**SUPPLEMENTARY MATERIAL**

Supplementary Methods.

Supplementary Table 1. Sociodemographic and clinical characteristics of the subjects at follow-up.

Supplementary Table 2. Baseline diagnoses predicting any diagnostic change and change to schizophrenia at follow-up.

Supplementary Table 3. Summary of hierarchical logistic regression analysis for baseline variables uniquely predicting diagnostic change to schizophrenia.

Supplementary Table 4. Univariable logistic regression predicting the effect of baseline variables on diagnostic change to schizoaffective disorder over the follow-up.

Supplementary Table 5. Univariable logistic regression predicting the effect of baseline variables on diagnostic change to bipolar disorder over the follow-up.

Supplementary Table 6. Summary of hierarchical logistic regression analysis for baseline variables uniquely predicting diagnostic change to bipolar disorder

SUPPLEMENTARY METHODS. METHODOLOGY FOR ASSESSING PREDICTOR AND OUTCOME VARIABLES

1. General methodology

The main instrument for assessing sociodemographics, baseline predictors, and clinical features were the different sections from the Comprehensive Assessment of Symptoms and History (CASH) (Andreasen, 1987; Andreasen, Flaum, & Arndt, 1992). For some relevant variables not included in the CASH, specific assessment instruments were employed (see below).

A major advantage of the CASH is that it provides broad descriptive coverage of the lifetime history, by means of a Life Chart Interview (LCI), symptoms and diagnoses. The information included within the CASH allows the diagnosis of the patients using a variety of criteria, which is especially important because of the changing diagnostic systems over the study period. In this manner, we could diagnose all the subjects at baseline using the DSM-III-R (American Psychiatric Association, 1987) or DSM-IV (APA, 1994) criteria and rediagnose them with the DSM-5 (APA, 2013) criteria. The CASH also contains the Scale for the Assessment of Positive Symptoms (SAPS) and the Scale for the Assessment of Negative Symptoms (SANS).

All the variables included in the CASH were rated using multiple sources of information, including interviews with the participants, clinical records, first-degree relatives, significant others, and, if necessary, information provided by the primary physician.

We assessed inter-rater reliability for most of the baseline predictors, which has been reported in several papers from our research group and found to be adequate for most of them (kappa or ICC >0.80). For the other predictors, consensus ratings were obtained by the research team through regular meetings as described elsewhere (Peralta et al., 2021).

2. Predictor variables

2.1. Sociodemographic factors

Sociodemographic factors were all assessed with the CASH and included gender, years of education, age at baseline and follow-up assessments, and length of follow-up.

2.2. Familial-genetic liability

Familial-genetic factors included Polygenic Risk Scores (PRS) and the family history for schizophrenia, bipolar disorder, and major depressive disorder.

Polygenic risk score Genome-wide genotyping was performed in a sample of 173 subjects, who consented to DNA extraction, using the Illumina Global Screening Array (730 059 genetic variants). The final quality-controlled dataset ready for imputation consisted of 164 subjects (94.8% of the initial sample) and 489 135 genetic markers (67.0% of the initial sample). The entire genotyping procedure has been described elsewhere (Cuesta et al., 2023). We calculated PRS for schizophrenia, bipolar disorder, major depressive disorder, and educational attainment (as a proxy measure of cognitive reserve), and only the first three were used in the present study. The PRS variables were categorized using the 75th quantile of the original continuous variables, with the highest quartiles (>75%) considered as genetic exposure risk states.

A familial history of schizophrenia spectrum disorders (SSD), bipolar disorder and major depressive disorder was assessed in the first-degree relatives of the participants by means of the Family History-Research Diagnostic Criteria (FH-RDC) (N. C. Andreasen, Endicott, Spitzer, & Winokur, 1977), which was administered at baseline and final follow-up interviews. The combined information of the two interviews was used to rate the family history. SSD included all non-affective psychotic disorders plus schizotypal personality disorder.

2.3. Antecedents

Antecedents included obstetric complications, developmental delay at year 3, childhood adversity, premorbid adjustment, premorbid cognitive reserve, premorbid social networks, drug abuse and acute psychosocial stressors.

Obstetric complications were assessed with the Lewis & Murray scale (Lewis S, Owen R, Murray R, 1989) and neurodevelopmental delay with the Shapiro et al., (1990) scale. The latter scale rates developmental milestones attainment at age 3, including sitting, standing, walking, talking words, talking sentences and urine/faces control. These two variables were rated using clinical records and information provided by the subjects’ mother, which was available in the majority of the cases.

Childhood adversity was assessed by means of the Global Family Environment Scale (GFES) (Rey et al., 1997), which indexes the global quality of the environment in which the child was raised. Raters use a hypothetical continuum from 1 (e.g., severe abuse, deprivation) to 90 (e.g., stable and secure nurturing) and formulate a single score reflecting the lowest quality of family environment to which the child has been exposed. The GFES was not available at the beginning of the baseline recruitment period; thus, in 28% of the cases ratings were made using the rich available background information on this variable.

The modified Gittelman-Klein scale (GKS), as included in the CASH, was used to rate premorbid psychosocial adjustment during childhood (ages 6-12) and adolescence (ages 13-18). For the present study, the GKS total score (childhood plus adolescence scores) was employed.

Premorbid cognitive reserve was estimated according to established proxy measures of premorbid intelligence, education and leisure activities (Amoretti et al., 2020; Barnett, Salmond, Jones, & Sahakian, 2006). Premorbid intelligence was assessed by means of the Word Accentuation Test (WAT), which is the Spanish equivalent of the National Adult Reading Test. We used the WAIS III full scale IQ equivalence of the WAT scores as reported by Gomar et al., (2011) to obtain the premorbid IQ scores. Educational level was assessed using the years of education completed beyond the compulsory education and the scholastic performance subscale from the Cannon-Spoor scale (Cannon-Spoor, Potkin, & Wyatt, 1982). Participation in leisure activities was rated according to the peer relationships and interests subscales from the GKS. Higher scores were arranged to denote better performance and a Principal Component Analysis was performed, which resulted in a single factor, to create a premorbid cognitive reserve score for each subject (Amoretti et al., 2020)

Premorbid social networks were assessed by means of the Sturtees’ social support scale (Sturtees, 1980), which rates several dimensions of close and diffuse social networks of the subject. The scale was applied to the adolescence period with higher ratings indicating poorer social support.

Proximal antecedents were conceptualized as trigger factors occurring within the 6 months before illness onset. They included acute psychosocial stressors rated per DSM-III Axis IV (APA, 1980), and severity of drug abuse as rated using the global rating severity score (range:0-9) from the Addiction Severity Scale (McLellan et al., 1985).

2.4. Illness-onset variables

The variables were assessed with the CASH and included age at illness onset, chronicity of onset, duration of untreated psychosis (DUP) and duration of untreated continuous psychosis (DUCP).

Age at onset was defined as the age at which the subject met DSM criterion A for schizophrenia. DUP was defined as the months that elapsed between the appearance of the first psychotic symptom and the first antipsychotic treatment. DUCP was defined as the months that elapsed between the appearance of the first continuous psychotic symptom (i.e., present most of the days) and the first antipsychotic treatment. Mode of onset was rated from 1 (acute, <1 month) to 4 (chronic, >6 months), indicating the time elapsed between the onset of any illness-related symptom and the development of the full psychotic syndrome.

2.5. First-episode characteristics

This set of variables included duration of index admission in weeks (a proxy for initial illness severity), primary neuromotor abnormalities assessed at the drug-naïve status, psychopathological syndromes assessed at intake, and DSM-5 diagnosis assessed 6 months after discharge.

Primary neurological abnormalities were assessed in those drug-naïve participants at index admission (n=194, 79.8% of the sample) by means of a structured neurological examination. We assessed spontaneous dyskinesia and parkinsonism using the Abnormal Involuntary Movements Scale (Guy, 1976) and the Simpson-Angus Rating Scale (Simpson & Angus, 1970). The Neurological Examination Scale (Buchanan & Heinrichs, 1989) was also administered to those drug-naïve participants who were able to collaborate in the exploration of NSS (n=179, 73.7% of the sample).

Psychopathological syndromes at intake were assessed with the current psychopathological status from the CASH considering the most severe psychopathology over the last month and included global ratings for reality-distortion (psychosis), disorganization, negative, catatonia and affective (mood) syndromes.

DSM-5 diagnosis was assessed using the CASH 6 months after discharge from the hospital. Diagnosis was made at this time point to account for the 6-month criterion duration of schizophrenia definition.

 2.6. Early treatment response

Early treatment response at discharge from index admission was assessed by means of the Clinical Global Improvement Index (Guy, 1976). Early treatment response at 6-month after discharge was rated using the Scale for the Assessment of Positive Symptoms (SAPS) and the Scale for the Assessment of Negative Symptoms (SANS) as included in the CASH, and the Remission in Schizophrenia Working Group criteria (Andreasen et al., 2005) were employed to define symptomatic remission. These criteria require a score ≤2 (mild or less) in the 8 SAPS and SANS global ratings and a period of at least 6 months of symptomatic remission.

3. Outcome measures at follow-up

3.1. Diagnosis

DSM-5 diagnoses (APA, 2013) were rated at follow-up considering the whole illness course. Lifetime diagnoses were established using a follow-back methodology and all available information including health and social records, and interviews with the subject and significant others at the follow-up assessment. Two research psychiatrists (DP and LJ) using the Life Chart Interview (LCI) from the CASH constructed chart records using levels of symptoms and functioning such as other illness-related characteristics to rate diagnostic status across the following time periods: baseline (from age at illness onset to the 6-month assessment after index admission, baseline diagnosis), year 2, year 5, year 10, year 15, year 20 and follow-up visit. The follow-up diagnosis was taken as the reference standard and information from these time periods served to determine when that diagnosis could be implemented for the first time (age at diagnostic change). If baseline and follow-up diagnoses coincided, the diagnosis was considered stable. The senior authors (VP and MJC), using all available information from the LCI and a consensus methodology determined final/lifetime DSM-5 diagnoses and time of diagnostic change, if appropriate.

3.2. Other outcome measures

The RSWG criteria (Andreasen et al., 2005) were employed to define symptomatic remission.

Functional recovery was rated by means of the Social and Occupational Functioning Assessment Scale (SOFAS) (APA, 1994). Functional recovery was defined as a SOFAS score ≥61 over the last year.

Personal recovery was rated using the 15-item version of the Questionnaire about the Process of Recovery (QPR-15)(Law, Neil, Dunn, & Morrison, 2014). The QPR-15 was developed in collaboration with service users and covers the major processes of personal recovery, including the establishment of identity, finding meaning in life, taking responsibility for recovery, and having a sense of purpose and hope. The QPR-15 is a self-rated scale where each item consists of a declarative statement with a five-point Likert scale that ranges from 0 (“strongly disagree”) to 4 (“strongly agree”), where higher scores indicate recovery.

Lastly, we rated the 5 psychosis syndromes as described above (section 2.5.).

REFERENCES

American Psychiatric Association (1987). *Diagnostic and Statistical Manual of Mental Disorders (3rd ed., revised)*. Washington: APA

American Psychiatric Association (1994). *Diagnostic and Statistical Manual of Mental Disorders (4th edition)*. Washington, DC: APA

American Psychiatric Association. (2013). *Diagnostic and Statistical Manual of Mental Disorders* *(5th Edition)*. Washington DC: APA

Amoretti, S., Rosa, A. R., Mezquida, G., Cabrera, B., Ribeiro, M., Molina, M., … PEPs Group. (2020). The impact of cognitive reserve, cognition and clinical symptoms on psychosocial functioning in first-episode psychoses. *Psychological Medicine*, 1-12. doi: 10.1017/S0033291720002226

Andreasen, N. C., Endicott, J., Spitzer, R. L., & Winokur, G. (1977). The family history method using diagnostic criteria. Reliability and validity. *Archives of General Psychiatry*, *34*(10), 1229-1235. doi: 10.1001/archpsyc.1977.01770220111013

Andreasen, N. C., Flaum, M., & Arndt, S. (1992). The Comprehensive Assessment of Symptoms and History (CASH). An instrument for assessing diagnosis and psychopathology. *Archives of General Psychiatry*, *49*(8), 615-623. doi: 10.1001/archpsyc.1992.01820080023004

Andreasen, Nancy C., Carpenter, W. T., Kane, J. M., Lasser, R. A., Marder, S. R., & Weinberger, D. R. (2005). Remission in schizophrenia: Proposed criteria and rationale for consensus. *The American Journal of Psychiatry*, *162*(3), 441-449. doi: 10.1176/appi.ajp.162.3.441

Barnett, J. H., Salmond, C. H., Jones, P. B., & Sahakian, B. J. (2006). Cognitive reserve in neuropsychiatry. *Psychological Medicine*, *36*(8), 1053-1064. doi: 10.1017/S0033291706007501

Buchanan, R. W., & Heinrichs, D. W. (1989). The Neurological Evaluation Scale (NES): A structured instrument for the assessment of neurological signs in schizophrenia. *Psychiatry Research*, *27*(3), 335-350. doi: 10.1016/0165-1781(89)90148-0

Cannon-Spoor, H. E., Potkin, S. G., & Wyatt, R. J. (1982). Measurement of premorbid adjustment in chronic schizophrenia. *Schizophrenia Bulletin*, *8*(3), 470-484. doi: 10.1093/schbul/8.3.470

Cuesta, M. J., Papiol, S., Ibañez, B., García de Jalón, E., Sánchez-Torres, A. M., Gil-Berrozpe, G. J., … SEGPEPs Group. (2023). Effect of polygenic risk score, family load of schizophrenia and exposome risk score, and their interactions, on the long-term outcome of first-episode psychosis. *Psychological Medicine*, 1-10. doi: 10.1017/S0033291723000351

Gomar, J. J., Ortiz-Gil, J., McKenna, P. J., Salvador, R., Sans-Sansa, B., Sarró, S., … Pomarol-Clotet, E. (2011). Validation of the Word Accentuation Test (TAP) as a means of estimating premorbid IQ in Spanish speakers. *Schizophrenia Research*, *128*(1-3), 175-176. doi: 10.1016/j.schres.2010.11.016

Guy W. (1976). Abnormal Involuntary Movement Scale (AIMS). En *ECDEU Assessment Manual for Psychopharmacology*. Washington: US Dept Health Education and Welfare.

Law, H., Neil, S. T., Dunn, G., & Morrison, A. P. (2014). Psychometric properties of the questionnaire about the process of recovery (QPR). *Schizophrenia Research*, *156*(2-3), 184-189. doi: 10.1016/j.schres.2014.04.011

Lewis S, Owen R, Murray R. (1989). Obstetric complications and schizophrenia. En *Schizophrenia: Scientific Progress* (SC S, CA T, eds., pp. 56-68). New York: Oxford University Press.

McLellan, A. T., Luborsky, L., Cacciola, J., Griffith, J., Evans, F., Barr, H. L., & O’Brien, C. P. (1985). New data from the Addiction Severity Index. Reliability and validity in three centers. *The Journal of Nervous and Mental Disease*, *173*(7), 412-423. doi: 10.1097/00005053-198507000-00005

Rey, J. M., Singh, M., Hung, S. F., Dossetor, D. R., Newman, L., Plapp, J. M., & Bird, K. D. (1997). A global scale to measure the quality of the family environment. *Archives of General Psychiatry*, *54*(9), 817-822. doi: 10.1001/archpsyc.1997.01830210061006

Shapiro, B. K., Palmer, F. B., Antell, S., Bilker, S., Ross, A., & Capute, A. J. (1990). Precursors of reading delay: Neurodevelopmental milestones. *Pediatrics*, *85*(3 Pt 2), 416-420.

Simpson, G. M., & Angus, J. W. (1970). A rating scale for extrapyramidal side effects. *Acta Psychiatrica Scandinavica. Supplementum*, *212*, 11-19. doi: 10.1111/j.1600-0447.1970.tb02066.x

Sturtees PG (1980). Social support, residual adversity and depressive outcome. British Journal of Psychiatry 15, 71–80.

Supplementary Table 1. Sociodemographic and clinical characteristics of the subjects at follow-up

|  |  |  |
| --- | --- | --- |
|  | N (%) | Mean (SD) |
| Gender (male) | 137 (56.4) |  |
| Civil status (single) | 150 (61.7) |  |
| Living: |  |  |
|  Own family | 71 (29.2) |  |
|  Other family members | 54 (22.2) |  |
|  Other persons | 17 (7.0) |  |
|  Supported housing | 37 (15.2) |  |
| Employment (paid working) | 80 (33) |  |
| DSM-5 illness course: |  |  |
|  Full remission | 73 (30.0) |  |
|  Partial remission | 149 (61.3) |  |
|  Chronic/continuous | 59 (24.3) |  |
| Comorbid drug use: | 121 (49.8) |  |
| Psychiatric medication: |  |  |
|  Antipsychotics | 182 (74.9) |  |
|  Mood stabilizers | 72 (29.6) |  |
|  Antidepressants | 81 (33.3) |  |
|  Anxiolytics/hypnotics | 105 (43.2) |  |
|  None | 41 (16.9) |  |
| Age, years |  | 48.5 (10.4) |
| No. of psychiatric admissions |  | 5.85 (6.24) |
| GAF |  | 64.0 (19.8) |
| SOFAS |  | 62.8 (21.4) |
| SAPS, global ratings total score |  | 2.86 (3.66) |
| SANS, global ratings total score |  | 5.89 (4.96) |

DSM-5= Diagnostic and Statistical Manual, fifth edition; GAF= Global Assessment of Functioning Scale; QPR= Questionnaire about the Process of Recovery; SAPS= Scale for the Assessment of Positive Symptoms; SANS= Scale for the Assessment of Negative Symptoms; SOFAS= Social and Occupational Functioning Assessment Scale.

Supplementary Table 2. Baseline diagnoses predicting any diagnostic change and change to schizophrenia at follow-up.

|  |  |  |
| --- | --- | --- |
|  | Any diagnostic change (n=127) | Change to schizophrenia (n=47) |
| OR | 95% CI | p | OR | 95% CI | p |
| Schizophrenia | 0.03 | 0.01 – 0.09 | **<0.001** |  |  |  |
| Schizophreniform disorder | 6.70 | 2.69 – 16.6 | **<0.001** | 3.69 | 1.76 – 7.74 | **0.001** |
| Brief psychotic disorder | 1.96 | 0.97 – 3.96 | 0.059 | 0.83 | 0.34 – 2.01 | 0.687 |
| Delusional disorder | 4.66 | 1.30 – 16.6 | **0.018** | 4.26 | 1.54 – 11.7 | **0.005** |
| Schizoaffective disorder | 1.29 | 0.39 – 4.19 | 0.667 | 3.21 | 0.97 – 10.6 | 0.056 |
| Mania/ bipolar disorder | 0.36 | 0.13 – 0.97 | 0.044 |  0.20 , | 0.02 – 1.55 | 0.124 |
| Major depression/ major depressive disorder | 2.67 | 1.13 – 6.30 | **0.025** | 0.44 | 0.12 – 1.54 | 0.202 |
| Psychosis not otherwise specified | 13.1† | 1.79 – 98.6 | **<0.001** | 6.68 | 2.01 – 22.1 | **0.002** |

† Estimate represents the Haldane-Anscombe OR correction because of zero cell.

In bold are presented the statistically significant associations after the Benjamini-Hochberg correction for multiple comparisons.

Supplementary Table 3. Summary of hierarchical logistic regression analysis for baseline variables uniquely predicting diagnostic change to schizophrenia

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  |  OR (95% CI) | p | R2 | % Correct |
| Step 1 (familial-genetic liability): |  |  |  |  |
|  Family history of schizophrenia | 3.48 (1.56 – 7.75) | 0.002 | 0.076 | 72.5 |
| Step 2 (very distal antecedents): |  |  |  |  |
|  Obstetric complications | 4.15 (1.57 – 10.9) | 0.004  | 0.140 | 76.0 |
| Step 3 (distal antecedents): |  |  |  |  |
|  Developmental delay | 4.03 (1.80 – 9.08) | 0.001 | 0.226 | 77.8 |
| Step 5 (proximal antecedents): |  |  |  |  |
|  Psychosocial stressors | 0.38 (0.16 – 0.88) | 0.025 | 0.291 | 79.5 |
| Step 7 (first-episode characteristics): |  |  |  |  |
|  Spontaneous dyskinesia | 8.53 (1.20 – 59.7) | 0.031 | 0.357 | 83.0 |
| Step 8 (early treatment response): |  |  |  |  |
|  6-month symptomatic remission | 0.13 (0.03 – 0.53)  | 0.004 | 0.414 | 86.0 |

Note: The time-ordered blocks of variables (steps) entered successively in the model were as follows: Step 1 (familial-genetic liability); Step 2 (very distal antecedents); Step 3 (distal antecedents); Step 4 (childhood/adolescence functioning); Step 5 (proximal antecedents); Step 6 (illness-onset features); Step 7 (first-episode characteristics); and Step 8 (early treatment response). No variables at Step 4 and at Step 6 were significantly related to diagnostic change.

Supplementary Table 4. Univariable logistic regression predicting the effect of baseline variables on diagnostic change to schizoaffective disorder over the follow-up

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | No diagnostic change (n=197) † | Diagnostic change (n=34) † | OR (95% CI) | p |
| **Socio-demographics** |  |  |  |  |
| Age at follow-up, high (≥ 47 years) | 102 (51.8) | 21 (61.8) | 1.50 (0.71 – 3.17) | 0.283 |
| Male gender | 110 (55.8) | 19 (55.9) | 1.00 (0.48 – 2.08) | 0.996 |
| Education, high school | 87 (44.2) | 16 (47.1) | 1.12 (0.54 – 2.33) | 0.754 |
| Age at baseline assessment, high (≥25 years) | 96 (48.7) | 21 (61.8) | 1.70 (0.80 – 3.58) | 0.163 |
| Length of follow-up, high (≥21 years) | 101 (51.3) | 23 (67.6) | 1.98 (0.91 – 4.29) | 0.081 |
| **Familial-genetic liability** |  |  |  |  |
| PRS for schizophrenia, high (≥0.66) | 30 (23.1) | 10 (35.7) | 1.85 (0.77 – 4.43) | 0.167 |
| PRS for bipolar disorder, high (≥0.64) | 29 (22.3) | 12 (42.9) | 2.61 (1.11 – 6.14) | 0.028 |
| PRS for major depression, high, (≥0.68) | 33 (25.4) | 6 (21.4) | 0.80 (0.29 – 2.14) | 0.660 |
| Family history of schizophrenia spectrum disorders | 36 (18.6) | 8 (23.5) | 1.37 (0.56 – 3.28) | 0.473 |
| Family history of bipolar disorder | 18 (9.1) | 2 (5.9) | 0.62 (0.13 – 2.80) | 0.537 |
| Family history of major depressive disorder | 29 (14.7) | 5 (14.7) | 0.99 (0.35 – 2.79) | 0.998 |
| **Antecedents** |  |  |  |  |
| Obstetric complications, any definite | 37 (18.8) | 1 (2.9) | 0.13 (0.01 – 0.98) | 0.049 |
| Developmental delay at year 3, any | 98 (49.7) | 10 (29.4) | 0.42 (0.19 – 0.92) | 0.032 |
| Childhood adversity score, high (<77) | 99 (50.3) | 15 (44.1) | 0.50 (0.37 – 1.62) | 0.509 |
| Premorbid adjustment score, poor (≥4) | 89 (45.2) | 13 (38.2) | 0.75 (0.35 – 1.58) | 0.453 |
| Premorbid cognitive reserve score, low (≥41) | 88 (44.7) | 16 (47.1) | 1.01 (0.53 – 2.28) | 0.796 |
| Premorbid social networks score, poor (<4) | 89 (45.2) | 10 (29.4) | 0.50 (0.23 – 1.11) | 0.090 |
| Drug abuse, any | 69 (35.0) | 9 (26.5) | 0.66 (0.29 – 1.51) | 0.332 |
| Acute psychosocial stressors, any | 65 (33.0) | 15 (44.1) | 0.50 (0.23 – 1.11) | 0.090 |
| **Illness-onset variables** |  |  |  |  |
| Age at illness onset, early (≤22 years) | 104 (52.8) | 13 (38.2) | 0.55 (0.26 – 1.16) | 0.120 |
| Chronicity of onset (> 6months)  | 75 (38.1) | 6 (17.6) | 0.34 (0.13 – 0.88) | 0.026 |
| DUP, long (≥ 2 months) | 101 (51.3) | 11 (32.4) | 0.45 (0.21 – 0.98) | 0.045 |
| DUCP, long (≥ 1 month) | 102 (51.8) | 15 (44.1) | 0.73 (0.35 – 1.52) | 0.411 |
| **First-episode characteristics**  |  |  |  |  |
| Spontaneous dyskinesia, Schooler & Kane criteria  | 21 (13.2) | 0 (0) | 0.48 (0.20 – 0.98)‡ | 0.049 |
| Spontaneous parkinsonism score, high (≥ 4)  | 29 (18.2) | 0 (0) | 0.36 (0.15 – 0.82)‡ | 0.018 |
| Neurological soft signs score, high (≥ 15) | 57 (45.3) | 14 (63.6) | 2.11 (0.83 – 5.34) | 0.113 |
| Psychosis syndrome, global rating score >2 | 158 (80.2) | 30 (88.2) | 1.85 (0.61 – 5.56) | 0.273 |
| Disorganization syndrome, global rating score >2 | 96 (48.7) | 11 (32.4) | 0.53 (0.23 – 1.08) | 0.081 |
| Negative syndrome, global rating score >2 | 39 (19.8) | 3 (8.8) | 0.39 (0.11 – 1.34) | 0.138 |
| Catatonia syndrome, global rating score >2 | 23 (11.7) | 4 (11.8) | 1.00 (0.32 – 3.12) | 0.998 |
| Affective syndrome, global rating score >2 | 55 (27.9) | 15 (44.1) | 2.03 (0.96 – 4.29) | 0.061 |
| Duration of index admission, long (≥3 weeks) | 112 (56.9) | 16 (47.1) | 0.67 (0.32 – 1.40) | 0.291 |
| **Early treatment response** |  |  |  |  |
| CGI-EI at index discharge, marked improvement  | 118 (59.9) | 23 (67.6) | 1.40 (0.64 – 3.03) | 0.394 |
| 6-month symptomatic remission after index admission  | 135 (68.5) | 28 (82.4) | 2.14 (0.84 – 5.44) | 0.109 |

† Data are number (and percentages) of the stated features.

‡ Estimate represents the Haldane-Anscombe OR correction because of zero cell.

In bold are presented statistically significant associations after the Benjamini-Hochberg correction for multiple testing.

CGI-EI = Clinical global impression-Efficacy Index; DUP = Duration of Untreated Psychosis; DUCP = Duration of Untreated Continuous Psychosis; PRS = Polygenic Risk Score.

Supplementary Table 5. Univariable logistic regression predicting the effect of baseline variables on diagnostic change to bipolar disorder over the follow-up

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | No diagnostic change(n=195) † | Diagnostic change (n=28) † | OR (95% CI) | p |
| **Socio-demographics** |  |  |  |  |
| Age at follow-up, high (≥ 47 years) | 102 (52.3) | 14 (50.0) | 0.91 (0-41 – 2.01) | 0.819 |
| Male gender | 110 (56.4) | 15 (53.6) | 0.89 (0.40 – 1.97) | 0.777 |
| Education, high school | 85 (43.6) | 14 (50.0) | 1.29 (0.58 – 2.86) | 0.524 |
| Age at baseline assessment, high (≥25 years) | 99 (55.8) | 12 (42.9) | 0.72 (0.32 – 1.68) | 0.435 |
| Length of follow-up, high (≥21 years) | 101 (51.8) | 12 (42.9) | 0.69 (0.31 – 1.55) | 0.378 |
| **Familial-genetic liability** |  |  |  |  |
| PRS for schizophrenia, high (≥0.66) | 28 (21.5) | 8 (40.0) | 2.42 (0.90 – 6.52) | 0.078 |
| PRS for bipolar disorder, high (≥0.64) | 29 (22.3) | 8 (40.0) | 2.32 (0.86 – 6.21) | 0.094 |
| PRS for major depression, high, (≥0.68) | 35 (26.9) | 5 (25.0) | 0.90 (0.30 – 2.67) | 0.856 |
| Family history of schizophrenia spectrum disorders | 45 (23.1) | 2 (7.1) | 0.25 (0.05 – 1.12) | 0.071 |
| Family history of bipolar disorder | 13 (6.7) | 3 (10.7) | 1.68 (0.44 – 6.30) | 0.442 |
| Family history of major depressive disorder | 30 (15.4) | 4 (14.3) | 0.91 (0.29 – 2.83) | 0.880 |
| **Antecedents** |  |  |  |  |
| Obstetric complications, any definite | 41 (21.0) | 0 (0) | 0.23 (0.09 – 0.66)‡ | **0.007** |
| Developmental delay at year 3, any | 100 (51.3) | 9 (32.1) | 0.45 (0.19 – 1.04) | 0.063 |
| Childhood adversity score, high (<77) | 110 (56.4) | 5 (17.9) | 0.16 (0.06 – 0.46) | **0.001** |
| Premorbid adjustment score, poor (≥4) | 104 (53.3) | 4 (14.3) | 0.14 (0.49 – 0.43) | **0.001** |
| Premorbid cognitive reserve score, low (≥41) | 100 (51.3) | 7 (25.0) | 0.31 (0.12 – 0.77) | **0.012** |
| Premorbid social networks score, poor (<4) | 97 (49.7) | 6 (21.4) | 0.27 (0.10 – 0.70) | **0.008** |
| Drug abuse, any | 66 (33.8) | 8 (28.6) | 0.78 (0.32 – 1.87) | 0.580 |
| Acute psychosocial stressors, any | 60 (30.8) | 17 (60.7) | 3.47 (1.53 – 7.87) | **0.003** |
| **Illness-onset variables** |  |  |  |  |
| Age at illness onset, early (≤22 years) | 103 (52.8) | 16 (57.1) | 1.19 (0.53 – 2.64) | 0.668 |
| Chronicity of onset (> 6months)  | 83 (42.6) | 4 (14.3) | 0.22 (0.07 – 0.67) | **0.008** |
| DUP, long (≥ 2 months) | 107 (54.9) | 10 (35.7) | 0.45 (0.20 – 1.04) | 0.062 |
| DUCP, long (≥ 1 month) | 115 (59.0) | 8 (28.6) | 0.27 (0.11 – 0.66) | **0.004** |
| **First-episode characteristics**  |  |  |  |  |
| Spontaneous dyskinesia, Schooler & Kane criteria  | 20 (12.9) | 0 (0) | 0.26 (0.05 – 1.11)‡ | 0.067 |
| Spontaneous parkinsonism score, high (≥ 4)  | 28 (18.1) | 2 (8.7) | 0.43 (0.09 – 1.95) | 0.275 |
| Neurological soft signs score, high (≥ 15) | 75 (51.7) | 5 (25.0) | 0.31 (0.10 – 0.90) | 0.031 |
| Psychosis syndrome, global rating score >2 | 165 (84.6) | 19 (67.9) | 0.38 (0.15 – 0.92) | 0.034 |
| Disorganization syndrome, global rating score >2 | 86 (44.1) | 15 (53.6) | 1.46 (0.66 – 3.23) | 0.349 |
| Negative syndrome, global rating score >2 | 42 (21.5) | 2 (7.1) | 0.28 (0.06 – 1.22) | 0.092 |
| Catatonia syndrome, global rating score >2 | 22 (11.3) | 5 (17.9) | 1.70 (0.59 – 4.95) | 0.323 |
| Affective syndrome, global rating score >2 | 48 (24.6) | 13 (46.4) | 2.65 (1.18 – 5.97) | 0.018 |
| Duration of index admission, long (≥3 weeks) | 111 (56.9) | 13 (46.4) | 0.65 (0.29 – 1.45) | 0.298 |
| **Early treatment response** |  |  |  |  |
| CGI-EI at index discharge, marked improvement  | 103 (52.8) | 23 (82.1) | 4.10 (1.50 – 11.2) | **0.006** |
| 6-month symptomatic remission after index admission  | 125 (64.1) | 26 (92.9) | 7.28 (1.67 – 31.5) | **0.008** |

† Data are number (and percentages) of the stated features.

‡ Estimate represents the Haldane-Anscombe OR correction because of zero cell.

In bold are presented statistically significant associations after the Benjamini-Hochberg correction for multiple testing. CGI-EI = Clinical global impression-Efficacy Index; DUP = Duration of Untreated Psychosis; DUCP = Duration of Untreated Continuous Psychosis; PRS = Polygenic Risk Score.

Supplementary Table 6. Summary of hierarchical logistic regression analysis for baseline variables uniquely predicting diagnostic change to bipolar disorder

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  |  OR (95% CI) | p | R2 | % Correct |
| Step 1 (very distal antecedents): |  |  |  |  |
|  Obstetric complications | 0.23 (0.09 – 0.66) | 0.007  | 0.091 | 69.7 |
| Step 2 (premorbid functioning): |  |  |  |  |
|  Poor premorbid adjustment | 0.13 (0.02 – 0.71) | 0.019 | 0.214 |  75.4 |
| Step 3 (proximal antecedents): |  |  |  |  |
|  Psychosocial stressors | 2.64 (1.09 – 6.40) | 0.031 | 0.250 | 82.9 |
| Step 5 (early treatment response): |  |  |  |  |
|  6-month symptomatic remission | 4.54 (1.01 – 20.4)  | 0.049 | 0.282 | 86.0 |

Note: The time-ordered blocks of variables (steps) entered successively in the model were as follows: Step 1 (very distal antecedents); Step 2 (premorbid functioning); Step 3 (proximal antecedents); Step 4 (illness-onset features); and Step 5 (early treatment response). No variables at Step 4 were significantly related to diagnostic change.