-median levels of predicted risk (i.e., $\geq$0.16), while cost-effective treatments with benign side effect profiles may be the best option for those whose predicted risk for psychosis is in the lower range.

Like a person at risk for cardiovascular disease or cancer, an individual with a prodromal risk syndrome is more interested in receiving information pertinent to his or her personal risk profile than information about the population at large. Publication of this risk calculator is intended to assist clinicians in providing such personalized risk estimates. It is of course possible for untrained individuals to access these tools and approximate their scores on the set of risk variables. If, in so doing, a high predicted risk of conversion were generated, this could lead to significant personal distress. To mitigate this possibility, we have built in a decision tree for the online calculator that requires confirmation of an interview-based SIPS diagnosis of a prodromal risk syndrome and confirmation that the ratings and test scores were obtained by a professional; if either one of these verifications are missing, the

decision tree opts out of making a prediction. The risk for loss of privacy or stigmatization based on access of the prediction tool by untrained users is also mitigated for these reasons.

In summary, a well-performing risk calculator for psychosis is available for application to new patients who meet criteria for a psychosis risk syndrome. Challenges to be addressed in the next phase of research include incorporating biological assays into the risk calculations, extending the analysis to predict likelihood of remission, extending the framework to calculate reductions in risk based on particular interventions, and investigating how patients and family members feel about and use this information.

AUTHOR AND ARTICLE INFORMATION

From the Departments of Psychology and Psychiatry, Yale University, New Haven, Conn.; the Department of Quantitative Health Sciences, Cleveland Clinic, Cleveland; the Hotchkiss Brain Institute, Department of Psychiatry, University of Calgary, Alberta; the Departments of Psychiatry and Biobehavioral Sciences and Psychology, UCLA, Los Angeles; the Department of Psychiatry, UCSD, San Diego; the Department of Psychiatry, Zucker Hillside Hospital, Glen Oaks, New York; NIMH, Bethesda, Md.; the Renaissance Computing Institute, University of North Carolina, Chapel Hill; the Department of Psychiatry, UCSF, San Francisco; the Department of Psychiatry, University of North Carolina, Chapel Hill; the Department of Psychiatry, Harvard Medical School at Beth Israel Deaconess Medical Center and Massachusetts General Hospital, Boston; the Center for Behavioral Genomics, the Department of Psychiatry, and the Institute of Genomic Medicine, UCSD, La Jolla; and the Departments of Psychology and Psychiatry, Emory University, Atlanta.

Address correspondence to Dr. Cannon (tyrone.cannon@yale.edu).

Supported by NIH grants U01 MH081902 (to Dr. Cannon), P50 MH066286 (to Dr. Bearden), U01 MH081857 (to Dr. Cornblatt), U01 MH82022 (to Dr. Woods), U01 MH066134 (to Dr. Addington), U01 MH081944 (to Dr. Cadenhead), R01 U01 MH066069 (to Dr. Perkins), R01 MH076989 (to Dr. Mathalon), U01 MH081928 (to Dr. Seidman), and U01 MH081988 (to Dr. Walker).

Dr. Cannon has served as a consultant for the Los Angeles County Department of Mental Health and Boehringer-Ingelheim Pharmaceuticals. Dr. Mathalon has served as a consultant for Boehringer-Ingelheim Pharmaceuticals. Dr. Perkins has served as a consultant for Genentech, Lundbeck, Otsuka, and Sunovion, has participated in educational activity for Otsuka, and has received research support from Genentech. Dr. Woods has received investigator-initiated research support from Pfizer and sponsor-initiated research support from Auspex and Teva; he has served as a consultant for Biomedisyn (unpaid), Boehringer-Ingelheim, and Merck and as an unpaid consultant to DSM-5; he has been granted a patent for a method of treating prodromal schizophrenia with glycine agonists and is named as an inventor on a patent pending for a method of predicting psychosis risk using blood biomarker analysis; and he has received royalties from Oxford University Press. The other authors report no financial relationships with commercial interests.

Received July 8, 2015; revisions received Dec. 14, 2015, and March 8, 2016; accepted March 30, 2016.

REFERENCES

1. Insel TR: The arrival of preemptive psychiatry. Early Interv Psychiatry 2007; 1:5–6
2. Fusar-Poli P, Bonoldi I, Yung AR, et al: Predicting psychosis: meta-analysis of transition outcomes in individuals at high clinical risk. Arch Gen Psychiatry 2012; 69:220–229
3. Cannon TD, Cadenhead K, Cornblatt B, et al: Prediction of psychosis in youth at high clinical risk: a multisite longitudinal study in North America. Arch Gen Psychiatry 2008; 65:28–37
4. Fusar-Poli P, Borgwardt S, Bechdolf A, et al: The psychosis high-risk state: a comprehensive state-of-the-art review. JAMA Psychiatry 2013; 70:107–120
5. Kattan MW, Yu C, Stephenson AJ, et al: Clinicians versus nomogram: predicting future technetium-99m bone scan positivity in patients with rising prostate-specific antigen after radical prostatectomy for prostate cancer. Urology 2013; 81:956–961
6. Lee TH, Marcantonio ER, Mangione CM, et al: Derivation and prospective validation of a simple index for prediction of cardiac risk of major noncardiac surgery. Circulation 1999; 100:1043–1049
7. Pfeiffer RM, Park Y, Kreimer AR, et al: Risk prediction for breast, endometrial, and ovarian cancer in white women aged 50 y or older: derivation and validation from population-based cohort studies. PLoS Med 2013; 10:e1001492
8. Ross PL, Gerigk C, Gonen M, et al: Comparisons of nomograms and urologists' predictions in prostate cancer. Semin Urol Oncol 2002; 20:82–88
9. Specht MC, Kattan MW, Gonen M, et al: Predicting nonsentinel node status after positive sentinel lymph biopsy for breast cancer: clinicians versus nomogram. Ann Surg Oncol 2005; 12:654–659
10. Addington J, Cadenhead KS, Cornblatt BA, et al: North American Prodrome Longitudinal Study (NAPLS 2): overview and recruitment. Schizophr Res 2012; 142:77–82
11. McGlashan TH, Walsh BC, Woods SW: The Psychosis-Risk Syndrome: Handbook for Diagnosis and Follow-Up. New York, Oxford University Press, 2010
12. First M, Spitzer RL, Gibbon M, et al: Structured Clinical Interview for DSM-IV Axis I Disorders, Patient Edition. New York, Biometrics Research Department, New York State Psychiatric Institute, 1995
13. Miller TJ, McGlashan TH, Rosen JL, et al: 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
14. Kirkbride JB, Fearon P, Morgan C, et al: Heterogeneity in incidence rates of schizophrenia and other psychotic syndromes: findings from the 3-center AeSOP study. Arch Gen Psychiatry 2006; 63:250–258
15. Cannon TD, Chung Y, He G, et al: Progressive reduction in cortical thickness as psychosis develops: a multisite longitudinal neuroimaging study of youth at elevated clinical risk. Biol Psychiatry 2015; 77:147–157
16. McGlashan TH, Hoffman RE: Schizophrenia as a disorder of developmentally reduced synaptic connectivity. Arch Gen Psychiatry 2000; 57:637–648
17. Walker EF, Walder DJ, Reynolds F: Developmental changes in cortisol secretion in normal and at-risk youth. Dev Psychopathol 2001; 13:721–732
18. Riecher-Rössler A, Pflueger MO, Aston J, et al: Efficacy of using cognitive status in predicting psychosis: a 7-year follow-up. Biol Psychiatry 2009; 66:1023–1030
19. Seidman LJ, Giuliano AJ, Meyer EC, et al: Neuropsychology of the prodrome to psychosis in the NAPLS consortium: relationship to family history and conversion to psychosis. Arch Gen Psychiatry 2010; 67:578–588
20. Giuliano AJ, Li H, Mesholam-Gately RI, et al: Neurocognition in the psychosis risk syndrome: a quantitative and qualitative review. Curr Pharm Des 2012; 18:399–415
21. Keefe RS, Goldberg TE, Harvey PD, et al: The Brief Assessment of Cognition in Schizophrenia: reliability, sensitivity, and comparison with a standard neurocognitive battery. Schizophr Res 2004; 68:283–297
22. Brandt J, Benedict RHB: Hopkins Verbal Learning Test–Revised (HVLT-R). Odessa, Fla, Psychological Assessment Resources, 1998
23. Cornblatt BA, Carrion RE, Auther A, et al: Psychosis prevention: a modified clinical high risk perspective from the Recognition and Prevention (RAP) Program. Am J Psychiatry 2015; 172:986–994
24. Cornblatt BA, Auther AM, Niendam T, et al: Preliminary findings for two new measures of social and role functioning in the

prodromal phase of schizophrenia. Schizophr Bull 2007; 33: 688–702
25. Holtzman CW, Shapiro DI, Trotman HD, et al: Stress and the prodromal phase of psychosis. Curr Pharm Des 2012; 18:527–533
26. Dohrenwend BS, Krasnoff L, Askenasy AR, et al: Exemplification of a method for scaling life events: the Peri Life Events Scale. J Health Soc Behav 1978; 19:205–229
27. Janssen I, Krabbendam L, Bak M, et al: Childhood abuse as a risk factor for psychotic experiences. Acta Psychiatr Scand 2004; 109: 38–45
28. Sullivan PF: The genetics of schizophrenia. PLoS Med 2005; 2:e212
29. Uno H, Cai T, Pencina MJ, et al: On the C-statistics for evaluating overall adequacy of risk prediction procedures with censored survival data. Stat Med 2011; 30:1105–1117
30. Carrion RE, Cornblatt B, Burton CZ, et al: Personalized prediction of psychosis: external validation of the NAPLS 2 psychosis risk calculator with the EDIPPP project. Am J Psychiatry (Epub ahead of print, July 1, 2016)
31. Ruhrmann S, Schultze-Lutter F, Salokangas RK, et al: 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
32. Kattan MW: Doc, what are my chances? A conversation about prognostic uncertainty. Eur Urol 2011; 59:224
33. Koutsouleris N, Riecher-Rossler A, Meisenzahl EM, et al: Detecting the psychosis prodrome across high-risk populations using neuroanatomical biomarkers. Schizophr Bull 2015; 41:471–482
34. Perkins DO, Jeffries CD, Addington J, et al: Towards a psychosis risk blood diagnostic for persons experiencing high-risk symptoms: preliminary results from the NAPLS project. Schizophr Bull 2015; 41: 419–428
35. Addington J, Cornblatt BA, Cadenhead KS, et al: At clinical high risk for psychosis: outcome for nonconverters. Am J Psychiatry 2011; 168: 800–805
36. Schlosser DA, Jacobson S, Chen Q, et al: Recovery from an at-risk state: clinical and functional outcomes of putatively prodromal youth who do not develop psychosis. Schizophr Bull 2012; 38:1225–1233
37. van der Gaag M, Smit F, Bechdolf A, et al: Preventing a first episode of psychosis: meta-analysis of randomized controlled prevention trials of 12 month and longer-term follow-ups. Schizophr Res 2013; 149:56–62
38. McGorry PD, Yung AR, Phillips LJ, et al: Randomized controlled trial of interventions designed to reduce the risk of progression to first-episode psychosis in a clinical sample with subthreshold symptoms. Arch Gen Psychiatry 2002; 59:921–928
39. McGlashan TH, Zipursky RB, Perkins D, et al: Randomized, double-blind trial of olanzapine versus placebo in patients prodromally symptomatic for psychosis. Am J Psychiatry 2006; 163:790–799
40. Amminger GP, Schäfer MR, Papageorgiou K, et al: Long-chain omega-3 fatty acids for indicated prevention of psychotic disorders: a randomized, placebo-controlled trial. Arch Gen Psychiatry 2010; 67:146–154
41. Miklowitz DJ, O'Brien MP, Schlosser DA, et al: Family-focused treatment for adolescents and young adults at high risk for psychosis: results of a randomized trial. J Am Acad Child Adolesc Psychiatry 2014; 53:848–858
42. Stafford MR, Jackson H, Mayo-Wilson E, et al: Early interventions to prevent psychosis: systematic review and meta-analysis. BMJ 2013; 346:f185

Personalized Prediction of Psychosis: External Validation of the NAPLS-2 Psychosis Risk Calculator With the EDIPPP Project

Ricardo E. Carrión, Ph.D., Barbara A. Cornblatt, Ph.D., M.B.A., Cynthia Z. Burton, Ph.D., Ivy F. Tso, Ph.D., Andrea M. Auther, Ph.D., Steven Adelsheim, M.D., Roderick Calkins, Cameron S. Carter, M.D., Ph.D., Tara Niendam, Ph.D., Tamara G. Sale, M.A., Stephan F. Taylor, M.D., William R. McFarlane M.D.

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

Method: Of the total EDIPPP sample of 210 subjects rated as being at clinical high risk based on the Structured Interview for Prodromal Syndromes, 176 had at least one follow-up assessment and were included in the construction of a new prediction model with six predictor variables in the NAPLS-2 psychosis risk calculator (unusual thoughts and suspiciousness, symbol coding test performance, verbal learning test performance, decline in social functioning, baseline age, and family history). Discrimination performance was assessed with the area under the receiver operating characteristic curve (AUC). The NAPLS-2 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 0.790 (95% CI=0.644–0.937). In addition, the personalized risk generated by the risk calculator provided a solid estimation of the actual conversion outcome in the validation sample.

Conclusions: Two independent samples of clinical high-risk patients converge to validate the NAPLS-2 psychosis risk calculator. This prediction calculator represents a meaningful step toward early intervention and the personalized treatment of psychotic disorders.

Am J Psychiatry 2016; 173:989–996; doi: 10.1176/appi.ajp.2016.15121565

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 cancers (4–6), type 2 diabetes (7, 8), and cardiovascular disease (9–11). Surprisingly, risk calculators for mental health conditions are almost nonexistent, even though serious mental illness costs the United States $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 for psychosis, nearly 30% of whom will develop a psychotic illness over 2 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 toward enhancing prediction. However, progress has been less rapid in determining how to individualize prediction and determine the probability of psychosis risk on a case-by-case basis. At present, in clinical settings, a mental health professional can derive a general estimate of risk of psychosis for a given patient from the presence of

See related feature: **Editorial** *by Dr. Carpenter (p. 949)*

traditional risk factors. However, unlike in other fields of medicine, there is no widely available tool for calculating a more precise estimate of risk for a help-seeking or referred individual.

In an article published concurrently with this one, Cannon and colleagues (20), as part of the second phase of the North American Prodrome Longitudinal Study (NAPLS-2) (21), report on a risk calculator for the individualized prediction of a psychotic disorder over 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 must be validated in samples of clinical high-risk patients collected independently of NAPLS-2 (22). In the present study, we evaluated the performance of the NAPLS-2 risk calculator in an external, independent sample of individuals at clinical high risk 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 (23, 24) in preventing the onset of psychosis (25). Over the span of 3 years (September 2007 through June 1, 2010), EDIPPP recruited a large sample (N=337) of adolescents and young adults who were at risk or were in the very early stages of a major psychotic disorder.

The EDIPPP sample offers the opportunity for a clear test of the applicability of the NAPLS-2 risk calculator to an independent sample of clinical high-risk subjects, as the two projects had key differences in goals, recruitment strategies, and ascertainment criteria. The EDIPPP project established a community education and outreach network at six urban and rural sites across the United States 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 Family-Aided Assertive Community Treatment intervention, to reduce the incidence of psychosis (26). In EDIPPP, allocation to treatment was based on clinical risk (higher versus lower), which was determined by a cutoff of 7 on the total attenuated positive symptom severity score as specified by the Scale of Prodromal Symptoms from the Structured Interview for Prodromal Syndromes (SIPS) (25–28). EDIPPP participants thus had a wide range of attenuated positive symptom severity levels, from no symptoms in the prodromal range to positive symptoms that reached the threshold for psychosis. Because the original EDIPPP sample was categorized differently, this external validation sample was reconfigured to match the NAPLS-2 intake criteria, which includes all subjects meeting the criteria for prodromal syndromes as defined by the SIPS (29–31).

In the present study, our aim was to assess the predictive ability of the NAPLS-2 psychosis risk calculator in an external, independent sample of patients at clinical high risk for psychosis. Thus, in this report, we refer to the NAPLS-2 sample as the development sample and the EDIPPP sample as the external validation sample. Given that the predictors in the NAPLS-2 calculator were based on theoretical considerations, we first evaluated the predictive ability of the components of the NAPLS-2 model. Six key predictor variables used in the NAPLS-2 psychosis risk calculator were used in the external validation sample to construct a new model predicting psychosis. Second, we assessed the performance of the NAPLS-2 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 clinical high-risk 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 (22).

METHOD

The data reported here were collected as part of EDIPPP, a large multisite clinical trial for preventing psychosis among young people, funded by the Robert Wood Johnson Foundation (2007–2011) (25–27). EDIPPP consisted of six participating sites: Portland Identification and Early Referral, Portland, Maine; the Recognition and Prevention Program, Zucker Hillside Hospital, Glen Oaks, N.Y.; the Michigan Prevents Prodromal Progression Program, Ann Arbor, Mich.; the Early Assessment and Support Team Program, Salem, Ore.; the Early Diagnosis and Preventive Treatment Clinic, Sacramento, Calif.; and Early Assessment and Resource Linkage for Youth, Albuquerque, N.M. Details of the study design, study implementation, assessments, psychosocial and pharmacological treatments, methods, and sample characteristics have been reported elsewhere (25, 26). Although the Zucker Hillside Hospital site was part of both NAPLS-2 and EDIPPP, the present analyses included only four overlapping subjects, none of whom converted to psychosis and whose outcomes did not have an impact on the study findings.

Clinical high-risk subjects met criteria for one of the three prodromal syndromes based on the SIPS (29–31): attenuated positive symptom syndrome, with the presence of one or more moderate, moderately severe, or severe attenuated positive symptoms (scores of 3, 4, or 5 on the Scale of Prodromal Symptoms, on a scale of 0–6); genetic risk and deterioration syndrome, with genetic risk for psychosis coupled with deterioration in functioning; and brief intermittent psychotic syndrome, with intermittent psychotic symptoms that are recent, brief in duration, and not seriously disorganizing or dangerous.

A total of 210 clinical high-risk subjects were included in the validation sample; 92% had attenuated positive symptom syndrome, 6.3% had brief intermittent psychotic syndrome, and 1.7% had genetic risk and deterioration syndrome. In addition to these subjects, the EDIPPP sample included 32 early first-episode psychosis and 95 low-risk comparison subjects, who were excluded from the present study.

The EDIPPP study included participants 12–25 years old. Exclusion criteria for the study were a current or previous frank psychotic episode; treatment with antipsychotic medication for ≥30 days at a dosage appropriate for treating a psychotic episode; an IQ <70; permanent residence outside the catchment area; lack of fluency in English; current incarceration; and psychotic symptoms due to an acute toxic or medical cause.

Patients age 18 or older provided written informed consent; for patients under 18, the parents provided written informed consent and the patient provided written assent. The research protocol was approved by all sites' institutional review boards.

Baseline Assessments

Details of the baseline clinical assessment have been reported previously (25). Prodromal symptoms were assessed by the SIPS and the companion Scale of Prodromal Symptoms (29–31). Social and role functioning was assessed using the Global Functioning: Social and Global Functioning: Role scales (32). In addition to several other clinical measures, the baseline assessment included the Measurement and Treatment Research to Improve Cognition in Schizophrenia Consensus Cognitive Battery (33). The present analyses utilized data from two of the tests—the symbol coding subtest of the Brief Assessment of Cognition in Schizophrenia (BACS) (34) and the Hopkins Verbal Learning Test–Revised (35).

Clinical Outcome

Of the initial 210 clinical high-risk subjects, 176 (83.8%) had at least one follow-up assessment. Of those, 12 (6.8%) transitioned to psychosis over 2 years of follow-up (25). Conversion to psychosis was defined according to the Presence of Psychosis Scale criteria on the SIPS: developing any psychotic-level-intensity positive symptom (score of 6) that is sustained for at least an hour per day, at an average of 4 days per week over 1 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 Analysis

All analyses were conducted using SPSS Statistics 20.0 (IBM, Armonk, N.Y.). Comparisons of demographic and clinical characteristics were performed with Student's t tests for continuous variables and chi-square tests for categorical variables (two-tailed, $p<0.05$). Overall, 2.4% of the data (25 of 1,056 values) were missing. Among participants followed, missing values were imputed using mean values for scores on the BACS symbol coding test and the Hopkins Verbal Learning Test–Revised (missing one value each), and modal values for family history (missing 14 values) and decline in Global Functioning: Social scale (missing nine 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 and 95% confidence intervals for risk of conversion to psychosis in the external validation sample. We evaluated the ability of six predictor variables used in the NAPLS-2 psychosis risk calculator to predict psychosis: baseline age; severity of SIPS items P1 and P2 (unusual thought content and suspiciousness) recoded (18); raw score on the BACS symbol coding test; the sum of trials 1–3 on the Hopkins Verbal Learning Test–Revised, (35); a decline in social functioning in the year prior to the baseline assessment, measured using the Global Functioning: Social scale (32, 36, 37); and having a first-degree relative with a psychotic disorder (critical alpha, 0.05). 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 characteristic curve (AUC, equivalent to Harrell's c-statistic) (38, 39). The NAPLS-2 calculator was then used to generate risk estimates for each case in the external validation sample. The gamma statistic was used to examine 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 (40). The gamma statistic ranges from −1 (perfect negative association) to +1 (perfect agreement), with a value of 0 indicating no association. Precision and bias of the NAPLS-2 calculator were assessed with the Brier score and mean prediction error, respectively. The Brier score is the mean squared difference between the observed outcome and the predicted risk score (41). Mean prediction error was calculated as the difference between the risk estimates generated by the external validation model and the NAPLS-2 psychosis risk calculator, providing a metric for the tendency for over- or underestimation of risk (42). For both measures, a lower score indicates higher precision and less bias; ≤15% was considered to be an acceptable level (43). 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, negative predictive value) of the NAPLS-2 calculator was examined across different levels of predicted risk.

RESULTS

Table 1 summarizes baseline demographic and clinical data for the clinical high-risk sample. There was no difference between subjects with and without follow-up on any major demographic or clinical variable, including baseline age, gender, race, and education level. Clinical high-risk subjects in the validation sample had a mean age of 16.6 years (SD=3.3), and majorities were male (58.5%) and white (64.1%). The populations in the external validation (EDIPPP) sample and the development (NAPLS-2) sample were similar on most of the major demographic and clinical features. However, patients in the external validation sample were markedly younger on average than those in the development sample (mean ages of 16.6 years and 18.5 years, respectively), most likely because cases were referred from school-based sources.

Table 2 presents the regression model constructed with the validation sample that includes the six key variables

TABLE 1. Baseline Demographic and Clinical Characteristics of the EDIPPP Validation Sample[a]

Characteristic	Followed (N=176)		Not Followed (N=34)	
	Mean	SD	Mean	SD
Age (years)	16.6	3.3	16.0	3.1
Education (years)	9.8	2.6	9.6	2.7
Modified SIPS items P1 + P2[b]	2.4	1.9	2.4	2.1
BACS symbol coding test, raw score (number correct)	54.8	12.5	50.9	8.7
Hopkins Verbal Learning Test–Revised, trials 1–3 summed	25.8	5.5	25.0	4.7
	N	%	N	%
Family history of psychosis	37	22.8	5	17.2
Decline in Global Functioning: Social scale score >0	94	53.4	14	41.2
Male	103	58.5	19	55.9
Race				
White	109	64.1	20	66.7
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

[a] No significant difference between groups on any variable.
[b] Modified such that all levels in the nonprodromal range (0–2 on the original scale) are recoded as 0, levels in the prodromal range (3–5 on the original scale) are recoded as 1–3, and psychotic intensity (6 on the original scale) is recoded as 4.

selected from the NAPLS-2 psychosis risk calculator. The overall model was significant when all six independent variables were entered simultaneously (χ^2=19.68, df=6, p=0.003). The base model of SIPS items P1 and P2 showed an acceptable discrimination performance, with an AUC of 0.67. Combining the additional five variables with SIPS items P1 and P2 increased the AUC by 0.12, resulting in an AUC of 0.79 (95% CI=0.644–0.937, p=0.001), indicating good discrimination performance (Figure 1). Scores on the neurocognitive tests and baseline age were associated with the largest increases in the AUC when added to the base model (i.e., SIPS items P1 and P2).

As also shown in Table 2, in terms of individual variables, score on SIPS items P1 and P2 and baseline age bordered on statistical significance (p=0.05), while scores on the neurocognitive tests (the symbol coding test and the Hopkins Verbal Learning Test–Revised) only approached significance (p<0.10). A decline in social functioning and having a first-degree relative with psychosis were not significant predictors of psychosis in the validation sample.

The NAPLS-2 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 and the Brier score were at acceptable levels (<15%) (Brier score=7.5%, SD=15.8; mean prediction error=9.5%, SD=12.14), suggesting that the NAPLS-2 calculator provided a reasonable estimation of psychosis risk when comparing the risk prediction generated by the validation model compared with observed outcomes.

In addition, the risk estimates generated by the external validation model and the NAPLS-2 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 NAPLS-2 calculator and the external validation model (gamma=0.7, p<0.001).

Table 3 summarizes the performance of the NAPLS-2 calculator when applied to the external validation sample across increasing levels of model-predicted risk. The sensitivity and specificity values for each threshold were comparable to those observed in the development sample. For example, 10% model predicted risk provided a sensitivity of 91% and a specificity of 37% with the external validation sample, compared with a sensitivity of 94.1% and a specificity of 23.6% in the development 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 development model (66.7% sensitivity and 72.1% specificity).

DISCUSSION

This 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. Six risk factors from the NAPLS-2 calculator—baseline age, unusual thought content and suspiciousness, family history of a psychotic disorder, verbal learning, processing speed performance, and social decline—were able to distinguish individuals who developed psychosis from those who did not with a good degree of accuracy in the EDIPPP validation sample. In addition, there was good agreement between the risk prediction from the NAPLS-2 model and observed outcomes in the EDIPPP sample. Thus, this novel approach to constructing a psychosis prediction model using theoretical predictors has now been validated in two independent clinical high-risk samples: the development sample initially used to test the theoretical model (NAPLS-2) and an external, independent clinical high-risk sample (EDIPPP). To the best of our knowledge, this type of validation has not been performed on a psychosis prediction model. It 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 it has taken so long for a tool for a psychiatric disorder to be developed. Lack of replications of well-performing models and complex biological findings with limited clinical applicability may have contributed to this delay (44). Our findings highlight the importance of building a predictor model that includes a set

TABLE 2. Performance of Key Predictors From the NAPLS-2 Psychosis Risk Calculator in the EDIPPP Validation Sample

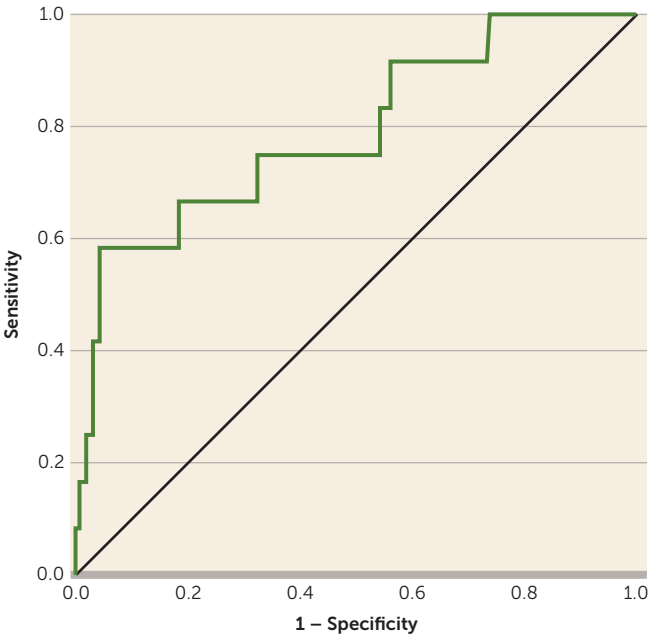
	Multivariable Model			AUC[a]	
Predictor	Hazard Ratio	95% CI	p	Decrement if Removed	Increase if added
Modified SIPS items P1 + P2	1.4	1.0–1.9	0.05	0.035	N/A[b]
Decline in Global Functioning: Social scale score	0.9	0.5–1.7	0.73	0.000	0.005
Hopkins Verbal Learning Test–Revised, trials 1–3 summed	0.9	0.8–1.0	0.10	0.013	0.092
BACS symbol coding test, raw score (number 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

[a] The receiver operating characteristic curve (AUC or c-statistic) was used to quantify the discrimination ability for separating converters and nonconverters. The AUC for the overall model was 0.790.
[b] The base model included only the modified SIPS items P1 and P2 scores; the AUC for the base model was 0.670.

of theoretically derived risk factors that have strong ties to vulnerability to the disease and can easily be applied in a clinical setting. The performance of the independent prediction model built with the EDIPPP sample using the risk factors included in the NAPLS-2 calculator showed good discrimination ability, comparable with that of the original development cohort; the overall model accuracy rates were 79% and 71%, respectively. Moreover, for the range of predicted risks that are adequately represented, the sensitivity and specificity for the levels of predicted risk generated by applying the NAPLS-2 calculator to the EDIPPP sample corresponded to levels of predicted risk of the development model seen in the Cannon et al. study (20). The values for positive predictive value in the present study were lower, however, than those observed in the development sample because of the lower conversion rate (i.e., prevalence) in the validation sample.

The discrimination accuracy of the base model (SIPS items P1 and 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 patients in clinical high-risk samples better than any individual predictor. It also potentially protects against 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 scores on neurocognitive tests 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 lower 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.

FIGURE 1. Receiver Operating Characteristic Curve for the EDIPPP Validation Model[a]



[a] EDIPPP=Early Detection, Intervention, and Prevention of Psychosis Program. 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. The AUC for this model was 0.790 (95% CI=0.644–0.937, p=0.001).

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 be used only by mental health professionals trained to a reliability standard in identifying the prodromal syndromes criteria with the SIPS/Scale of Prodromal Symptoms 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

TABLE 3. Prediction Statistics for Conversion to Psychosis Across Various Levels of NAPLS-2 Model-Predicted Risk

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] Positive predictive value=sensitivity × prevalence/sensitivity × prevalence + (1 − specificity) × (1 − prevalence).
[d] Negative predictive value=specificity × (1 − prevalence)/(1 − sensitivity) × prevalence + specificity × (1 − prevalence).

should discuss with the patient the cost-benefit ratio of treatment 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-making process between clinician and patient (45). Risk should not be overestimated, but at the same time, these estimates convey important information; for example, the difference between risks of 5% and 25% should have a meaningful impact on treatment decisions (46). Fourth, and of particular relevance to intervention, treatment recommendations should take into account the possible adverse effects, as well as the magnitude of the estimate (i.e., high versus low risk) and 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 (47). On the other hand, high-risk individuals or those 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 the balance of risks and benefits (1).

Next Steps: Integration Into Clinical Practice

The NAPLS-2 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. (22), four steps are involved in establishing a validated predictive tool and decision rules for use in clinical practice: selecting variables (level 4), validation at a single site or in a small prospective sample (level 3), validation at different sites (level 2), and then evaluation of 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 (22), which would involve demonstrating that the prediction tool can both change clinician behavior and 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 NAPLS-2 psychosis risk calculator in an independent clinical high-risk sample, the performance of the risk calculator needs further replications. The performance of the model should be evaluated on subcohorts (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 (48). 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 risk calculator to predict later illness are needed to validate this range as the appropriate target for intervention. Second, it is unclear how the NAPLS-2 risk calculator would perform on clinical high-risk cohorts obtained outside of North America, such as in 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 NAPLS-2 psychosis risk calculator, available on the Internet and incorporating a set of theoretically derived risk factors, can be replicated in a separate, independently collected clinical high-risk 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 it may provide a foundation for the personalized treatment of clinical high-risk individuals.

AUTHOR AND ARTICLE INFORMATION

From the Division of Psychiatry Research, Zucker Hillside Hospital, Northwell Health, Glen Oaks, N.Y.; the Center for Psychiatric Neuroscience, Feinstein Institute for Medical Research, Northwell Health, Manhasset, N.Y.; the Departments of Psychiatry and Molecular Medicine, Hofstra Northwell School of Medicine, Hempstead, N.Y.; the Department of Psychiatry, University of Michigan, Ann Arbor; the Department of

Psychiatry, Stanford University, Palo Alto, Calif.; the Imaging Research Center and the Center for Neuroscience, University of California Davis, Sacramento, Calif.; Portland State University Regional Research Institute, Portland, Ore.; the Mid-Valley Behavioral Care Network, Marion County Health Department, Salem, Ore.; Tufts University School of Medicine, Boston; and Maine Medical Center Research Institute, Portland.

Address correspondence to Dr. Carrión (rcarrion@northwell.edu).

Supported by NIMH grants MH61523 and MH081857 (to Dr. Cornblatt); NIH grants 5KL2TR000434-08 (to Dr. Tso), 5R01MH059883-11 (to Dr. Carter), R21MH101676 (to Dr. Taylor), K23MH087708 (to Dr. Niendam), and MH074543 (to the Zucker Hillside Hospital NIMH Advanced Center for Intervention and Services Research for the Study of Schizophrenia; John M. Kane, principal investigator); grant 67525 from the Robert Wood Johnson Foundation; and additional institutional support from the Maine Medical Center Research Institute and the state of Maine. Dr. Carrión received funding from the Brain and Behavior Research Foundation (NARSAD): Young Investigator Grant 19740 and Let the Sun Shine Run/Walk.

Dr. Cornblatt was the original developer of the Continuous Performance Test–Identical Pairs version. Dr. Taylor has received research support from Neuronetics, St. Jude's Medical, and Vanguard Research Group. Dr. McFarlane provides training on request to public and not-for-profit clinical services implementing psychosis early intervention programs. The other authors report no financial relationships with commercial interests.

Received Dec. 15, 2015; revision received March 11, 2016; accepted May 2, 2016; published online July 1, 2016.

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
2. Moons KG, Royston P, Vergouwe Y, et al: Prognosis and prognostic research: what, why, and how? BMJ 2009; 338:b375
3. D'Agostino RB Sr, Pencina MJ, Massaro JM, et al: Cardiovascular disease risk assessment: insights from Framingham. Glob Heart 2013; 8:11–23
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, et al: Assessing individual risk for prostate cancer. J Clin Oncol 2007; 25:3582–3588
6. Wells BJ, Kattan MW, Cooper GS, et al: 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
7. Hippisley-Cox J, Coupland C, Robson J, et al: Predicting risk of type 2 diabetes in England and Wales: prospective derivation and validation of QDScore. BMJ 2009; 338:b880
8. Lindström J, Tuomilehto J: The diabetes risk score: a practical tool to predict type 2 diabetes risk. Diabetes Care 2003; 26:725–731
9. Ridker PM, Buring JE, Rifai N, et al: Development and validation of improved algorithms for the assessment of global cardiovascular risk in women: the Reynolds Risk Score. JAMA 2007; 297:611–619
10. Ridker PM, Paynter NP, Rifai N, et al: C-reactive protein and parental history improve global cardiovascular risk prediction: the Reynolds Risk Score for men. Circulation 2008; 118:2243–2251
11. Pencina MJ, D'Agostino RB Sr, Larson MG, et al: Predicting the 30-year risk of cardiovascular disease: the Framingham Heart Study. Circulation 2009; 119:3078–3084
12. Insel TR: Assessing the economic costs of serious mental illness. Am J Psychiatry 2008; 165:663–665
13. Fusar-Poli P, Borgwardt S, Bechdolf A, et al: The psychosis high-risk state: a comprehensive state-of-the-art review. JAMA Psychiatry 2013; 70:107–120
14. Cannon TD, Cadenhead K, Cornblatt B, et al: Prediction of psychosis in youth at high clinical risk: a multisite longitudinal study in North America. Arch Gen Psychiatry 2008; 65:28–37
15. Nelson B, Yuen HP, Wood SJ, et al: Long-term follow-up of a group at ultra high risk ("prodromal") for psychosis: the PACE 400 study. JAMA Psychiatry 2013; 70:793–802
16. Ruhrmann S, Schultze-Lutter F, Salokangas RK, et al: 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
17. Cornblatt BA, Carrión RE, Auther A, et al: Psychosis prevention: a modified clinical high risk perspective from the Recognition and Prevention (RAP) program. Am J Psychiatry 2015; 172:986–994
18. Perkins DO, Jeffries CD, Cornblatt BA, et al: Severity of thought disorder predicts psychosis in persons at clinical high-risk. Schizophr Res 2015; 169:169–177
19. Michel C, Ruhrmann S, Schimmelmann BG, et al: A stratified model for psychosis prediction in clinical practice. Schizophr Bull 2014; 40: 1533–1542
20. Cannon TD, Yu C, Addington J, et al: An individualized risk calculator for research in prodromal psychosis. Am J Psychiatry 2016; 173:980–988
21. Addington J, Cadenhead KS, Cornblatt BA, et al: North American Prodrome Longitudinal Study (NAPLS 2): overview and recruitment. Schizophr Res 2012; 142:77–82
22. McGinn TG, Guyatt GH, Wyer PC, et al: Users' guides to the medical literature, XXII: how to use articles about clinical decision rules. JAMA 2000; 284:79–84
23. 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; spring (53):43–54
24. Fjell A, Bloch Thorsen GR, Friis S, et al: Multifamily group treatment in a program for patients with first-episode psychosis: experiences from the TIPS project. Psychiatr Serv 2007; 58:171–173
25. McFarlane WR, Levin B, Travis L, et al: 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
26. McFarlane WR, Cook WL, Downing D, et al: Early detection, intervention, and prevention of psychosis program: rationale, design, and sample description. Adolesc Psychiatry 2012; 2:112–124
27. Lynch S, McFarlane WR, Joly B, et al: Early Detection, Intervention, and Prevention of Psychosis Program: community outreach and early identification at six US sites. Psychiatr Serv 2016; 67:510–516
28. McFarlane WR, Cook WL, Downing D, et al: 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
29. Miller TJ, McGlashan TH, Rosen JL, et al: 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
30. Miller TJ, McGlashan TH, Rosen JL, et al: 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
31. Miller TJ, McGlashan TH, Woods SW, et al: Symptom assessment in schizophrenic prodromal states. Psychiatr Q 1999; 70:273–287
32. Cornblatt BA, Auther AM, Niendam T, et al: Preliminary findings for two new measures of social and role functioning in the prodromal phase of schizophrenia. Schizophr Bull 2007; 33:688–702
33. Green MF, Nuechterlein KH, Gold JM, et al: Approaching a consensus cognitive battery for clinical trials in schizophrenia: the NIMH-MATRICS conference to select cognitive domains and test criteria. Biol Psychiatry 2004; 56:301–307
34. Keefe R: Brief Assessment of Cognition in Schizophrenia (BACS) Manual–A: Version 2.1. Durham, NC, Duke University Medical Center, 1999

35. Benedict RHB, Schretlen D, Groninger L, et al: Hopkins Verbal Learning Test–Revised: normative data and analysis of inter-form and test-retest reliability. Clin Neuropsychol 1998; 12:43–55
36. Carrión RE, McLaughlin D, Goldberg TE, et al: Prediction of functional outcome in individuals at clinical high risk for psychosis. JAMA Psychiatry 2013; 70:1133–1142
37. Cornblatt BA, Carrión RE, Addington J, et al: Risk factors for psychosis: impaired social and role functioning. Schizophr Bull 2012; 38:1247–1257
38. Hanley JA, McNeil BJ: The meaning and use of the area under a receiver operating characteristic (ROC) curve. Radiology 1982; 143:29–36
39. Harrell FE: Regression Modeling Strategies: With Applications to Linear Models, Logistic Regression, and Survival Analysis. New York, Springer-Verlag, 2001
40. Siegel S, Castellan NJJ: Nonparametric Statistics for the Behavioral Sciences, 2nd ed. New York, McGraw-Hill, 1988
41. Brier GW: Verification of forecasts expressed in terms of probability. Mon Weather Rev 1950; 78:1–3
42. Sheiner LB, Beal SL: Some suggestions for measuring predictive performance. J Pharmacokinet Biopharm 1981; 9:503–512
43. Steyerberg E: Clinical Prediction Models: A Practical Approach to Development, Validation, and Updating, New York, Springer, 2008
44. 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
45. Stone NJ, Robinson JG, Lichtenstein AH, et al: 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(suppl 2):S1–S45
46. Van Calster B, Steyerberg EW, Harrell FH: Risk prediction for individuals. JAMA 2015; 314:1875
47. Loewy R, Fisher M, Schlosser DA, et al: Intensive auditory cognitive training improves verbal memory in adolescents and young adults at clinical high risk for psychosis. Schizophr Bull (Epub ahead of print, Feb 22, 2016)
48. DeFilippis AP, Young R, Carrubba CJ, et al: An analysis of calibration and discrimination among multiple cardiovascular risk scores in a modern multiethnic cohort. Ann Intern Med 2015; 162:266–275

of Neurology and Psychiatry, Tokyo; Psychiatrist Team Alkmaar West, Alkmaar, the Netherlands; the Department of Psychiatry, Nihon University School of Medicine, Tokyo; the Department of Affective Disorders, Institute of Psychiatry and Neurology, Warsaw; Maudsley Hospital and the Institute of Psychiatry, Psychology, and Neuroscience, King's College London, London; the Department of Psychiatry, Marmara University Hospital, Istanbul, Turkey; and the Institute of Mental Health, Peking University, Beijing.

Address correspondence to Dr. Benedetti (benedetti.francesco@hsr.it).

Dr. Avery has written articles for UpToDate. Dr. Henriksen is a shareholder in Chrono Chrome AS. Dr. Kasper received grants, research support, consulting fees, and/or honoraria within the last 3 years from Angelini, AOP Orphan Pharmaceuticals AG, AstraZeneca, Eli Lilly, Janssen, KRKA-Pharma, Lundbeck, Neuraxpharm, Pfizer, Pierre Fabre, Schwabe, and Servier. Dr. Lam has received speakers honoraria from the Canadian Network for Mood and Anxiety Treatments, the Canadian Psychiatric Association, Lundbeck, and Pfizer; he has been a consultant for or served on advisory boards of Akili, Allergan, the Asia-Pacific Economic Cooperation, the Canadian Depression Research and Intervention Network, the Canadian Network for Mood and Anxiety Treatments, the CME Institute, Janssen, Lundbeck, Medscape, Otsuka, and Pfizer; he has received research funds (through the University of British Columbia) from the BC Leading Edge Foundation, Brain Canada, the Canadian Institutes of Health Research, the Canadian Network for Mood and Anxiety Treatments, Janssen, Lundbeck, the Movember Foundation, Pfizer, St. Jude Medical, the University Health Network Foundation, the Vancouver Coastal Health Research Institute, and the VGH Foundation; he owns a patent related to the Lam Employment Absence and Productivity Scale (LEAPS); he receives royalties from Cambridge University Press, Informa Press, and Oxford University Press; and he owns stock in Mind Mental Health Technologies. Dr. Winkler has received lecture fees from Angelini, Lundbeck, and Pfizer. The other authors report no financial relationships with commercial interests.

Accepted March 28, 2018.

Am J Psychiatry 2018; 175:905–906; doi: 10.1176/appi.ajp.2018.18020231

Light Therapy and Risk of Hypomania, Mania, or Mixed State Emergence: Response to Benedetti et al.

TO THE EDITOR: We chose midday light for our randomized controlled trial of patients with bipolar disorder because of the findings from our pilot study (1). Three of our first four women with depression treated with antimanic drugs rapidly developed mixed states, which necessitated discontinuation of morning light therapy. However, we have recommended morning light therapy for patients who do not respond to 45–60 minutes of midday light therapy. The interpretation that morning light therapy is contraindicated is not consistent with our publications (1, 2).

Morning light therapy can elicit abrupt, large circadian rhythm phase advances that may precipitate bipolar switching, as has been described after eastward jet travel. Midday light therapy is far less likely to induce similar phase shifts and is a conservative initial treatment. The gradual emergence of group differences in our controlled study of midday light therapy (2) contrasts with the rapid improvement often seen with morning light therapy, which may reflect the relative circadian rhythm potency associated with the timing of light therapy.

The claim of "proven efficacy and safety" of early morning bright light treatment for bipolar depression is overstated. Many of the publications on morning light therapy and bipolar disorder in Dr. Benedetti's review (3) included studies of seasonal depression and patients with both unipolar and bipolar disorder. Other studies were constrained by open trial design, lack of a comparator group, brief duration, inclusion of antidepressants with adjunctive light therapy, and light therapy combined with sleep deprivation. Assessing hypomanic or manic symptoms with a valid measure is necessary to quantify the rate of their emergence (4). Only 12 of 43 studies (3) included the administration of a mania scale, which will bias the results toward underestimating the occurrence of mixed states and hypomania.

With due respect to our colleagues, the extensive list of authors who "have used [morning light therapy] in everyday clinical practice" (as have we) cannot supplant controlled clinical trial data. In his comprehensive survey (3), Dr. Benedetti reported that morning light therapy has been compared with placebo for bipolar disorder in only three studies. Using the Young Mania Rating Scale in two of the studies, symptoms were absent or rare, while the third study lacked a standard mania measure. With midday light therapy, we did not observe any mixed states, hypomania or mania, or significant differences in scores on the mania rating scale. Direct comparisons of midday and morning light therapy in a randomized controlled trial, with attention to gender-specific rates and predictors of hypomania or mania and mixed state emergence, would be a valuable contribution.

REFERENCES

1. Sit D, Wisner KL, Hanusa BH, et al: Light therapy for bipolar disorder: a case series in women. Bipolar Disord 2007; 9:918–927
2. Sit DK, McGowan J, Wiltrout C, et al: Adjunctive bright light therapy for bipolar depression: a randomized double-blind placebo-controlled trial. Am J Psychiatry 2018; 175:131–139
3. Benedetti F: Rate of switch from bipolar depression into mania after morning light therapy: a historical review. Psychiatry Res 2018; 261: 351–356
4. Angst J, Adolfsson R, Benazzi F, et al: The HCL-32: towards a self-assessment tool for hypomanic symptoms in outpatients. J Affect Disord 2005; 88:217–233

Dorothy K. Sit, M.D.
Michael Terman, Ph.D.
Katherine L. Wisner, M.D., M.S.

From the Department of Psychiatry and Behavioral Sciences, Feinberg School of Medicine, Northwestern University, Chicago; and the Department of Psychiatry, Columbia University and New York State Psychiatric Institute, New York.

Address correspondence to Dr. Sit (dorothy.sit@northwestern.edu).

The authors' disclosures accompany the original article.

Accepted March 28, 2018.

Am J Psychiatry 2018; 175:906; doi: 10.1176/appi.ajp.2018.18020231r

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

TO THE EDITOR: A web-based risk calculator (http://riskcalc.org:3838/napls/) for use in clinical high-risk populations was developed in the second phase of the North American Prodrome Longitudinal Study (NAPLS-2) (1). This calculator

integrated baseline age, unusual thoughts and suspiciousness, symbol coding, verbal learning test performance, functional decline, and family history of psychosis variables and achieved a concordance index of 0.71 for predicting psychosis. A study in an independent U.S. sample validated the risk calculator and provided supporting evidence for its application and dissemination (2). Should robust cross-validations occur in different countries with different populations, this would strengthen its potential use clinically in early identification and intervention programs treating individual clinical high-risk cases. There are many steps needed before such tools can be implemented. At this point, cross-validation in other independent samples is an important step that would strengthen the evidence base for use of the risk calculator.

TABLE 1. Comparison of Characteristics of Clinical High-Risk Subjects Who Were in the SHARP Program or in the NAPLS-2[a]

Variable	NAPLS-2 (N=596; followed)		SHARP (N=199; followed)		Statistical Analysis	
	Mean	SD	Mean	SD	t	p
Age (years)	18.5	4.3	19.1	5.1	1.695	0.092
Modified P1 and P2 SIPS items[b]	2.6	1.6	3.1	1.5	5.168	<0.001
Hopkins Verbal Learning Test–Revised (raw score)	25.6	5.2	23.5	5.4	−5.534	<0.001
Brief Assessment of Cognition in Schizophrenia symbol coding test (raw score)	56.8	13.1	57.9	10.0	1.493	0.137
	N	%	N	%	χ^2	p
Male	344	57.7	94	47.2	6.625	0.010
Family history of psychosis	96	16.1	17	8.5	7.001	0.008

[a] SHARP=Shanghai At Risk for Psychosis; NAPLS-2=second phase of the North American Prodrome Longitudinal Study.
[b] Represents the severity of unusual thought content and suspiciousness (items P1 and P2 in the Structured Interview for Prodromal Syndromes [SIPS]). P1 or P2 items rated 0–2 on the original scale are recoded as 0; items rated 3–6 on the original scale are recoded as 1–4.

An important question is how the NAPLS-2 risk calculator will work in other parts of the world, such as in Asian samples, that have different cultural and social backgrounds. From a validity standpoint, it is ideal for such replications to measure the same risk factors using very similar inclusion criteria and assessments comparable to the NAPLS. Such a study does exist. In 2010, the Shanghai At Risk for Psychosis (SHARP) study was launched at the Shanghai Mental Health Center, the largest outpatient mental health clinic in China (3, 4). The Chinese SHARP research and clinical team has been working closely with a U.S. team that was led by Larry J. Seidman, Ph.D., who was also the principal investigator at the Harvard Medical School site of the NAPLS project. Together, these teams have implemented methods similar to those used in the NAPLS for the identification of clinical high-risk individuals in mainland China in studies jointly funded by the National Institute of Mental Health and Chinese agencies.

A total of 300 clinical high-risk youths were identified using the Structured Interview for Prodromal Syndromes (SIPS). Among them, 228 (76.0%) completed neurocognitive assessments at baseline, 199 (87.3%) clinical high-risk youths had at least a 1-year follow-up assessment, and 46 (23.1%) converted to full psychosis. Details of the study procedures, study setting, implementation of the measurement, and assessment are reported elsewhere (3, 4). The clinical high-risk youths in the SHARP and NAPLS-2 samples were compared on demographic and clinical variables (Table 1). The six key predictor variables were entered into the NAPLS-2 risk calculator by two persons independently, and a new risk ratio variable for the Chinese clinical high-risk population was constructed. The only difference was that the Global Functioning: Social scale in the NAPLS-2 risk calculator was replaced by the Global Assessment of Functioning Scale (GAF) change score, which also measures functional deterioration (score relative to the previous 12 months). A GAF score that has declined to 5% or less of the previous best GAF score is recoded as 0. Declines of 5%–15% are recoded as 1, 15%−25% as 2, 25%−35% as 3, 35%−45% as 4, 45%−55% as 5, and 55%−65% as 6. Our data highlight the importance of a declining GAF score in the prediction of psychosis (4); that is, we found a significant positive association (r_s=0.884, N=200, p<0.001, Spearman rank-order correlation) and comparability for predicting psychosis by the receiver operating characteristic analysis between the GAF and the Global Functioning: Social scale in later samples that were acquired. Another reason for using the GAF score is that cultural differences have not been examined and may affect the validity of social functioning scales; otherwise, the GAF scores can be derived from the SIPS assessment and have been widely used in China for many years.

We investigated whether probability risk estimates provided by the NAPLS-2 risk calculator for each individual in the SHARP validation sample could discriminate converters from nonconverters. When conversion to psychosis is the principal endpoint, the receiver operating characteristic analysis resulted in an area under the curve (AUC) value of 0.631 (95% CI=0.542–0.721, p=0.007) for the probability risk estimates. Frequency distributions of predicted risks for converters and nonconverters in the SHARP sample are in good agreement with those obtained using the NAPLS-2 risk calculator. Converters occur at a proportionally higher rate than nonconverters at a predicted risk of 0.20 (χ^2=4.450, p=0.035) (Figure 1). In addition, Table 2 summarizes the performance of probability risk estimated by the NAPLS-2 risk calculator for the SHARP sample.

The aim of this study was to cross-validate the NAPLS-2 risk calculator in a Chinese clinical high-risk sample. To the best of our knowledge, this is the first attempt to verify the

FIGURE 1. Frequency Distributions of Predicted Risk Among Nonconverters and Converters in the Shanghai At Risk for Psychosis (SHARP) Sample

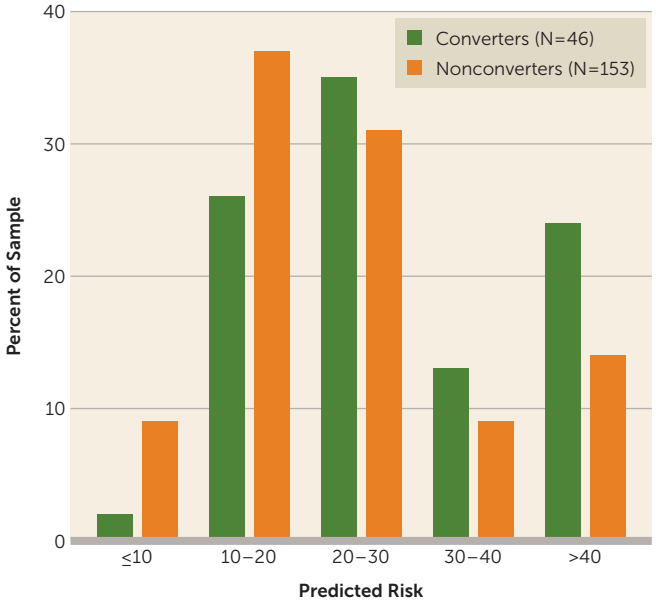


TABLE 2. Psychometric Property Values of the Predicted Risk Index for Conversion to Psychosis (or Nonrecovered)

Predicted Risk (%)	Sensitivity	Specificity	Positive Predictive Value[a]	Negative Predictive Value[b]
≥10	97.8	9.2	24.5	93.3
≥20	71.7	45.8	28.5	84.3
≥30	37.0	77.1	32.7	80.3
≥40	21.7	86.3	32.3	78.6
≥50	15.2	93.5	41.3	78.6

[a] Positive predictive value represents the proportions of positive results in a risk class at the specified level or higher that are true positive.
[b] Negative predictive value represents the proportions of negative results in a risk class at the specified level or higher that are true negative.

NAPLS-2 risk calculator using a comparable data set from an Asian sample and only the second to do so with a non-NAPLS sample, although the NAPLS-2 risk calculator did not fit our SHARP data as well as it fit the original sample. We believe that our slightly lower AUC is to be expected given a completely independent sample, which may be subject to the issue of statistical "shrinkage" (i.e., less good fit when applying regression models to new samples). This result suggests that the NAPLS-2 risk calculator has some generalizability to an Asian country and may have usefulness in clinical applications in China. This information provides a critical first step in the implementation of the NAPLS-2 risk calculator for the clinical high-risk population in China and supports the validity of the risk calculator in novel samples. However, as emphasized by Cannon et al. (1) and Carrión et al. (2) in the October 2016 issue of the *Journal*, the risk calculator remains experimental. At this point, it should be used only in research settings and with clinicians who have had rigorous SIPS training (SIPS scores being at the core of the model) and not yet used in general clinical settings with individuals until its clinical utility and properties are validated more firmly.

REFERENCES

1. Cannon TD, Yu C, Addington J, et al: An individualized risk calculator for research in prodromal psychosis. Am J Psychiatry 2016; 173:980–988
2. Carrión RE, Cornblatt BA, Burton CZ, et al: Personalized prediction of psychosis: external validation of the NAPLS-2 psychosis risk calculator with the EDIPPP project. Am J Psychiatry 2016; 173:989–996
3. Zhang T, Li H, Woodberry KA, et al: Prodromal psychosis detection in a counseling center population in China: an epidemiological and clinical study. Schizophr Res 2014; 152:391–399
4. Ren W, Ma J, Li J, et al: Repetitive transcranial magnetic stimulation (rTMS) modulates lipid metabolism in aging adults. Front Aging Neurosci 2017; 9:334

TianHong Zhang, M.D., Ph.D.
HuiJun Li, Ph.D.
YingYing Tang, Ph.D.
Margaret A. Niznikiewicz, Ph.D.
Martha E. Shenton, Ph.D.
Matcheri S. Keshavan, M.D.
William S. Stone, Ph.D.
Robert W. McCarley, M.D.
Larry J. Seidman, Ph.D.
JiJun Wang, M.D., Ph.D.

From the Shanghai Mental Health Center, Shanghai Jiaotong University School of Medicine, Shanghai Key Laboratory of Psychotic Disorders, Shanghai; the Department of Psychology, Florida A&M University, Tallahassee; the Department of Psychiatry, Harvard Medical School at Beth Israel Deaconess Medical Center, Boston; the Departments of Psychiatry and Radiology, Brigham and Women's Hospital, Boston; and the VA Boston Healthcare System, Boston.

Address correspondence to Dr. Wang (jijunwang27@163.com).

Dr. Seidman died in September 2017. Dr. McCarley died in May 2017.

Supported by a National Key R&D Program of China grant (2016YFC1306803) to Dr. Wang, by National Natural Science Foundation of China grants (81671329, 81671332) to Dr. Zhang and Dr. Wang, by an R21 Fogarty/NIMH grant (1R21 MH093294-01A1) to Dr. Li, and by a U.S.-China Program for Biomedical Collaborative Research grant (1R01 MH 101052-01) to Dr. Seidman.

Drs. Seidman and McCarley were founders and core members of the Shanghai At Risk for Psychosis (SHARP) project.

Funding agencies and organizations had no role in the design, analysis, interpretation, or publication of this study.

The authors report no financial relationships with commercial interests.

Accepted June 18, 2018.

Am J Psychiatry 2018; 175:906–908; doi: 10.1176/appi.ajp.2018.18010036

Understanding the Risk of Treatment Failure After Discontinuation of Long-Term Antipsychotic Treatment

TO THE EDITOR: Tiihonen et al. (1) should be commended for their follow-up investigation, published in the August 2018 issue of the *Journal*, of discontinuation of antipsychotic treatment in schizophrenia, using a nationwide cohort and a state-of-the-art sampling design. We think that some clarifications are warranted to interpret the statement that the "risk of treatment failure increases with duration of antipsychotic treatment before discontinuation" and the conclusion that long-term antipsychotic exposure "makes discontinuation more difficult when exposure has been longer."