Prediction of Psychosis in Adolescents and Young Adults at High Risk

Results From the Prospective European Prediction of Psychosis Study

Stephan Ruhrmann, MD; Frauke Schultze-Lutter, PhD; Raimo K. R. Salokangas, MD, PhD, MSc; Markus Heinimaa, MD; Don Linszen, MD; Peter Dingemans, PhD; Max Birchwood, PhD; Paul Patterson, PhD; Georg Juckel, MD; Andreas Heinz, MD; Anthony Morrison, PhD; Shôn Lewis, MD; Heinrich Graf von Reventlow, DiplPsych; Joachim Klosterkötter, MD

Context: Indicated prevention is currently regarded as the most promising strategy to attenuate, delay, or even avert psychosis. Existing criteria need improvement in terms of specificity and individual risk assessment to allow for better targeted and earlier interventions.

Objective: To develop a differential predictive clinical model of transition to first-episode psychosis.

Design: Prospective multicenter, naturalistic field study with a total follow-up time of 18 months.

Setting: Six early-detection outpatient centers in Germany, Finland, the Netherlands, and England.

Participants: Two hundred forty-five help-seeking patients in a putatively prodromal state of psychosis according to either ultra-high-risk (UHR) criteria or the basic symptom–based criterion cognitive disturbances (COGDIS).

Main Outcome Measure: Incidence of transition to psychosis.

Results: At 18-month follow-up, the incidence rate for transition to psychosis was 19%. Combining UHR and COGDIS yielded the best sensitivity. A prediction model was developed and included positive symptoms, bizarre thinking, sleep disturbances, a schizotypal disorder, level of functioning in the past year, and years of education. With a positive likelihood ratio of 19.9, an area under the curve of 80.8%, and a positive predictive value of 83.3%, diagnostic accuracy was excellent. A 4-level prognostic index further classifying the general risk of the whole sample predicted instantaneous incidence rates of up to 85% and allowed for an estimation of time to transition.

Conclusions: The prediction model identified an increased risk of psychosis with appropriate prognostic accuracy in our sample. A 2-step risk assessment is proposed, with UHR and cognitive disturbance criteria serving as first-step criteria for general risk and the prognostic index as a second-step tool for further risk classification of each patient. This strategy will allow clinicians to target preventive measures and will support efforts to unveil the biological and environmental mechanisms underlying progression to psychosis.

Arch Gen Psychiatry. 2010;67(3):241-251

Prevention is most important in psychiatry today. An indicated approach—targeting intervention—when minimal though detectable signs or symptoms are already present but diagnostic criteria are not met is currently considered the most promising strategy to attenuate, delay, or prevent psychosis. The European Prediction of Psychosis Study (EPOS), a unique prospective, multicenter, naturalistic field study, was designed to develop a prediction model of psychosis within a period of 18 months based on a sufficiently large cohort of potentially prodromal subjects. Two concepts of early detection of help-seeking patients at short-term risk of psychosis are the focus of current research: the ultra-high-risk (UHR) criteria and the basic symptom approach. While retrospective studies have confirmed an average prodromal period of 5 to 6 years, the introduction of the UHR criteria has significantly advanced the possibility of indicated prevention of psychosis. Ultra-high-risk criteria include 3 different syndromes: attenuated positive symptoms, brief limited intermittent psychotic symptoms (BLIPS), or a combination of genetic risk indicators and recent functional deterioration. Supporting their conceptualization as criteria of an imminent risk of psychosis, initial studies reported transition rates to psychosis of 35% to 54% within 12 months. However, recent
studies indicate that the UHR criteria may actually capture a wider variance of individual risk in terms of magnitude and time.16-21

Basic symptoms are subtle, subjective, subclinical disturbances in several mental domains not recognized in measurements of positive symptoms.9,22 In the prospective Cologne Early Recognition Study, 78% of patients with at least 1 cognitive-perceptive basic symptom at baseline developed schizophrenia within a mean follow-up of 9.6 years.6 Further analyses resulted in a second basic symptom criterion based on 9 cognitive disturbances (COGDIS) that was associated with a transition rate to psychosis of 23.9% at 12 months and 46.3% at 24 months.23 Therefore, we predicted that the COGDIS criterion when combined with the UHR criteria would improve the prediction of transition to psychosis. Against this background, EPOS aims to develop a clinical model that will maximize the prediction of psychosis with sufficient sensitivity and specificity and enable a differentiation of individual risk in terms of magnitude and time, thus providing risk-adapted inclusion criteria for future randomized trials of early detection and intervention.24

METHODS

RECRUITMENT

Local ethics committees of the participating universities or health care agencies approved the study; patients were recruited from consecutive referrals to early-detection services at the 6 participating centers.5 Written informed consent was obtained from all participants and their parents if they were minors. Referral sources included psychiatrists, psychologists, general practitioners, outreach clinics, counseling services, and teachers; patients also initiated contact. Knowledge about early warning signs (eg, concentration and attention disturbances, unexplained functional decline) and inclusion criteria was disseminated (through local workshops, articles in professional journals and newsletters, informational flyers, and Web sites) to mental health professionals as well as institutions and persons who might be contacted by at-risk persons seeking help.

Inclusion criteria were composed of COGDIS, assessed by the Bonn Scale for the Assessment of Basic Symptoms—Prediction List (BSABS-P),23 and UHR criteria, assessed by the Structured Interview for Prodromal Syndromes, version 3.0 (SIPS).26 The COGDIS criterion requires the presence of at least 2 of 9 cognitive basic symptoms of at least moderate severity (minimum score of ≥3) during the last 3 months and, independent of severity, a first occurrence at least 1 year before intake. The basic symptoms include inability to divide attention; thought interference, pressure, and blockage; and disturbances of receptive and expressive speech, disturbance of abstract thinking, unstable ideas of reference, and captivation of attention by details of the visual field.

The UHR approach consists of 3 alternative criteria. Attenuated positive symptoms were defined by at least 1 of the following symptoms with a moderate to severe but not psychotic (3-5) SIPS score appearing several times per week for at least 1 week within the last 3 months: unusual thought content/delusional ideas, suspiciousness/persecutory ideas, grandiosity, perceptual abnormalities/hallucinations, disorganized communication, and odd behavior/appearance. Brief limited intermittent psychotic symptoms were defined by hallucinations, delusions, or formal thought disorders that occurred within the last 3 months, resolved spontaneously within 1 week, scored as severe and psychotic (score of 6) on the SIPS, and scored as at least moderate on the Positive and Negative Syndrome Scale for Schizophrenia.27 Genetic risk and functional deterioration were defined by (1) a 30% or greater reduction in score on the modified version of the Global Assessment of Functioning Scale (GAF-M)26,28,29 for at least 1 month within the previous year compared with the highest level of previous functioning and (2) having a first- or second-degree relative with a history of any DSM-IV30 psychotic disorder or having a DSM-IV schizotypal personality disorder. The age range was 16 to 35 years.

Exclusion criteria were having had a psychotic episode for more than 1 week, ie, fulfilling DSM-IV criteria of a brief psychotic episode not only for more than 1 day but for more than 7 days, assessed with the Structured Clinical Interview for DSM-IV; having symptoms relevant for inclusion arising from a known general medical disorder or drugs or alcohol dependency30; and having a low verbal IQ (<85).

Between August 2002 and April 2006, 661 persons were screened. eTable 1 (available at http://www.archgenpsychiatry.com) shows the reasons for nonparticipation. Of 513 patients who fulfilled inclusion criteria, 245 consented to participate (Table 1 and Table 2).

ASSSESSMENTS

Present analyses use data assessed by SIPS, including GAF-M, BSABS-P, the Beck Depression Inventory,31 the Composite International Diagnostic Interview—Sections on Alcohol and Drugs,32 and for sociodemographic data, the EPOS Basic Data Form.33 At follow-ups, any new medical condition or prescription was assessed.

The 4 SIPS subscales (Positive, Negative, Disorganization, and General Symptoms) include 4 to 6 items (in total 19) rated on a 7-point severity scale (score, 0-6). The EPOS investigators received extensive training by one of the scale’s authors (Tandy J. Miller, PhD). Pairwise interrater concordance for SIPS was 77%, which was determined acceptable by the training team. The BSABS-P, an abbreviated item list of the Schizophrenia Proneness Instrument,34 includes 3 subscales, totaling 33 cognitive, perceptual, and motor disturbances assessed on a 7-point severity scale (score, 0-6), with maximum frequency of occurrence during the preceding 3 months as the guiding criterion. Every item corresponds to a single symptom; the BSABS-P differs from structure in the SIPS, in which items are mostly defined by multiple SIPS items. The EPOS investigators received repeated training by one of the scale’s authors (Frauke Schultz-Lutter, PhD). The concordance rate with expert rating (F. S-L.) was 87.9%. Past or current psychosis as part of the exclusion criteria as well as psychotic diagnosis was assessed by the Structured Clinical Interview for DSM-IV.35

FOLLOW-UP

Follow-up assessments took place 9 and 18 months from baseline. Transition to psychosis was operationalized as a continuation of BLIPS, ie, any single item on the Positive subscale of SIPS (SIPS-Positive) with a score of 6 for more than 7 days,8,10,21 Following identification of full-blown psychotic symptoms in SIPS, the diagnostic category of transition was determined by applying DSM-IV criteria for psychotic disorders and affective disorders with psychotic features. The different time threshold of criterion B of a brief psychotic disorder was adapted to the definition of BLIPS.

STATISTICAL ANALYSIS

To determine the risk of transition, we used Kaplan-Meier survival analysis for calculating the cumulative hazard rate. This
rate measures the incidence rate at exactly time t and is thus called the instantaneous incidence rate (iIR). Subjects with survival times exceeding the 18-month follow-up were considered censored at the end of month 18. Survival curves were compared using the log-rank test.

The effect of covariates on survival time, ie, time to transition, was estimated with the Cox proportional hazard model. Continuous data were entered into analyses as raw as well as categorized data (eTable 2). Item scores were dichotomized according to the cutoff values of the inclusion criteria of the respective scale, generally at a score of more than 2. Summary scores were dichotomized at their respective cutoffs, which were determined by explorative receiver operating characteristic curve analyses to combine a high specificity with a sensitivity of 0.25 or greater.

Predictors were selected in several steps. First, covariates were computed individually and chosen for further analyses when changes of the –2 log-likelihood of the model and the Wald statistic became significant (P < .10). Next, backward multivariate Cox regression analyses were performed within each domain (eTable 2) at a liberal level of significance (P < .15). We entered retained covariates into a multivariate backward regression (P < .05) across domains, introducing domains block-wise. For the resulting covariates, interactions were calculated and kept in the model if they were significant (P < .05). Finally, the remaining covariates were analyzed forward and backward to exclude effects of blocking. The center that participants attended was entered as a strata variable in all Cox regressions.

Applying the final Cox model to each case, we generated individual prognostic scores, allowing a 4-level prognostic index (PI) to be developed. Continuous variables had to be centered around the mean for this calculation.

In addition to Cox regression–derived PIs and hazard ratios (HRs) serving as comparative measures of survival during the 18-month period, more common, end point–related predictive accuracy measures—sensitivity, specificity, positive and negative likelihood ratios, and positive and negative predictive values—were calculated for the Cox model by entering the individual prognostic scores into a binary logistic regression. Areas under the receiver operating characteristic curve were estimated for the predicted logistic probabilities as cutoff-independent measures of the ability to discriminate between transition and nontransition. For group comparisons not including survival times, we used Mann-Whitney, χ², and Fisher exact tests. The impact of psychopathology on GAF-M scores was explored through linear regression analysis. Normal distribution was ascertained by the Kolmogorov-Smirnov test. Unless indicated otherwise, a 2-sided α less than .05 was considered significant. Bonferroni correction was applied to adjust for multiple testing. We used SPSS, version 15.0.1.1, and OpenEpi, version 2.2.41.
Table 2. Clinical Sample Characteristics

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Total Sample (N = 245)</th>
<th>Known Outcomea (n = 183)</th>
<th>Censored Before End of Follow-upb (n = 62)</th>
<th>U/χ²</th>
<th>P Valuea</th>
</tr>
</thead>
<tbody>
<tr>
<td>SIPS score, mean (SD)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>34.4 (14.2)</td>
<td>35.1 (14.4)</td>
<td>32.3 (13.6)</td>
<td>5096.0</td>
<td>.26</td>
</tr>
<tr>
<td>Positive</td>
<td>9.6 (4.4)</td>
<td>9.7 (4.4)</td>
<td>9.4 (4.5)</td>
<td>5419.0</td>
<td>.60</td>
</tr>
<tr>
<td>Negative</td>
<td>11.7 (6.6)</td>
<td>12.1 (6.6)</td>
<td>10.5 (6.6)</td>
<td>4924.0</td>
<td>.13</td>
</tr>
<tr>
<td>Disorganization</td>
<td>4.6 (3.1)</td>
<td>4.7 (3.2)</td>
<td>4.3 (2.8)</td>
<td>5356.0</td>
<td>.55</td>
</tr>
<tr>
<td>General</td>
<td>8.5 (4.1)</td>
<td>8.6 (4.2)</td>
<td>8.1 (4.0)</td>
<td>5206.5</td>
<td>.36</td>
</tr>
<tr>
<td>BSABS-P score, mean (SD)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>22.8 (15.5)</td>
<td>22.4 (15.6)</td>
<td>24.1 (15.3)</td>
<td>5240.5</td>
<td>.37</td>
</tr>
<tr>
<td>Cognitive thought disturbances</td>
<td>13.9 (8.9)</td>
<td>13.8 (9.9)</td>
<td>13.9 (8.9)</td>
<td>5604.0</td>
<td>.89</td>
</tr>
<tr>
<td>Cognitive perception disturbances</td>
<td>8.0 (7.8)</td>
<td>7.6 (7.9)</td>
<td>9.4 (7.5)</td>
<td>4580.5</td>
<td>.02</td>
</tr>
<tr>
<td>Cognitive motor disturbances</td>
<td>1.0 (2.0)</td>
<td>1.0 (2.1)</td>
<td>0.8 (1.9)</td>
<td>5377.0</td>
<td>.42</td>
</tr>
<tr>
<td>GAF-M score, mean (SD)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>At inclusion</td>
<td>51.1 (11.7)</td>
<td>50.3 (11.4)</td>
<td>53.4 (12.4)</td>
<td>4764.0</td>
<td>.06</td>
</tr>
<tr>
<td>Highest score in the year before inclusion</td>
<td>65.4 (14.0)</td>
<td>64.0 (14.2)</td>
<td>69.3 (12.1)</td>
<td>4394.0</td>
<td>.01</td>
</tr>
<tr>
<td>% of reduction in GAF-M score, mean (SD)</td>
<td>20.4 (15.7)</td>
<td>19.9 (15.0)</td>
<td>21.9 (17.6)</td>
<td>7785.5</td>
<td>.54</td>
</tr>
<tr>
<td>Persons with a GAF-M score reduction ≥30%, No. (%)</td>
<td>70 (28.6)</td>
<td>52 (28.4)</td>
<td>18 (29.0)</td>
<td>5368.0</td>
<td>.57</td>
</tr>
<tr>
<td>BDI score, mean (SD)</td>
<td>20.3 (10.9)</td>
<td>20.2 (10.4)</td>
<td>20.6 (12.3)</td>
<td>5187.0</td>
<td>.94</td>
</tr>
<tr>
<td>First- or second-degree relative with psychotic disorder, No. (%)</td>
<td>57 (23.3)</td>
<td>47 (25.7)</td>
<td>10 (16.1)</td>
<td>2.4</td>
<td>.12</td>
</tr>
<tr>
<td>First-degree relative with psychotic disorder, No. (%)</td>
<td>40 (16.3)</td>
<td>32 (17.5)</td>
<td>8 (12.9)</td>
<td>0.7</td>
<td>.39</td>
</tr>
<tr>
<td>Only second-degree relative with psychotic disorder, No. (%)</td>
<td>16 (6.5)</td>
<td>14 (7.7)</td>
<td>2 (3.2)</td>
<td>1.5</td>
<td>.22</td>
</tr>
<tr>
<td>Schizotypal personality disorder, No. (%)</td>
<td>33 (13.5)</td>
<td>28 (15.3)</td>
<td>5 (8.1)</td>
<td>2.1</td>
<td>.15</td>
</tr>
<tr>
<td>Obstetric complications, No. (%)</td>
<td>39 (16.3)</td>
<td>28 (15.7)</td>
<td>11 (17.7)</td>
<td>0.1</td>
<td>.71</td>
</tr>
<tr>
<td>Any drug abuse, No. (%)</td>
<td>53 (23.0)</td>
<td>35 (20.1)</td>
<td>18 (32.1)</td>
<td>3.6</td>
<td>.06</td>
</tr>
<tr>
<td>Alcohol abuse, No. (%)</td>
<td>69 (30.0)</td>
<td>51 (29.7)</td>
<td>18 (31.0)</td>
<td>0.04</td>
<td>.84</td>
</tr>
<tr>
<td>Distribution of inclusion criteria, No. (%)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>UHR and COGDIS positive</td>
<td>146 (59.6)</td>
<td>106 (57.9)</td>
<td>40 (64.5)</td>
<td>0.8</td>
<td>.36</td>
</tr>
<tr>
<td>UHR positive and COGDIS negative</td>
<td>74 (30.2)</td>
<td>59 (32.2)</td>
<td>15 (24.2)</td>
<td>1.4</td>
<td>.23</td>
</tr>
<tr>
<td>COGDIS positive and UHR negative</td>
<td>25 (10.2)</td>
<td>18 (9.8)</td>
<td>7 (11.3)</td>
<td>0.1</td>
<td>.74</td>
</tr>
</tbody>
</table>

Abbreviations: BDI, Beck Depression Inventory; BSABS-P, Bonn Scale for the Assessment of Basic Symptoms—Prediction List; COGDIS, basic symptom criterion cognitive disturbances; GAF-M, Global Assessment of Functioning Scale, modified version; SIPS, Structured Interview for Prodromal Syndromes; UHR, ultra-high-risk criteria.

aAll persons transitioning to psychosis during 18-month follow-up or completing it.

bPersons with observation times shorter than 18 months and unknown outcomes.

cContinuous variables were compared using the Mann-Whitney U test; categorical variables were compared using χ² or Fisher exact test.

dSignificant at a Bonferroni-adjusted α level < .0016.

eCalculated as [highest score in the year before inclusion – score at inclusion]/highest score in the year before inclusion] × 100.

fOnly 2 patients had first- and second-degree relatives with psychosis.

gAccording to the SIPS definition: DSM-IV criteria for schizotypal personality disorder fulfilled for more than 1 year.

hNo information in 5 cases (known outcome sample).

iAccording to the Composite International Diagnostic Interview. Data about drug abuse were only available in 174 subjects with known outcomes and 56 subjects censored before the end of follow-up; data about alcohol abuse were available in 172 and 58 subjects, respectively.

jAt least 1 UHR criterion had to be fulfilled to be considered positive.

**EFFECTS OF INCLUSION CRITERIA**

Distribution of UHR and COGDIS is displayed in Table 2. The mean survival time for participants at risk of psychosis according to COGDIS criteria and negative according to UHR criteria was 528.6 days (SE, 19.0 days; 95% CI, 491.4-565.8; iIR, 4.5%); for participants positive according to UHR and negative according to COGDIS, 500.7 days (SE, 15.0 days; 95% CI, 471.3-530.2; iIR, 18.0%); and for participants positive according to both UHR and COGDIS, 489.1 days (SE, 11.6 days; 95% CI, 466.7-512.1; iIR, 21.9%). Differences were nonsignificant (log-rank test; χ² = 2.82, P = .24). No single UHR criterion had a significant effect on time to transition. When both UHR and COGDIS criteria were considered irrespective of each other (as done in all earlier studies), subjects reporting UHR criteria had a 20.6% transition rate; those reporting COGDIS had a rate of 19.1%.
EFFECTS OF COVARIATES ON SURVIVAL TIME

Six variables were included in the final Cox regression model (Table 3). The resulting equation was $[1.571 \times \text{SIPS–Positive score} >16] + [0.865 \times \text{SIPS bizarre thinking score}] + [0.793 \times \text{SIPS sleep disturbance score}] + [1.037 \times \text{SIPS Schizotypal personality disorder score}] + [0.933 \times \text{highest GAF-M score in the past year} - 34.64]) + [0.250 \times \text{years of education} - 12.52]$. The 4.81 HR emerging for the variable SIPS–Positive score greater than 16 was the highest, indicating that patients with a total subscale score above 16 transitioned to psychosis at a rate 4.81 times higher than subjects with a lower score. In the univariate analyses, GAF-M score and years of education showed the best predictive performance as continuous variables. For consistency with the scaling of the other predictors, they had to be inverted (Table 3). An inversion of the highest GAF-M score in the past year of, eg, 70 resulted in an HR of 3.21. Compared with the mean of 34.64 (HR, 1.03; 34.64 equals a converted (Table 3). An inversion of the highest GAF-M score (highest GAF-M score in the past year $- 34.64])$. Comparing a rank of 16, equaling 9 years of education, with the mean rank of 12.52 (HR, 1.28), an HR of 2.39 emerged.

The impact of psychopathology on GAF-M score at inclusion was estimated, with Beck Depression Inventory total score and SIPS and BSABS-P subscale scores as independent variables. Stepwise linear regression retained SIPS–Positive and SIPS–Negative scores, explaining 14.9% of variance and thus indicating that GAF-M scores were predominantly determined by deterioration of role functioning.

PROGNOSTIC INDEX

The final equation resulting from Cox regression procedures was applied to each patient to calculate individual prognostic scores. As these scores followed a normal distribution, mean and first positive and negative SD served to stratify the scores into 4 classes, thus establishing a prognostic index (PI) for risk classification (Table 4 and Figure 2). Figure 3 displays the corresponding HR curves. The log-rank test is calculated across all classes and became significant because of class IV ($X^2 = 58.68, P<.001$); in post hoc tests, class IV differed significantly from the other classes.

EFFECTS OF ANTIPSYCHOTIC OR ANTIDEPRESSANT MEDICATION

Pharmacological treatment was added to the prediction model to estimate its additional effect during follow-up. Antidepressants were prescribed to 46 patients (18.8%), antipsychotics to 32 (13.1%), and both were prescribed to 22 (9.0%); no reliable information was available for 31 (12.7%) patients.

Adding treatment to the model, only antipsychotics were kept in the equation ($\beta = 1.261; SE, 0.456; P = .006$). All 6 variables of the initial predictor set (Table 3) continued to contribute significantly to the equation, and changes in HRs were minor and unsystematically occurred in either direction at decimal places only.

To further evaluate this result, GAF-M, Beck Depression Inventory, SIPS, and BSABS-P baseline total and subscale scores were analyzed with regard to antipsychotic prescription. Patients prescribed antipsychotic drugs displayed significantly poorer GAF scores at inclusion ($U = 1916.5, P = .01$) and SIPS-Positive scores ($U = 1948.5, P = .02$). Significant results did not survive Bonferroni adjustment for multiple testing (11 analyses; $P<.0045$, adjusted).

PROGNOSTIC ACCURACY MEASURES

As prognostic accuracy measures cannot be analyzed with censored data, calculations had to rely on the subsample with a known state of transition (n=183 [74.7%]). After adjustment for multiple testing, this subsample displayed no significant difference from the subsample censored earlier, ie, the dropouts, with regard to demographic or clinical variables (Table 1 and Table 2) or even prognostic scores ($U = 5054.5, P = .62$).

The prognostic accuracy measures for the prediction model and, for comparison, for inclusion criteria are given in Table 5. Neither patients who were positive according to UHR and negative according to COGDIS criteria for an at-risk state of psychosis nor patients who were positive according to COGDIS and negative according to UHR criteria produced an area under the receiver operating characteristic curve above the threshold for an acceptable discrimination or an adequate (ie, very positive37 positive likelihood ratio; the positive predictive value was low. Only patients positive for both UHR and

Table 3. Cox Proportional Hazard Model

<table>
<thead>
<tr>
<th>Predictor Variable</th>
<th>$\beta$</th>
<th>SE</th>
<th>Wald,</th>
<th>HR (95% CI)</th>
<th>$P$ Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>SIPS Positive subscale score $&gt;16$</td>
<td>1.571</td>
<td>0.428</td>
<td>13.46</td>
<td>4.81 (2.078-11.134)</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Bizarre thinking score $&gt;2$ on SIPS</td>
<td>0.865</td>
<td>0.387</td>
<td>4.99</td>
<td>2.38 (1.112-5.074)</td>
<td>.03</td>
</tr>
<tr>
<td>Sleep disturbances score $&gt;2$ on SIPS</td>
<td>0.793</td>
<td>0.387</td>
<td>4.19</td>
<td>2.21 (1.034-4.717)</td>
<td>.04</td>
</tr>
<tr>
<td>Schizotypal personality disorder according to SIPS</td>
<td>1.037</td>
<td>0.423</td>
<td>6.01</td>
<td>2.82 (1.231-6.464)</td>
<td>.01</td>
</tr>
<tr>
<td>GAF-M score, highest in the past year</td>
<td>0.033</td>
<td>0.015</td>
<td>5.02</td>
<td>1.03 (1.004-1.064)</td>
<td>.03</td>
</tr>
<tr>
<td>Years of education, including university</td>
<td>0.250</td>
<td>0.086</td>
<td>8.35</td>
<td>1.28 (1.084-1.521)</td>
<td>.004</td>
</tr>
</tbody>
</table>

Abbreviations: CI, confidence interval; GAF-M, Global Assessment of Functioning Scale, modified version; HR, hazard ratio; SIPS, Structured Interview for Prodromal Syndromes.

aTo keep all $\beta$ coefficients positive, GAF-M scores have been inverted by subtracting 100 (higher score means lower level, eg, an original score of 30 equals 70) and years of education have been inverted by rank transformation in a scale from 1 to 17 (higher number means fewer years of education, eg, 17 = 8 years of education, 16 = 9 years of education). Only algebraic signs were changed owing to this procedure.
In this uniquely large European sample of help-seeking individuals putatively at risk of psychosis, we observed iIRs of transitions to psychosis of 14% after 12 months and 19% after 18 months. Relating this incidence for illustrative reasons to the recently reported 12-month incidence rate in a UK general population sample of 0.0348% for any psychosis,^{42} we found that the 18-month relative risk in EPOS was 364. With regard to other early detection studies using the UHR paradigm^{16-21} or similar approaches,^{43} the extent and course of EPOS transition rates are consistent with recent reports of lower 12-month transition rates^{17-21} and of a further progression of rates beyond 12 months.^{6,8,17,18,21,44} Predicted rates, however, are not yet satisfactory, thus calling for further improvement of risk assessment.

In somatic medicine such as oncology or pneumology,^{45-47} a well-established and widespread risk-modeling procedure is the use of PIs for a multivariate clinical staging. In EPOS, this approach was adopted for the first time in early-detection research. The EPOS prediction model, ie, the Cox regression equation, included 6 variables. Based on the resulting prognostic scores, a multivariate PI (EPOS-PI) for further classifying the risk of psychosis in selected risk populations into 4 classes was suggested. Thereby, the EPOS-PI is not intended to replace the inclusion criteria, but to facilitate better, more individualized risk estimation and PI-related future targeting of interventions in predefined clinical at-risk samples.

Table 4. Classification of Predefined Risk

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Risk Class of the Prognostic Index</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>I</td>
</tr>
<tr>
<td>Prognostic score&lt;sup&gt;a&lt;/sup&gt;</td>
<td>&lt;−0.50</td>
</tr>
<tr>
<td>No. of patients (%)&lt;sup&gt;b&lt;/sup&gt;</td>
<td>32 (13.5)</td>
</tr>
<tr>
<td>Estimated time to transition, mean (SEM), d</td>
<td>597.1 (10.7)&lt;sup&gt;c&lt;/sup&gt;</td>
</tr>
<tr>
<td>iIR of transition to psychosis, %</td>
<td>3.5</td>
</tr>
<tr>
<td>At month 9</td>
<td>3.5</td>
</tr>
<tr>
<td>At month 12</td>
<td>3.5</td>
</tr>
<tr>
<td>Distribution of inclusion criteria, %</td>
<td>COGDIS 18.8</td>
</tr>
<tr>
<td></td>
<td>UHR 25.0</td>
</tr>
<tr>
<td></td>
<td>Both 56.3</td>
</tr>
</tbody>
</table>

Abbreviations: COGDIS, basic symptom criterion cognitive disturbances; GAF-M, Global Assessment of Functioning Scale, modified version; iIR, instantaneous incidence rate; SIPS, Structured Interview for Prodromal Syndromes; SPD, schizotypal personality disorder; UHR, ultra-high-risk criteria.

<sup>a</sup>The prognostic score is calculated as $[1.571 \times \text{SIPS-Positive score} + 16] + [0.865 \times \text{bizarre thinking score}] + [0.793 \times \text{sleep disturbances score}] + [1.037 \times \text{SPD score}] + [0.033 \times (\text{highest GAF-M score in the past year} – 34.64)] + [0.250 \times (\text{years of education} – 12.52)]$. The highest GAF-M score in the past year and years of education are continuous variables (see footnote “a” in Table 3 for details); all others are dichotomous variables.

<sup>b</sup>Of 238 subjects.

<sup>c</sup>Ninety-five percent confidence interval: 516.1-558.1.

<sup>d</sup>Ninety-five percent confidence interval: 501.0-541.2.

<sup>e</sup>Ninety-five percent confidence interval: 492.3-536.9.

<sup>f</sup>Ninety-five percent confidence interval: 287.3-425.5.

<sup>g</sup>The relative risk at 18 months, referring illustratively to a 0.0348% 12-month incidence of psychosis in the general population,^{36} is 67 for class I, 153 for class II, 352 for class III, and 1630 for class IV.
Table 5. Prognostic Accuracy in the Group With Known Outcomes at Month 18 (n=183)

<table>
<thead>
<tr>
<th>Variable</th>
<th>No. of Subjects</th>
<th>Sensitivity</th>
<th>Specificity</th>
<th>AUC (SE), 95% CI</th>
<th>P Value</th>
<th>PPV</th>
<th>NPV</th>
<th>LR+</th>
<th>LR-</th>
</tr>
</thead>
<tbody>
<tr>
<td>EPOS prediction modelc</td>
<td>179d</td>
<td>41.7</td>
<td>97.9</td>
<td>80.8 (4.6), 71.7-89.9</td>
<td>&lt;.001</td>
<td>83.3</td>
<td>87.0</td>
<td>19.9</td>
<td>0.6</td>
</tr>
<tr>
<td>UHR positive and COGDIS negativea</td>
<td>183</td>
<td>29.7</td>
<td>67.1</td>
<td>51.6 (5.3), 41.2-61.9</td>
<td>.62</td>
<td>18.6</td>
<td>79.0</td>
<td>0.9</td>
<td>1.1</td>
</tr>
<tr>
<td>COGDIS positive and UHR negative</td>
<td>183</td>
<td>2.7</td>
<td>88.4</td>
<td>54.4 (5.1), 44.6-64.4</td>
<td>.40</td>
<td>5.6</td>
<td>78.2</td>
<td>0.2</td>
<td>1.0</td>
</tr>
<tr>
<td>UHR and COGDIS positiveb</td>
<td>183</td>
<td>67.6</td>
<td>45.6</td>
<td>56.4 (5.2), 46.2-66.6</td>
<td>.23</td>
<td>23.8</td>
<td>84.8</td>
<td>1.2</td>
<td>0.7</td>
</tr>
</tbody>
</table>

Abbreviations: AUC, area under the curve; CI, confidence interval; COGDIS, basic symptom criterion cognitive disturbances; EPOS, European Prediction of Psychosis Study; LR+, positive likelihood ratio; LR-, negative likelihood ratio; NPV, negative predictive value; PPV, positive predictive value; UHR, ultra-high-risk.

a Null hypothesis: AUC = 50.0; AUC and its P values are derived from receiver operating characteristic curve analysis. Predictive discrimination is considered acceptable if 0.7 ≤ AUC < 0.8 and excellent if 0.8 ≤ AUC < 0.9.4

b An LR of 3 or more is considered moderately positive, and 10 or more, very positive; an LR of 0.3 or less is considered moderately negative; 0.1 or less, extremely negative.42

c In the subsample of only medication-naive patients, the regression analysis produced the following prognostic accuracy values: sensitivity = 40.0; specificity = 98.6; PPV = 80.0; NPV = 92.2; LR+ = 28.8; LR- = 0.61; AUC (SE) = 84.4 (7.4); 95% CI, 69.9 to 99.0; P > .001.

d Lower population number owing to missing values in single variables.

Previous approaches using regression models and single cutoffs to improve prediction by UHR criteria resulted in an unfavorable loss of sensitivity and, consequently, in an undesirable exclusion of patients below the cutoff who might benefit from early intervention.5,10,21 In this respect, the multilevel PI approach that also relies on a regression model but results from applying the regression equation to each subject and enables a risk classification of the whole sample has a major advantage: no matter his or her prognostic score, a patient continues to be generally considered at risk once screened positive for inclusion criteria. Thus, no loss of sensitivity occurs with this approach. Instead of a general yet, with regard to a help-seeking individual, rather undifferentiated risk estimate (as given by inclusion criteria), however, a patient’s PI provides further classification or staging of the current risk. Once such a PI has been validated in an independent sample, shown to be sensitive to dynamic change in risk over time, and related to different treatment strategies varying, eg, from monitoring and psychotherapy to low-dose antipsychotic medication, it could be re-assessed in short intervals to enable adaptation of intervention to the patient’s actual needs.

Such a strategy, however, is clinically sensible only if the risk differs meaningfully between classes. This was true for the 4 EPOS-PI classes, in which the iIR and the related relative risk increased by more than 100% with each higher level, though distributions of transition rates over time differed statistically only between class IV and any other class but not between the first 3 classes. The iIR of class IV in particular was much higher than any iIR of classes I through III and, with a completely distinct 95% CI, was associated with a significantly shorter time to transition. Compared with the overall 18-month iIR of 19% based on the inclusion criteria alone, the iIR was differentiated from 3.5% to 85% after introduction of the EPOS-PI classes.

The advantage of additionally introducing the PI as the second step in a 2-step procedure is supported by the diagnostic accuracy measures. The logistic model based on the prognostic scores underlying the 4 PI classes had excellent general threshold-independent discriminative accuracy and good to excellent classic diagnostic accuracy measures in addition to a high positive likelihood ratio. Thus, this is the first predictive model of psychosis possessing a positive likelihood ratio of sufficient size, ie, above 10,19 which indicates a clear change from pretest to posttest probability of nearly 20 times, raising the odds for psychosis. Despite a high negative predictive value, a prevalence-dependent measure,48 the model’s negative likelihood ratio indicated that it could not effectively rule out an increased risk of psychosis in patients with a negative test result. The dropout rate, a common limitation of follow-up studies, was 25.6% at 18 months. Because the group lost to follow-up did not differ from the one with known outcomes with regard to the prognostic score nor demographical or clinical variables, dropouts seem to have occurred at random; thus an additional bias was probably not introduced into our data.

With regard to the inclusion criteria, ie, the first step preceding the model’s application, 10.2% of the sample was included by COGDIS alone, 30.2% by UHR criteria alone (Table 3). Because EPOS allowed separating COGDIS and UHR criteria, it revealed that both criteria produced transition rates below the expected level when analyzed with adjustment for one another. Yet, even the observed rate of 4.5% in the COGDIS group still corresponds to a relative risk of 84.9.42 As shown recently in a study that reported equally UHR-adjusted COGDIS transition rates, higher transition rates for this criterion alone can be expected with longer follow-up: a 21.1% transition rate was observed in up to 37 months.23 In both studies, however, the cooccurrence of UHR criteria and COGDIS at baseline was responsible for the decidedly highest proportion of inclusions and exhibited the highest transition rate.

A stepwise procedure, using UHR and COGDIS criteria for risk screening first and the EPOS prediction model for further risk classification second, will be a superior strategy. In line with current indicated approaches to risk detection and prevention,6,17-19,21,43 the suggested procedure will apply only to those already seeking help; in the general population, prognostic accuracy would substantially drop for epidemiological reasons.40 A decrease in accuracy would also result, if the first step (screening by the inclusion criteria) were skipped. In an unselected sample of subjects seeking help, exclusive use of the PI...
could result in labeling persons who are characterized only by sleep disturbances, a decrease in psychosocial functioning, and a lower educational level as at risk. Yet, in terms of psychosis prediction, these rather unspecific characteristics become meaningful only in the presence of any of the inclusion criteria (eg, class I in Table 4).

As to the predictors themselves, all 6 variables included in the EPOS prediction model (SIPS-Positive score, bizarre thinking, sleep disturbances, schizotypal personality disorder [according to SIPS], highest GAF-M score in the past year, and years of education), have already been linked to psychosis and thus are clinically plausible.

The association between a higher SIPS-Positive score, ie, a higher degree of psychosis-like experiences, and an increased risk of psychosis is in line with most recent studies. Our finding also partially corresponds to the results of the North American Prodrome Longitudinal Study (NAPLS), as 2 positive SIPS items, unusual thought content and suspiciousness/paranoia, were among our 5 predictors.

The SIPS item bizarre thinking is characterized by strange, fantastic, or bizarre ideas that are distorted, illogical, or patently absurd. In a first-episode sample, 53.3% of the patients with schizophrenia exhibited odd and bizarre ideas. This symptom has been conceptualized as a formal thought disorder that is associated with cognitive disturbances in schizophrenia, especially verbal memory, which in turn has been reported to predict psychosis.

Sleep disturbances—common in psychiatric disorders—were associated with more severe psychotic states, electrophysiological disturbances, and dopaminergic as well as GABAergic dysregulations in psychosis. In retrospective studies on first-episode schizophrenia, sleep disturbances were a most common prodromal symptom and preceded first-episode psychosis in 77% to 100% of patients. Thus, albeit nonspecific, their inclusion in the EPOS model is well supported.

The inclusion of SIPS schizotypal personality disorder (minimum symptom duration of 1 year) as a predictor is in line with studies on schizotypal disorder according to International Statistical Classification of Diseases, 10th Revision (minimum symptom duration of 2 years) or with studies on schizotypy reporting an increased risk of psychosis. Furthermore, SIPS schizotypal personality disorder is part of NAPLS predictor genetic risk/recent deterioration in functioning, and a combination of attenuated positive symptoms and the UHR trait criterion that includes schizotypal personality disorder was among the 5 predictors in a study from the Australian Personal Assessment and Crisis Evaluation clinic. However, this particular combination was insignificant in our univariate testing and excluded from further analysis.

Loss of functioning was also predictive in the Australian study. Some overlap with the NAPLS predictors is apparent, as highest GAF-M score in the past year is part of genetic risk/recent deterioration in functioning; another NAPLS predictor, social functioning, is part of the determinants of highest GAF-M score in the past year. In line with our finding, deterioration of role function-

ing is a well-known characteristic of the prespsychotic phase and develops well before the first psychotic symptom. In our sample, the selection of highest GAF-M score in the past year suggests a functional deterioration starting well before the 12 months prior to baseline. At 65, it is only just above the maximum score of 60 used as a threshold in defining serious impairment and within the 0 to 70 range found to be predictive of psychiatric caseness. Removal of current functioning at baseline and relative 12-month functional decline from the model indicates that a long-standing deterioration is more strongly associated with the risk of transition to psychosis than a recent significant decline.

In first-episode psychosis, years of education correlated significantly with working memory and verbal learning. In schizophrenia, years of education were associated inversely with symptom severity and positively with neuropsychological performance. The years of education—associated measure decline in school functioning frequently occurred in the prodromal phase in retrospective first-episode studies as well as in a clinical high-risk sample. Contrary to the NAPLS finding on genetic risk/recent deterioration in functioning, family history of psychosis was not retained in the EPOS model, neither alone nor combined with a reduction in GAF score. This may be related to the comparably lower proportion of patients with a positive family history in the EPOS sample, which, nevertheless, was in the range expected for a clinically defined high-risk sample. The proportion of genetic high-risk subjects may, in fact, be the most striking difference between the NAPLS and EPOS samples.

The EPOS was not primarily designed to study the benefits of pharmacological intervention. This requires randomized controlled trials, while the focus of EPOS was broad and medication only one of several dimensions. Thus, drug treatment was recorded retrospectively at follow-up, and detailed information on medication could not always be obtained. With this limitation presumed, only antipsychotics had a significant but marginal impact on the model’s predictors and the prognostic accuracy in only the medication-naive subsample remained well comparable with that of the sample. Accordingly, treatment did not limit the validity of the EPOS findings. This and the similar result from NAPLS appear counterintuitive to findings of controlled studies, indicating a beneficial effect of these substances in the prodromal phase. Most likely, the NAPLS and our findings reflect the clinicians’ decision to prescribe antipsychotics to patients presenting with more severe, psychosis-like symptoms who are already on the edge of transition. This argument is further supported by higher SIPS-Positive and lower GAF-M scores in the antipsychotic-treated group at baseline.

Although the incidence rates of psychosis in early detection studies, including EPOS, are several hundred times higher than in the general population, the high proportion of seemingly false-positives—at least within shorter follow-up periods—has fostered ethical concerns. However, the progression of transition rates over longer follow-up periods supports the assump-
tion that increasing numbers of patients are being referred earlier to early-detection centers and, consequently, a change in the average time to transition but not necessarily in the risk has occurred in these growing cohorts. The flattening of the incidence curve over time is a frequently observed phenomenon.\textsuperscript{8,18,44,75,76} This may be connected to a bias toward an artificial early incidence peak resulting from the proportion of patients who are only referred late, when symptoms have become more extreme and psychotic-like experiences have already occurred.\textsuperscript{8} In line with this view, the EPOS-PI class IV displayed the shortest mean survival time.

Two methodological limitations of EPOS have to be addressed: overfit and the naturalistic study design. As fitted models always perform in an “optimistic manner”\textsuperscript{40} on the model-development data, cross-validation in an independent sample is needed to control for tailormade modeling. Yet, despite the overall large sample size, the limited number of transitions did not allow for data splitting for statistical reasons. Existing or future samples of comparable size and risk definition are required to finally validate the EPOS prediction model.

The naturalistic design has also been a criticism of NAPLS.\textsuperscript{24} With regard to treatment effects, randomized controlled trials may be better, and the lack of concurrent recording of applied treatments is an unfortunate caveat of the current study. However, even if treatment had been sufficiently recorded, the number of patients who received medication would still have remained too small to enable clear conclusions when compared with randomized controlled trials. Yet, randomized controlled trials suffer from potential recruitment bias introduced by narrow criteria and requiring agreement with randomization, which, compared with the population of individuals at risk, may result in nonrepresentative samples.\textsuperscript{80} This is illustrated by a randomized controlled trial\textsuperscript{73} in which subjects positive under UHR criteria who agreed to randomization and entered the control condition showed a 12-month transition rate of 35.7%, whereas those refusing randomization showed a rate of only 18.2%.

In conclusion, our findings signal a methodological advancement in the early detection of psychosis: UHR and COGDIS criteria together serve as a first-step detection tool for a generally increased risk of psychosis, and the prognostic scores serve as a multivariate second-step tool for further risk classification in terms of magnitude and time. This procedure might enable not only a differentiated, possibly change-sensitive estimation of current individual risk, but also risk-adapted treatment approaches and a risk-enrichment strategy necessary for clinical trials.\textsuperscript{24,81} Next, it will have to be investigated whether a multilevel model including additional neurocognitive, neurobiological, sociobiographical, and environmental variables will increase the predictive accuracy even further. In addition, future analyses will have to test whether this model also applies to the prediction of psychosis within longer time frames.

Submitted for Publication: July 6, 2008; final revision received May 13, 2009; accepted June 29, 2009.

Author Affiliations: Department of Psychiatry and Psychotherapy, University of Cologne, Cologne, Germany (Drs Ruhrmann, Schulze-Lutter, and Klosterkötter and Mr Graf von Reventlow); Department of Psychiatry, University of Turku, Turku, Finland (Drs Salokangas and Heinimaa); Department of Psychiatry, Academic Medical Center, University of Amsterdam, Amsterdam, the Netherlands (Drs Linszen and Dingemans); School of Psychology, University of Birmingham, Birmingham, England (Drs Birchwood and Patterson); Department of Psychiatry, Psychotherapy, and Psychosomatic Medicine, Ruhr-University Bochum, Bochum, Germany (Dr Juckel); Department of Psychiatry and Psychotherapy, Charité University Medical Center, Berlin, Germany (Drs Juckel and Heinz); and School of Psychological Sciences (Dr Morrison) and Department of Psychiatry (Dr Lewis), University of Manchester, Manchester, England.

Correspondence: Stephan Ruhrmann, MD, Department of Psychiatry and Psychotherapy, University of Cologne, Kerpener Strasse 62, 50924 Cologne, Germany (stephan.ruhrmann@uk-koeln.de).

Financial Disclosure: None reported.

Funding/Support: This study was funded by grant QLG4-CT-2001-01081 from the 5th Framework Program Quality of Life and Management of Living Resources of the European Commission. Additional local funding was provided by the Social Insurance Institution of Finland and the Medical Faculty of Cologne (Köln Fortune Program).

Role of the Sponsor: The funding sources had no influence on the design or conduct of the study; collection, management, analysis, or interpretation of the data; or preparation, review, or approval of the manuscript. All authors had full access to all the data in the study.

Previous Presentations: Reported findings were presented in part at the 16th European Congress of Psychiatry; Nice, France; April 9, 2008; and the First Schizophrenia International Research Society Conference; Venice, Italy; June 24, 2008.


Additional Contributions: The authors appreciate the helpful support of Walter Lehmacher, PhD, head of the Institute for Medical Statistics, Informatics and Epidemiology of the University of Cologne. The authors thank the following scientists for their great work for EPOS: Heinz Picker, PhD, Meike Neumann, DiplPsych, Anke Brockhaus-Dumke, MD, and Ralf Pukrop, PhD (Department of Psychiatry and Psychotherapy/FETZ, University of Cologne, Cologne, Germany); Tanja Svirskis, MD, PhD, Jukka Huttunen, BM, Tiina Laine, MSc, Tuula Ilonen, PhD, Terja Riski, MA, and Jarmo Hietala, MD, PhD (Department of Psychiatry, University of Turku, Turku, Finland); Hiske Becker, MD, and Doreen Nieman, PhD (Academic Medical Centre de Meren, Department of Psychiatry, University of Amsterdam, Amsterdam, the Netherlands); Amanda Skeate, DClin (Early Detection and Intervention Team, Early Intervention Service, Birmingham, England); Yehonala Gudowsk, DiplPsych (Charité University Medical Center, Berlin, Germany); Seza Oszárdal, DiplPsych, and Henning Wittaus, MD (Charité, Berlin, and Ruhr-University, Bochum, Germany); and Paul French, PhD, and Helen...
Stevens, BSc (School of Psychological Sciences, University of Manchester, Manchester, England).

REFERENCES


