Towards a Psychosis Risk Blood Diagnostic for Persons Experiencing High-Risk Symptoms: Preliminary Results From the NAPLS Project

The Harvard community has made this article openly available. Please share how this access benefits you. Your story matters.

Citation

Published Version
doi:10.1093/schbul/sbu099

Citable link
http://nrs.harvard.edu/urn-3:HUL.InstRepos:14065477

Terms of Use
This article was downloaded from Harvard University’s DASH repository, and is made available under the terms and conditions applicable to Other Posted Material, as set forth at http://nrs.harvard.edu/urn-3:HUL.InstRepos:dash.current.terms-of-use#LAA
Towards a Psychosis Risk Blood Diagnostic for Persons Experiencing High-Risk Symptoms: Preliminary Results From the NAPLS Project


1Department of Psychiatry, University of North Carolina, Chapel Hill, NC; 2Renaissance Computing Institute, University of North Carolina, Chapel Hill, NC; 3Department of Psychiatry, Hotchkiss Brain Institute, University of Calgary, Alberta, Canada; 4Departments of Psychiatry and Biobehavioral Sciences and Psychology, University of California, Los Angeles, Los Angeles, CA; 5Department of Psychiatry, University of California, San Diego, San Diego, CA; 6Department of Psychology, Yale University, New Haven, CT; 7Department of Psychiatry, Yale University, New Haven, CT; 8Department of Psychiatry, Zucker Hillside Hospital, Long Island, NY; 9Department of Psychiatry, University of California, San Francisco, San Francisco, CA; 10Department of Psychiatry, Harvard Medical School at Beth Israel Deaconess Medical Center and Massachusetts General Hospital, Boston, MA; 11Department of Psychiatry, Center for Behavioral Genomics, Institute of Genomic Medicine, University of California, San Diego, La Jolla, CA; 12Departments of Psychology and Psychiatry, Emory University, Atlanta, GA; 13Division of Adult Translational Research and Treatment Development, National Institute of Mental Health, Bethesda, MD

*To whom correspondence should be addressed; Department of Psychiatry, University of North Carolina at Chapel Hill, CB 7160, Chapel Hill, NC 27599, US; tel: 919-962-1401, fax: 919-445-0414, e-mail: diana_perkins@med.unc.edu

Introduction: A barrier to preventative treatments for psychosis is the absence of accurate identification of persons at highest risk. A blood test that could substantially increase diagnostic accuracy would enhance development of psychosis prevention interventions. Methods: The North American Prodrome Longitudinal Study project is a multisite endeavor that aims to better understand predictors and mechanisms for the development of psychosis. In this study, we measured expression of plasma analytes reflecting inflammation, oxidative stress, hormones, and metabolism. A “greedy algorithm” selected analytes that best distinguished persons with clinical high-risk symptoms who developed psychosis (CHR-P; n = 32) from unaffected comparison (UC) subjects (n = 35) and from those who did not develop psychosis during a 2-year follow-up (CHR-NP; n = 40). Results: The classifier included 15 analytes (selected from 117), with an area under the receiver operating curve for CHR-P vs UC of 0.91 and CHR-P vs CHR-NP of 0.88. Randomly scrambled group membership followed by reconstructions of the entire classifier method yielded consistently weak classifiers, indicating that the true classifier is highly unlikely to be a chance occurrence. Such randomization methods robustly imply the assays contain consistent information distinguishing the groups which was not obscured by the data normalization method and was revealed by classifier construction. These results support the hypothesis that inflammation, oxidative stress, and dysregulation of hypothalamic-pituitary axes may be prominent in the earliest stages of psychosis. Conclusion: If confirmed in other groups of persons at elevated risk of psychosis, a multiplex blood assay has the potential for high clinical utility.

Key words: clinical high risk/psychosis/prodrome/multiplex/risk prediction/modified low-density lipoprotein (MDA-LDL)/immune/inflammation/oxidative stress

Introduction

It is well established that early intervention is associated with better clinical outcomes in persons with schizophrenia,1 raising the hope that treatment during the prodromal phase of illness could prevent the development of a psychotic disorder and thus reduce risk of chronic symptoms and disability. Substantial progress has been made in establishing clinical criteria to identify persons at high risk (CHR) for psychosis, with about 20%–25% developing psychosis within a year and 30%–35% within 2 years.2,4 An “attenuated psychosis syndrome” was considered for inclusion in the most recent Diagnostic and Statistical Manual, Fifth Edition (DSM-5).5 However, the expert panel decided to place the syndrome in the DSM-5 Appendix due to concerns that specificity would be low and that a minority of persons meeting syndrome...
criteria would actually progress to develop a psychotic disorder.\textsuperscript{6,7} This decision reinforces the need for further research to identify biological validators of a high-risk state for psychosis warranting preventative interventions.

In general, biomarker tests are most useful for persons at elevated risk for a disease because the proportion of patients who test positive and who actually get the disease (positive predictive value, PPV) varies dramatically depending on the actual rate of disease in the tested population.\textsuperscript{8} If the rate of disease is low, eg, if 1\% of tested persons are actually at risk, then even a test with high sensitivity and specificity (~0.80) will have 20 times as many false positives as true positive (supplement S1). To minimize the risk of embarking on an inappropriate treatment program, any biomarker test with less than perfect specificity needs be applied in a relatively high-risk population. Applied to persons meeting CH\textsuperscript{R} criteria, a biomarker test with reasonable sensitivity and specificity (~0.8) can achieve a PPV of about 0.63; equivalently, about two-thirds of persons identified by such a test as at risk would truly be on a trajectory to develop psychosis. Of, perhaps, equal significance is that such a diagnostic test could have a negative predictive power of about 0.90, increasing a clinician’s confidence in predicting who is likely not to develop psychosis.

Patients with schizophrenia are reported to have anomalous levels of markers of inflammation, oxidative stress, metabolism, and hormonal status.\textsuperscript{9–15} For this reason, we investigated whether plasma analytes representing these components could be used to construct a classifier that distinguishes persons with CH\textsuperscript{R} risk symptoms who developed psychosis from persons with CH\textsuperscript{R} symptoms who did not develop psychosis during a 2-year follow-up period, as well as from unaffected comparison subjects.

Methods

Subjects

The aims and methods of North American Prodrome Longitudinal Study (NAPLS 2) were described in detail previously.\textsuperscript{16} To summarize, NAPLS 2 is an 8-site observational study of the predictors and mechanisms of conversion to psychosis in persons meeting the Criteria of Prodromal States (COPS).\textsuperscript{17} The NAPLS 2 cohort includes 765 clinical high-risk and 280 demographically similar unaffected comparison (UC) subjects aged between 12 and 35. The study was approved by the Institutional Review Board at each site, and each subject provided written informed consent or assent, with a parent or guardian also consenting for minor subjects.

The plasma analysis reported here was conducted in March 2012. We included all CH\textsuperscript{R} subjects with plasma samples either who had progressed to psychosis (CH\textsuperscript{R}-P, n = 32) or who had been followed 2 years and not progressed to psychosis (CH\textsuperscript{R}-NP, n = 40) as of February 2012. Note that because this is a nested case-control study conducted before completion of study follow-up, no inference can be made about conversion rates in this report. The unaffected comparison subjects (UC, n = 35) did not meet CH\textsuperscript{R} criteria or have a history of a psychotic disorder and were chosen to be demographically similar to the CH\textsuperscript{R} subjects.

Assessments

Clinical assessments were done every 6 months and subjects followed for up to 2 years. Participants were screened using the Structured Interview for Prodromal Syndromes (SIPS) and rated with the Scale of Prodromal Symptoms (SOPS) for the presence of one or more prodromal syndromes: attenuated psychotic symptoms, brief intermittent psychotic symptoms, substantial functional decline combined with a first-degree relative with a psychotic disorder, or schizotypal personality disorder in individuals younger than 18 years.\textsuperscript{17} Details regarding the SIPS, SOPS, and COPS are included in supplementary methods. The Structured Clinical Interview for DSM IV\textsuperscript{18} was used to determine psychiatric diagnoses.

Depressive symptoms were evaluated with the Calgary Depression Scale for Schizophrenia\textsuperscript{19} (CDSS) and anxiety symptoms with the Self-Rated Anxiety Scale\textsuperscript{20} (SAS). Data on prescription medications were based on self-reports. Socioeconomic status was estimated by maximum years of education of mother or father. Substance use was evaluated with the Alcohol/Drug Use Scales\textsuperscript{21} as a dichotomous variable, and by frequency of use (0 = no use, 5 = daily use, or for cigarettes > 25 per day).

Plasma Collection

Blood samples used in this analysis were drawn at the baseline visit in Becton Dickenson P100 blood collection tubes with ethylene diamine tetra-acetic acid as anticoagulant, proprietary protein stabilizers, and a mechanical separator. All samples were processed within 120 minutes (mean time to freezer = 28 minutes, SD = 2 minutes) and stored at −80°C until analysis.

Plasma Assay

Plasma samples were sent on dry ice to Myriad Rules Based Medicine, a biomarker testing laboratory that has maintained clinical laboratory improvement amendments–accreditation by the Commission on Office Laboratory Accreditation since 2006. Samples were analyzed with the Human Discovery Map assay, a Luminex bead-based multiplex immunoassay that included 185 analytes involved in hormonal responses, inflammation, growth, oxidative stress, and metabolism, all according to Rules-Based Medicine standard operating procedures. Technicians ran assays without knowledge of clinical status of the subjects.

Exclusion of Analytes

The assay included 185 analytes. We excluded 23 analytes that were not detected in ≥ 20\% of the subjects. Most
remaining analytes (80%) were detected in at least 90% of the 107 subjects (supplement S2).

Persons exhibiting CHR symptoms are often prescribed various medications by their health care provider, and in our CHR subjects, 19% were prescribed antipsychotics and 28% were prescribed antidepressants (table 1). In addition CHR subjects were more likely to use marijuana and nicotine, compared with UC subjects, and alcohol use occurred in almost half of subjects (table 1). Antipsychotics, antidepressants, marijuana, nicotine, and alcohol may impact inflammation, oxidative stress, and hormonal pathways, raising the possibility that drug use could confound differences in analyte expression comparisons between groups. To minimize risk of confounding, we excluded analytes with a possible relation to prescription of antipsychotics or antidepressants, or with self-reported current use of marijuana, nicotine, or alcohol, based on comparisons of analyte expression levels between persons using vs not using the substance. Consequently, 45 additional analytes were excluded.

Table 1. Demographic and Clinical Characteristics of Study Subjects

<table>
<thead>
<tr>
<th></th>
<th>Unaffected Comparison (UC) N = 35</th>
<th>Clinical High Risk, Not Psychotic (CHR-NP) N = 40</th>
<th>Clinical High Risk, Psychotic (CHR-P) N = 32</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age, average (SD)</td>
<td>20 (4.5)</td>
<td>19.5 (4.6)</td>
<td>19.2 (3.7)</td>
</tr>
<tr>
<td>Ancestry</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Caucasian, %</td>
<td>60</td>
<td>65</td>
<td>55</td>
</tr>
<tr>
<td>African, %</td>
<td>31</td>
<td>17.5</td>
<td>21</td>
</tr>
<tr>
<td>Asian, %</td>
<td>9</td>
<td>17.5</td>
<td>24</td>
</tr>
<tr>
<td>Sex, female, %</td>
<td>34</td>
<td>37.5</td>
<td>30.3</td>
</tr>
<tr>
<td>Socioeconomic status, average (SD)</td>
<td>4.8 (1.8)</td>
<td>4.5 (2.3)</td>
<td>4.5 (1.8)</td>
</tr>
<tr>
<td>Scale of Prodromal Symptom scores, average (SD)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Totala</td>
<td>5.06 (5.11)</td>
<td>36.63 (13.03)</td>
<td>43.81 (14.11)</td>
</tr>
<tr>
<td>Positivea</td>
<td>1.46 (1.84)</td>
<td>12.28 (4.74)</td>
<td>14.22 (3.92)</td>
</tr>
<tr>
<td>Negativea</td>
<td>1.23 (1.72)</td>
<td>11.38 (6.32)</td>
<td>13.58 (6.23)</td>
</tr>
<tr>
<td>Disorganizeda</td>
<td>0.91 (1.17)</td>
<td>5.05 (2.79)</td>
<td>6.45 (3.88)</td>
</tr>
<tr>
<td>Generalb</td>
<td>1.46 (1.80)</td>
<td>7.93 (4.46)</td>
<td>10.52 (4.44)</td>
</tr>
<tr>
<td>Calgary Depression Scale for Schizophrenia scores, average (SD)</td>
<td>0.89 (1.71)</td>
<td>5.00 (4.72)</td>
<td>6.88 (4.88)</td>
</tr>
<tr>
<td>Zung Self-Rated Anxiety Scale, average (SD)</td>
<td>28.79 (4.15)</td>
<td>45.80 (13.35)</td>
<td>48.39 (12.55)</td>
</tr>
<tr>
<td>Time blood draw, average (SD)</td>
<td>12:12 pm (1.85 h)</td>
<td>12:39 pm (2.0 h)</td>
<td>11:59 am (1.79 h)</td>
</tr>
<tr>
<td>Prescription medication</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Antipsychoticc</td>
<td>0%</td>
<td>25%</td>
<td>13%</td>
</tr>
<tr>
<td>Antidepressantd</td>
<td>1%</td>
<td>30%</td>
<td>25%</td>
</tr>
<tr>
<td>Stimulant</td>
<td>0%</td>
<td>8%</td>
<td>6%</td>
</tr>
<tr>
<td>Mood stabilizer</td>
<td>0%</td>
<td>5%</td>
<td>3%</td>
</tr>
<tr>
<td>Benzodiazepinee</td>
<td>0%</td>
<td>5%</td>
<td>13%</td>
</tr>
<tr>
<td>Nonsteroidal anti-inflammatory drug</td>
<td>0%</td>
<td>0%</td>
<td>0%</td>
</tr>
<tr>
<td>Antibiotic</td>
<td>0%</td>
<td>0%</td>
<td>0%</td>
</tr>
<tr>
<td>Substance use</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tobacco usef</td>
<td>9%</td>
<td>30%</td>
<td>44%</td>
</tr>
<tr>
<td>Alcohol use</td>
<td>46%</td>
<td>48%</td>
<td>38%</td>
</tr>
<tr>
<td>Marijuana useg</td>
<td>9%</td>
<td>25%</td>
<td>31%</td>
</tr>
<tr>
<td>Current comorbid Diagnostic and Statistical Manual of Mental Disorders IV diagnosis</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Depression disordersh,i</td>
<td>0%</td>
<td>45%</td>
<td>50%</td>
</tr>
<tr>
<td>Anxiety disordersi</td>
<td>3%</td>
<td>60%</td>
<td>56%</td>
</tr>
</tbody>
</table>

CHR-P vs UC t test P value < .0001, CHR-NP vs UC t test P value < .0001.
CHR-P vs CHR-NP t test P value = .02.
CHR-P vs UC Fisher Exact Test (FET) P value = .047, CHR-NP vs UC FET P value = .001.
CHR-P vs UC FET P value = .011, CHR-NP vs UC FET P value = .002.
CHR-P vs UC FET P value = .047.
CHR-P vs UC FET P value = .001, CHR-NP vs UC FET P value = .02.
CHR-P vs UC FET P value = .020, CHR-NP vs UC FET P value = .056.
CHR-P vs UC FET P value < .0001, CHR-NP vs UC FET P value < .0001.
Depression disorders include major depression, depressive disorder not otherwise specified, and dysthymic disorder. Anxiety disorders include obsessive compulsive disorder, post-traumatic stress disorder, panic disorder, agoraphobia, social phobia, specific phobia, and generalized anxiety disorder.
CHR-P vs UC FET P value < .0001, CHR-NP vs UC FET P value < .0001.
We conducted further analyses in the remaining 117 analytes.

**Normalization**

The normal plasma concentrations varied analyte-to-analyte up to 1 000 000-fold. So that results could be viewed on the same scale, we standardized (z score) the results for each analyte to the average and SD values of the unaffected comparison subjects.

**Reproducibility**

We had resources to analyze one technical replicate. The intraclass correlation coefficient (ICC) of the 117 analyte levels for the duplicate samples was 0.83; for the 15 analytes selected by the greedy algorithm described below, the ICC was 0.87 (supplement S3).

**Data Analyses**

Statistical analyses were performed with Excel (Microsoft), Unscrambler (CAMO Software AS), SAS (SAS), or MedCalc Statistical Software. We used EasyFit to test for rejection of Gaussian (normal) distribution by Kolmogorov-Smirnov (K-S) and Anderson-Darling (A-D) tests. We used the Excel add-in significance analysis of microarrays (SAM) to compare expression using the Wilcoxon statistic with 5000 permutations, and the Excel add-in Real Statistics Using Excel to calculate ICCs and Fisher Exact Tests.

Greedy algorithms are extensively used in bioinformatics to solve a variety of problems. In particular, they are well known to be capable of selecting collectively informative markers from large candidate sets. They linearly build marker selections and avoid brute force examination of all possible subsets of markers. In addition, with greedy algorithms weighting of individual analytes can be avoided, thus minimizing the risk of overfitting that can occur from techniques such as machine learning or statistical modeling that develop real-number weights for each analyte to optimize prediction.

The program first selected the very best single analyte for distinguishing the 3 groups. Then a second analyte was added that best improved performance, if possible. Additional analytes were selected and added until no further selection of any analyte improved performance. The chosen set of analytes created an “index” defined as the sum of the z scores of selected analytes. In effect, this was a weighted sum with all weights being 1; this was possible because most of analytes that moderately distinguished the 3 groups had higher values in CHR-P than CHR-NP and higher values in CHR-NP than UC. Performance was defined to be minimization of the sum of the squares of the Student’s t test P values for CHR-P vs CHR-NP and CHR-P vs UC.

Noting the general principles and pitfalls of classifier construction, classifier analyses on the results of the greedy algorithm were executed using 5-by-5-by-5-fold cross-validation with repeated random subsampling, implemented in Excel with macros and add-ins. Subjects in each group were randomly partitioned into ~20% subgroups. Four of the CHR-P, 4 of the CHR-NP, and 4 of the UC subgroups were used to train a classifier that was tested on the complementary subgroups. There are 125 possible different combinations of such tests (see supplementary methods). The entire process was repeated 20 times with 20 initial selections of the random 20% subgroups, for a total of 20 × 125 or 2500 executions of the greedy algorithm. A review of all the preliminary classifiers easily led to a final, integrated classifier because in the 125 tests several analytes were repeatedly selected, while other analytes were seldom or never selected. To evaluate consistency, we also applied a leave-one-out cross-validation procedure and compared results with those of the 5-fold cross-validation methods.

The area under the resulting receiver operating curve (AUC) assessed the capacity of the index to distinguish CHR-P from the other 2 groups using the exact (no smoothing) Hand-Till method with calculations of SE following the methods of Handley-McNeil. A receiver operating curve (ROC) is a plot of sensitivity (ie, test correctly predicts true positive declarations/all positives) vs the false positive rate (ie, 1 specificity, test correctly predicts true negative declarations/all negatives). Various threshold settings yield the points along the curve.

An AUC of 1.0 indicates perfect classification, and for large samples (> 100), an AUC of 0.5 indicates random classification. For 30 cases and controls, uniformly random numbers yield Hand-Till AUCs averaging 0.54 (SD = 0.03). Moreover, any classifier method applied to real data even after random relabeling of cases and controls can be expected to find weak patterns with AUCs > 0.5. This implies the necessity of checking a proposed classifier as follows.

Randomly reassigning our raw samples into bins of the sizes of the 3 groups and applying the very same classifier construction used in true classification should repeatedly lead to AUCs smaller than that of true classification. Such permutation tests of classifiers have a lengthy history and are considered powerful techniques of model validation.

**Evaluating Analyte Function**

The RBM panel of analytes was chosen because of its focus on inflammation, oxidative stress, metabolism, and hormonal status, factors consistently associated with schizophrenia. Thus, it is not appropriate to use tools that discover “enrichment” of chosen analytes when interpreting these results. Instead, we considered it reasonable to review the published literature for each selected analyte to evaluate the function of the analyte, relationships...
between the analytes (eg, cytokines regulating hypothalamic-pituitary function), and reports of associations with schizophrenia or psychosis. We also found overlap with previously published reports on plasma analytes that distinguished persons with schizophrenia using the same multianalyte platform\textsuperscript{45-47} (reported in supplements S4 and S5).

**Results**

**Participant Characteristics**

Table 1 provides study subject characteristics. All CHR subjects met attenuated positive symptom diagnostic criteria. The psychosis diagnoses at follow-up of the 32 CHR-P subjects included 13 with psychosis, not otherwise specified, 14 with schizophrenia, 2 with major depression with psychotic features, and 1 each with schizoaffective, delusional, or bipolar disorder.

**Plasma Analytes and Development of Psychosis**

Controlling for multiple comparisons, there were no analytes that were differentially expressed for CHR-P subjects compared with CHR-NP or UC subjects (supplement S6). Comparison of CHR-P with UC subjects using permutation analyses with SAM\textsuperscript{32} assuming a false discovery rate of less than 10% identified 18 analytes as differentially expressed (supplement S5). We can thus expect that of the 18 analytes about 2 are false positives. Conducting the same permutation analysis for CHR-P to CHR-NP identified only 2 differentially expressed analytes (supplement S5).

**Plasma Analytes and psychosis Risk Prediction**

An analyte could be chosen up to 125 times with each of the 20 runs (2500 total executions) in the cross-validation procedure. While somewhat different combinations of analytes were chosen every time, certain analytes were frequently chosen (figure 1, supplement S7). Likely, the most confidence in the informativeness of analytes should attend those most frequently chosen. For example, malondialdehyde-modified low-density lipoprotein was selected in almost all of the 2500 executions of the greedy algorithm. However, after the 15th most popular analyte, the frequency fell by 15%, suggesting a cut-off point and hence a selection of 15 analytes.

Best practices of classifier development mandate use of an external test set that is not used to derive the classifier, with a minimum of 20 subjects in each group.\textsuperscript{39} The number of subjects in this study was too small to set aside an external test set, so we applied the 15-analyte index to the full data for all 107 subjects (figures 2A and 2B). As shown in table 2, using the true data the sum of the most frequently chosen 15 analytes gave the highest AUC,

---

**Fig. 1.** Shown analytes were the most frequently appearing in 20 5-by-5-by-5 cross-validation trials, each trial testing 125 partitions to generate ~80% subsets of UC, CHR-NP, and CHR-P samples.
followed closely by AUCs using just the 9 or even 6 most frequently chosen analytes.

Classifier methods generally can find patterns in random data, but the AUCs of pseudo-classification should be weak compared with the AUC of a true classifier of true data. Thus, we repeatedly applied exactly the same classifier development to pseudo-data obtained by randomly scrambling all the 107 samples into bins of sizes 35, 40, and 32 (based on the number of subjects in each group). As shown in figures 3A and 3B, the AUCs of the true classifiers of UC vs CHR-P and CHR-NP vs CHR-P were greater than all AUCs of 100 pseudo-classifiers, respectively, applied to their pseudo-data. Fitting data with beta distributions yielded extrapolations with low P values, as shown.

The results of the leave-one-out validation procedure were essentially the same as the 5-fold cross-validation (supplement S8).

**Table 2.** Area Under the Receiver Operating Curve (AUC) Using Cut-off Points Chosen Based On Clustering of Number of Times Selected by Greedy Algorithm for Analytes Included in the Blood Analyte Classifier

<table>
<thead>
<tr>
<th>Real Data</th>
<th>CHR-P vs CHR-NP</th>
<th>CHR-P vs UC</th>
</tr>
</thead>
<tbody>
<tr>
<td>AUC (SE)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sum best 15 analytes</td>
<td>0.88 (0.043)</td>
<td>0.91 (0.036)</td>
</tr>
<tr>
<td>Sum best 9 analytes</td>
<td>0.86 (0.046)</td>
<td>0.87 (0.045)</td>
</tr>
<tr>
<td>Sum best 6 analytes</td>
<td>0.83 (0.051)</td>
<td>0.84 (0.048)</td>
</tr>
</tbody>
</table>

Fig. 2. Fifteen-analyte receiver operating curves and 95% confidence intervals: (A) for CHR-P vs UC and (B) CHR-P vs CHR-NP.

**Correlation of 15-Analyte Index With Symptoms**

Within the CHR subjects, the 15-analyte index was significantly correlated (Pearson) with baseline scores from the SOPS for total ($r = .26, P = .03$) and positive ($r = .23, P = .05$) scales, with trends for disorganized ($r = .22, P = .07$), general ($r = .19, P = .10$), and negative scales ($r = .19, P = .11$). Specific symptoms from the SOPS with strong correlations with the 15-analyte index included P1-unusual thought content/delusional ideas ($r = .28, P = .02$), D3-trouble with focus and attention, ($r = .25, P = .03$), G1-sleep disturbance ($r = .23, P = .05$), and G2-dysphoric (depressive, irritable, or anxious) moods ($r = .24, P = .04$). There was no relation between the 15-analyte index and severity of anxiety symptom as measured by the SAS ($r = .11, P = .36$) or depressive symptoms as measured by the CDSS ($r = .14, P = .25$).

**Confounder Analyses**

Baseline demographic variables (age, sex, ancestry, SES) and time of blood draw did not differ between groups and did not confound the relationship of the 15-analyte index with group status, nor did baseline measures of anxiety or depression or baseline (supplement S9). Given how common comorbid depression and anxiety disorders are in CHR persons,48 we also evaluated the potential confound of current diagnosis with a depression or anxiety disorder and found none (supplement S9).

There was no influence of cannabis, nicotine, or alcohol use on the strength of the relationship of the 15-analyte index to distinguish CHR-P vs CHR-NP or UC subjects, or antipsychotic, or antidepressant drug use with the strength of the relationship of the 15-analyte index in classifying CHR-P vs CHR-NP subjects. Only 3 UC subjects used marijuana or nicotine, only 1 UC subject was prescribed antidepressant, and none were prescribed antipsychotics, making evaluation of confounding of drug use and the 15-analyte index impossible to evaluate in CHR-P vs UC comparisons.

**Investigation of Analyte Functions**

A summary of functions of the 15 analytes included in the index is given in supplement S4 together with references. From the original panel of immune, hormonal, oxidative stress, and metabolism biomarkers, most of
the analytes included in the 15-analyte index are immunomodulatory: as cytokines (interleukin-1B, growth hormone, KIT ligand, interleukin-8, interleukin-7, resistin, chemokine [c-c motif] ligand 8) or as proteins involved in modulating inflammation including blood-brain barrier integrity (matrix metalloproteinase-7, immunoglobulin E, and coagulation factor VII). Three of the analytes (thyroid stimulating hormone, growth hormone, and cortisol) are part of hypothalamic-pituitary axes. Interestingly, 5 of the included cytokines (interleukin-1B, interleukin-7, interleukin-8, KIT ligand, and resistin) are known to regulate hypothalamic-pituitary axes. These findings indicate some dysregulation of immune-hypothalamic-pituitary interactions in the emergence of psychosis.

The most frequently chosen analyte, malondialdehyde-modified low-density lipoprotein, measures a lipoprotein damaged by free radicals and is thus a measure of oxidative stress. Elevation of apolipoprotein D, found at higher levels in CHR-P relative to CHR-NP and UC subjects, is also associated with oxidative stress. Oxidative stress and inflammation are intricately linked, and both are associated with schizophrenia. In summary, the above selection of 15 analytes may reflect interacting alterations in immune system function, hypothalamic-pituitary function, and oxidative stress.

**Discussion**

We report preliminary findings of a blood biomarker assay that, if confirmed, could improve determination of psychosis risk in persons experiencing the attenuated psychosis risk syndrome. Thus, given a new patient, the index value of a blood sample processed for 15 analytes as above could be compared with historic data to help evaluate likelihood of psychosis. For example, the ROC (figure 2B) suggests that if we accept a sensitivity of ~0.6, the index would have a specificity of ~0.9. As applied to persons meeting COPS criteria, who have a ~30% risk of psychosis in the next 2 years, this yields a PPV of 0.72, meaning that 72% of persons identified by the test as at high risk are truly at risk, and a negative predictive value (NPV) of 0.84, meaning that 84% of persons identified by the test as at low risk are truly at low risk of psychosis. Persons above this high-risk cut-off could be considered eligible for interventions where the benefits are high but the risks/costs may also be high, eg, antipsychotic medications or relatively intense individual/family therapies. Other cut-off scores could also have clinical utility. For example, using a cut-off score with a sensitivity of ~0.95 and thus a specificity of ~0.4, yields a PPV of 0.40 and a NPV of 0.95; persons scoring below the “high risk” cut-off score, but above the lower cut-off score, might appropriately receive brief counseling and more intense clinical monitoring. Thus, this index could provide a framework for segmenting the clinical high-risk population along a dimension of biological risk and then testing step-wise approaches to care.

There are several unique aspects regarding data analysis that warrant comment. Given that the analyte concentrations vary ~1 000 000-fold, we created a z score for each analyte, based on the average and SD of the UC subjects. The z score normalization allows the creation of an unweighted scale. We propose that the elimination of weighting minimizes the problem of overfitting sometimes encountered with more complex algorithms that use various weighting schemes. Also, we recognize that the relevant patient population will frequently be treated with various medications, especially antidepressants and antipsychotics or used nicotine, marijuana, and alcohol, and that such drugs may influence analyte levels. Thus, we elected to eliminate analytes that showed a possible relation to these substances and furthermore
verified that substance use did not confound relationships with the 15-analyte index. An alternative could be to develop different classifiers for persons depending on prescription medication treatment, but larger numbers of subjects would be required.

A major strength of this study is the prospective design. However, the parent project was not designed around plasma biomarker discovery, thus there are several methodological issues that could impact results, causing type I error if biased and underestimation of true difference (Type II error) if not biased. 52 Relation to meals, physical exertion, menstrual cycle, body weight, and presence of a cold or allergy were not considered or measured during data collection. While all samples were run in duplicate, limited resources allowed the analysis of only one occult technical replicate. Finally, with larger samples, normalization could be refined to include sex, age, and other factors that could impact analyte expression. 53

The results support our initial hypothesis that activation/dysregulation of the immune system may play a central role in the development of psychosis. Our findings are consistent with the emerging body of evidence linking inflammation and schizophrenia. 14, 15 Hypothalamic-pituitary axis dysregulation is also implicated by our results, although we found elevations of growth and thyroid hormones, in contrast to studies in schizophrenia subjects that report reductions. 53 In support of the involvement of hypothalamus and pituitary in the onset of psychosis, imaging studies find increases in hypothalamic 54 and pituitary 55 volumes in persons at elevated risk for psychosis compared with unaffected subjects. In addition, activation of the hypothalamic-pituitary-adrenal axis, as evidenced by elevated salivary cortisol, has also been reported in CHR who developed psychosis. 11, 56, 57

Finally, our finding that elevations in low-density lipoprotein damaged by oxidative stress predict psychosis is in concert with the substantial literature documenting oxidative damage in persons with schizophrenia. 10

A critical step is testing in other subjects at high risk for psychosis, as well as establishing the reproducibility of the assay. More work is needed, however, as there are likely many other combinations of analytes with utility in psychosis risk prediction, and a blood assay could be combined with other clinical, imaging, or electrophysiological measures associated with progression to psychosis in clinical high-risk subjects. 58 A better understanding of biomarkers predictive of psychosis could advance our ability to discriminate symptomatic persons truly at high risk for psychosis from those with minimally elevated risk. If confirmed, blood biomarkers have potential to influence strategies for preventative interventions, eg, targeting inflammation, stress reactivity, or oxidative stress directly. Our results highlight the potential promise to identify new targets for prevention for psychosis and support the need for further research in this area.

Supplementary Material
Supplementary material is available at http://schizophreniabiuletin.oxfordjournals.org.

Funding
National Institute of Mental Health grants (U01MH0818902 to T.D.C., U01MH081984 to J.M.A., P50MH066286 to C.E.B., U01MH082022 to K.S.C., U01MH081857 to B.A.C., U01MH082004 to D.O.P., U01MH081928 to L.J.S., U01MH081988 to E.F.W., U01MH066160 to S.W.W.). The analysis of plasma samples was funded by a gift from an anonymous donor administered through the San Francisco Foundation (Perkins).

Acknowledgments
We acknowledge and thank the study participants and their family members. We also thank Katie Lansing, Nicholas Walter, and Kees Frelinguer for their contributions toward the NAPLS plasma biobank. D.O.P., C.D.J., J.M.A., C.E.B., K.S.C., T.D.C., B.A.C., D.H.M., T.H.M., L.J.S., M.T.T., E.F.W., and S.W.W. are listed as inventors on a provisional patent application based in part on the work presented in this manuscript. R.H. has no financial disclosures or conflict of interests to declare.

References
Towards a Psychosis Risk Blood Diagnostic


