WILEY

ORIGINAL ARTICLE   OPEN ACCESS

A Ten-Year Cohort Study on the Stability and Trajectories of Psychiatric Diagnosis in Adolescents and Young Adults in Lazio, Italy

Anna Forastiere[1] ⬤ | Enrico Calandrini[2] | Giulia Cesaroni[2] | Anna Maria Bargagli[2] | Nera Agabiti[2]

[1]Department of Mental Health, Local Health Authority Roma 2, Rome, Italy | [2]Department of Epidemiology of the Regional Health Service of Lazio, Rome, Italy

Correspondence: Anna Forastiere (annaforastiere@gmail.com)

ABSTRACT
Introduction: We aimed to evaluate the diagnostic stability of a large spectrum of psychiatric diagnoses in adolescents and young adult patients attending specialised facilities in the Lazio Region (Italy).
Methods: A total of 3871 subjects (11–35 years) at their first hospitalisation with a psychiatric diagnosis were selected and followed up for 10 years on subsequent hospital admissions to psychiatric care.
Results: A total of 1145 patients were readmitted to a hospital for a psychiatric disorder in the following 10 years. Among these subjects, the concordance between the first and last diagnosis was 57.8% with a weighted kappa of 0.47 (95% CI: 0.42–0.51); the repeated stability (the same diagnosis in at least 75% of admissions) was 31.2%. The diagnostic categories with the highest values of prospective concordance and kappa were schizophrenia spectrum and other functional psychoses (70%, kappa 0.53), substance use disorders (54%, kappa 0.57), and eating disorders (80.9%, kappa 0.76).
Conclusions: In a population study, the stability over time of the first psychiatric diagnosis in a hospital varied according to the specific diagnostic categories, and overall, it was lower than previously reported. The trajectories were disorder-specific, and the stability was influenced by several factors, including the individuals' characteristics, the disorder's severity, and the diagnostic setting.

1 | Introduction

The evaluation of the stability of the clinical diagnoses following the hospitalisation of a psychiatric patient has implications for clinical practice, research, and public health. At least two systematic reviews (Fusar-Poli et al. 2016; Palomar-Ciria et al. 2019) have indicated the relative stability of the diagnosis of schizophrenia. In the first study (Fusar-Poli et al. 2016), a high concordance was found between the first and last diagnosis of schizophrenia, affective spectrum psychosis, and schizoaffective disorder. In contrast, a weaker concordance was found for substance-induced psychotic disorder, delusional

disorder, acute psychotic disorder, and brief transient disorder. There was a low concordance for psychosis not otherwise specified (NOS) and schizophreniform disorder. In a subsequent meta-analysis focussed on schizophrenia, Palomar-Ciria et al. (2019) indicated that several factors were most commonly associated with diagnostic stability: male gender, older age at onset, late stages of illness, family history of mental illness, more compromised clinical status and extended hospital stay. Also, bipolar disorder has been well-studied (Cegla-Schvartzman et al. 2019; Cano-Ruiz et al. 2020; Cegla-Schvartzman et al. 2021) with the suggestion of relative stability. In contrast, the available evaluations for other mental

disorders are relatively scanty, with valuable population-based studies for depression (Kessing 2005), anxiety (Hovenkamp-Hermelink et al. 2016), personality disorders (Kjaer et al. 2016), eating disorders (Milos et al. 2005; Fichter and Quadflieg 2007). A recent longitudinal study evaluated a small group of first-episode psychosis with a follow-up of 21 years. It showed an overall diagnostic stability of 47.7%, with schizophrenia and bipolar disorder as the most stable diagnoses, with other categories having low stability (Peralta et al. 2023).

Most available investigations have been conducted in clinical settings under the control of several conditions with a brief interval between the first and the second evaluation (1 or 2 years). The stability in these situations is likely to be high and does not reflect the actual condition at the population level. Therefore, studies on the general population are preferable, and a long follow-up time allows for evaluating the long-term stability of the diagnoses. A recent population-based cohort study by Høj Jørgensen et al. (2023) analysed data on all patients experiencing their first contact with psychiatric hospitals in Denmark from 1995 to 2008 and mapped subsequent diagnoses over a 10-year period following an initial psychiatric diagnosis at hospital admittance. Over the follow-up, 46.9% of all patients had at least one subsequent diagnosis that differed from their initial diagnosis. The likelihood of receiving a psychiatric diagnosis differed from the initial one across diagnostic categories. The highest variability occurred with persistent delusional disorders, mental and behavioural disorders due to multiple drug use and use of other psychoactive substances, and acute and transient psychotic disorders. The lowest variability occurred with eating disorders, non-organic sexual dysfunction, and somatoform disorders.

Given the paucity of population studies, we conducted a longitudinal study at a population level to assess the diagnostic stability of a broad spectrum of psychiatric diagnoses in adolescents and young adults after their first hospitalisation up to 10 years. We evaluated the diagnostic trajectories for different categories. A secondary aim was to assess the personal and contextual determinants that contribute to the stability/instability of the diagnosis.

The work has been conducted on a large population and various diagnoses for subjects aged 11–35. This age range was selected to study people at their first hospitalisation, avoiding including individuals with long-lasting conditions that are usually more difficult to characterise.

2 | Methods

2.1 | Area of the Study and Sources of Information

In Italy, mental health care is organized at the local health district (LHD) level, with each district housing a mental health department (MHD) responsible for providing both hospital and community-based care within a defined catchment area. This integrated system promotes continuity of care, as hospital and community services operate under a unified management structure. In this model, the same mental health professionals who oversee outpatient care also participate in hospitalization decisions. Typically, a patient's admission to a hospital is determined by their treating mental health team, ensuring a smooth transition between community and inpatient care. However, in cases of acute psychiatric crises, hospitalization may also occur through emergency services. Private hospitals and mental health facilities in Italy can operate under agreements with the National Health Service (NHS) through an accreditation system managed by regional health authorities. This system allows private providers to deliver publicly funded mental health services while maintaining NHS standards. The NHS reimburses private hospitals using a Diagnosis-Related Group (DRG) system, which assigns fixed rates based on the type of treatment and complexity of care.

The study was conducted in the Lazio Region of Italy (5.7 million inhabitants in 2024), including the metropolitan area of Rome. Different data sources were integrated: the Regional Health Registry with personal identification of the residents, the Hospital Information System, and the Mortality Registry. For each person, data on hospitalisations in any health facility (public or private), containing codes of primary and secondary diagnoses recorded at hospital discharges, according to the IX International Classification of Diseases (ICD 9), has been available since 1995. Since 1983, the Regional Mortality Registry has provided information on the vital status and the cause of death (coded according to the ICD 9).

2.2 | Study Design

The study design is a population-based cohort (longitudinal study). We enrolled all the residents of the Lazio Region who were admitted (acute hospital admissions) to psychiatric facilities for adults or neuropsychiatric centres for children with a principal psychiatric diagnosis (principal diagnosis, see Supporting Information: Table S1) in the 5 years 2005–2009, aged 11–35 years. To select only incident cases, we excluded all subjects hospitalised in the Region for a psychiatric diagnosis (in any position, including secondary diagnoses) in the previous 8 years preceding the first hospitalisation.

The following diagnostic categories were chosen according to the ICD-9 CM: Schizophrenia and other functional psychoses (with the following subcategories: Schizophrenia, Schizophreniform Disorder, Schizoaffective Disorder, Delusional Disorder, Brief Psychotic Disorder or Other Reactive Psychosis, NOS Psychosis), Bipolar Disorders, Depression, Anxiety and Dissociative and Somatoform Disorders, Obsessive Compulsive Disorder, Personality and Behavioural Disorders, Substance Use Disorders, Post-Traumatic Stress Disorder, Eating Disorders, Autism, ADHD, Behaviour and Conduct Disorders (see Supporting Information: Table S1).

For each subject, a 10-year follow-up for hospitalisation due to psychiatric disorders (principal diagnosis) in the period 2005–2019 was conducted. All hospital admissions beyond the first (occurring at least 30 days after the initial admission) in psychiatric wards across the Region were considered. In practice, all treatment episodes involved inpatient care. Three

stability indicators, calculated only for subjects with at least one admission after the first one, were used.

1. 'Repeated stability': the same diagnosis was observed in at least 75% of the subsequent hospitalisations Høj Jørgensen et al. (2023) (2/2, 2/3, or 3/4 hospitalisations and so on).

2. 'Prospective concordance' was assessed by cross-tabulating the diagnostic categories of the first and last hospitalisations (Baca-Garcia et al. 2007). The prospective concordance is equivalent to the proportion of individuals in a specific diagnostic category who maintain the same diagnosis in the last hospitalisation at follow-up. This would correspond to the positive predictive value if the first diagnosis were considered the gold standard.

3. Retrospective concordance' (Schwartz et al. 2000), when the proportion of individuals in a specific category at the last hospitalisation received the same diagnosis at the first hospitalisation, assuming that the first diagnosis represents the truth. This is conceptually similar to sensitivity.

2.3 | Statistical Analysis

In the evaluation of repeated diagnostic stability, 95% confidence intervals (CI) of a proportion were calculated. The agreement between diagnoses at the first and last evaluation was measured as crude agreement (prospective and retrospective agreement) and with the Kappa statistic.

An analysis of the determinants of diagnostic stability (prospective concordance, unstable vs. stable individuals) was also carried out. A logistic regression analysis was performed, calculating the Odds Ratio (OR) and relative 95% CI on the total number of individuals. The following characteristics at baseline were evaluated in association with instability (dependent variable): gender, age (in four classes), citizenship (Italian, foreign), educational level, type of health facility at first admission (Main Hospital, University Hospital, Scientific Institute, Local Hospital; to note that local hospitals were mainly private hospital with an agreement with the NHS), presence or absence of concomitant psychiatric diagnoses in secondary diagnosis, length of stay at the first hospitalisation (0–5 days, 6–15 days, 16+ days), and lag time between the first and the last hospitalisation (< 2 years, 2–5 years, 6+ years).

The patients' trajectories from the first diagnostic category to the last one were illustrated using 'alluvial graphs'. An alluvial graph is a type of data visualization used to represent flows and connections between different categories over time. Key features are 'flows', bands that connect different categories, representing the magnitude of the relationship or transition between them. The thickness of the bands corresponds to the magnitude.

3 | Results

A total of 6990 individuals aged 11–35 years admitted to hospital facilities in the Lazio region in 2005–2009 for the chosen diagnoses (Supporting Information: Table S1, only

principal diagnoses) were initially selected. Then, 3119 individuals admitted for psychiatric disorders in the previous eight years (principal and secondary diagnoses) were considered prevalent cases and therefore excluded; 3871 individuals were included in the study. Table 1 illustrates the frequencies of these subjects by diagnostic categories. The most frequent were schizophrenia and other functional psychoses (29.6%), bipolar disorder (11.7%), depression (8.7%), anxiety, dissociative and other somatic disorders (18.2%), personality disorders (12.3%), substance use disorders (17.4%). The number of cases for some diagnoses - obsessive-compulsive disorders, post-traumatic stress disorders, autism and ADHD - was too low, and these categories were no longer considered. Some patients (99, 2.6%) died during the follow-up, particularly among those diagnosed with schizophrenia (3.7%), depression (3.9%) and personality disorders (4.0%). A total of 1599 (41.3%) had no hospitalisation during the 10-year follow-up, and 1028 (26.3%) had a hospitalisation but with a non-psychiatric diagnosis. Then, 1145 individuals (29.6% of the initial cohort) had at least a second hospital admission with a psychiatric diagnosis during the follow-up and were enrolled to evaluate the diagnostic stability. Among these, the mean number of hospitalisation was 3.6 (SD: 4.9). The proportion of individuals with at least a second psychiatric hospitalisation was higher for schizophrenia (41.5%) and bipolar disorders (37.6%) and lower for anxiety disorders (13.2%) and substance use disorders (14.3%).

Table 2 shows the individual characteristics of the patients enrolled in the study and their selection process from the initial enrolment to the final inclusion. The 1145 subjects were, on average, 26-year-old (median 26 years, interquartile range, 20–31 ys.), there was a majority of males (55.0%), a minority of foreign citizens (135 subjects in the final study population) compared to those of Italian nationality, a majority of people with an educational qualification of middle and high school (84.1%) and a minority of graduates (3.1%), a minority of patients with psychiatric comorbidity (13.8%).

Table 3 shows the percentages of subjects with repeated stability for the various diagnostic categories. In the total sample, the repeated stability was 31.2% (95% CI: 28.5%–33.9%, based on a mean of 3.6 admissions). Some categories reached higher stability: schizophrenia (49.7%), bipolar disorders (50%), and eating disorders (50%), while other categories had lower repeated stability, such as schizophreniform disorders (8.1%) and anxiety disorders (17.2%).

Table S2 in the supplementary material presents the cross-tabulation of diagnoses at first and last admission. It should be noted that the category schizophrenia and other functional psychoses comprise six analytical subcategories. Based on these data, 'alluvial graphs' were drawn for all the categories (Figure 1), specifically for schizophrenia spectrum and bipolar disorders (Figure 2) and all the other individual categories (Figures S1 in the supplementary material), representing the diagnostic trajectories from the first to the last admission. General stability can be observed (Figure 1) in the number of diagnoses of schizophrenia and eating disorders, with a decrease in the frequency of other diagnoses. At the same time, there was a clear increase in the frequency of diagnoses of bipolar disorder

TABLE 1 | First hospitalisations, deaths, and subsequent admissions of patients enrolled by diagnostic category (Lazio 2005–2009).

Diagnostic categories	First hospital admission		Deaths		At least one subsequent admission		Mean (SD) subsequent admissions
	N	%[a]	N	%	N	%[b]	
TOTAL	3871		99	2.6%	1145	29.6%	3.61 (4.99)
1. Schizophrenia and other functional psychoses	1146	29.6%	33	2.9%	476	41.5%	3.56 (4.91)
1.1 schizophrenia	383	9.9%	14	3.7%	157	41.0%	3.47 (5.36)
1.2 schizophreniform disorder	188	4.9%	6	3.2%	86	45.7%	3.6 (5.89)
1.3 schizoaffective disorder	120	3.1%	3	2.5%	54	45.0%	4.63 (6.1)
1.4 delusional disorder (paranoia)	48	1.2%	—	—	22	45.8%	3.32 (4.29)
1.5 psychotic brief disorder or other non-specified reactive psychoses	40	1.0%	—	—	19	47.5%	4.89 (4.38)
1.6 psychoses NOS	102	2.6%	3	2.9%	27	26.5%	2.93 (3.06)
2. Bipolar disorders	452	11.7%	10	2.2%	170	37.6%	3.14 (3.43)
3. Depression	337	8.7%	13	3.9%	101	30.0%	3.83 (3.86)
4. Anxiety. Dissociative and somatoform disorders	706	18.2%	8	1.1%	93	13.2%	3.85 (6.92)
5. Obsessive-compulsive disorders	—	—	—	—	—		—
6. Personality and behavioural disorders	475	12.3%	19	4.0%	175	36.8%	4.67 (6.81)
7. Alcoholism and substance use disorders	672	17.4%	14	2.1%	96	14.3%	2.59 (3.06)
8. Post-traumatic stress syndrome	2	0.1%	—	—	1	50.0%	1 (.)
9. Eating disorders	69	1.8%	2	2.9%	26	37.7%	4.62 (5.46)
10. Autism	2	0.1%	—	—	1	50.0%	13 (.)
11. ADHD	—	—	—	—	—	—	—
12. Behavioural disorders	10	0.3%	—	—	6	60.0%	3.5 (3.08)

[a]Column percentages.
[b]Row percentages.

(from 169 subjects to 227 subjects, from 14.8% to 19.8%) and personality disorder diagnoses (from 174 to 212 subjects, from 15.2% to 18.5%). Specific trajectories of change were detected and illustrated here as an example (Figure 2): for the schizophrenic spectrum, there was stability for 70% of cases and a change in the remaining fraction, particularly towards bipolar disorder (14.3%) and depression (14.3%). Bipolar disorder remained relatively stable (59.4%) or changed in a fair proportion of cases (21.7%) for a diagnosis of schizophrenia.

The crude prospective and retrospective concordance and the Kappa index were calculated for each diagnostic category, and the values are illustrated in Table 4. Generally, a moderate concordance (57.8%) was observed in the total number of subjects with a weighted Kappa value of 0.47 (95% CI: 0.42, 0.52). Higher values of prospective concordance and kappa were reached by the spectrum of schizophrenia and other functional psychoses (70%, kappa 0.53), alcoholism and addiction (54%, kappa 0.57), and eating disorders (80.9%, kappa 0.76). Concordance was intermediate for schizophrenia (59%, kappa 0.41), schizoaffective disorders (35.2%, kappa 0.41), bipolar disorders (59.4%, kappa 0.4) and personality disorders (54.9%, kappa 0.38). Finally, concordance was modest for other conditions, including depression (36.6%, kappa 0.31), anxiety, and dissociative and somatic symptom disorders (22.6, kappa 0.28). Very low concordance and kappa values were recorded for delusional disorders (22.7%, kappa 0.17) and schizophreniform disorders (12.8%, kappa 0.08).

Table 5 presents the logistic regression analysis of all diagnostic groups to evaluate the association between diagnostic instability (prospective non-concordance) and various risk factors. The univariate and multivariate Odds Ratios (OR) - with their 95% CIs - for all characteristics in the table are reported. The factors most related to instability, after multiple adjustments in the multivariate model, were male gender (reduced instability, OR 0.75, 95% CI 0.58–0.98), younger age, 11–19 years (increased instability, OR 1.42, 95% CI 0.93–2.17), first hospital admission at university hospital versus a directly managed hospital (reduced instability, OR 0.62, 95% CI 0.42–0.90), psychiatric comorbidity (increased instability, OR 1.35, 95% CI 0.95–1.95), a short length of hospital stay at the baseline (> 15 days vs. < 6 days) (reduced instability, OR 0.66, 95% CI 0.47–0.93), and a short time lag between the first and the last hospitalisation (< 2 years vs. > 6 years) (reduced instability, OR 0.31, 95% CI 0.23–0.43).

TABLE 2 | Descriptive characteristics of the patients enrolled in the study at their first hospitalisation.

Variables	First hospital admission N.	%[a]	Deaths N.	%[b]	At least one subsequent admission N.	%[b]	Mean (SD) subsequent admissions
TOTAL	3871		99	2.6%	1145	29.6%	3.66 (5.08)
Gender							
Male	2002	51.7%	68	3.4%	630	31.5%	3.68 (4.89)
Female	1869	48.3%	31	1.7%	515	27.6%	3.64 (5.32)
Age (median, IQR)	26 (20–31)		29 (22–32)		26 (20–31)		
Citizenship							
Italian	3132	80.9%	89	2.8%	1010	32.2%	3.76 (5.16)
Foreign	739	19.1%	10	1.4%	135	18.3%	2.94 (4.44)
Education							
No education	182	4.7%	5	2.7%	47	25.8%	3.15 (4.43)
Primary school	323	8.3%	8	2.5%	85	26.3%	5.56 (7.73)
Middle school	1.913	49.4%	48	2.5%	583	30.5%	3.78 (5.42)
High school	1318	34.0%	36	2.7%	385	29.2%	3.16 (3.76)
University degree	119	3.1%	2	1.7%	41	34.5%	3.07 (3.51)
Missing	16	0.4%	0	0.0%	4	25.0%	6 (5.94)
Hospital facility							
Main hospital	672	17.4%	20	3.0%	201	29.9%	3.2 (3.79)
University hospital	493	12.7%	7	1.4%	158	32.0%	3.92 (5.16)
Research hospital	132	3.4%	3	2.3%	55	41.7%	4.75 (4.69)
Local hospital	2574	66.5%	69	2.7%	731	28.4%	3.65 (5.39)
Psychiatric comorbidity							
Yes	534	13.8%	27	5.1%	201	37.6%	3.48 (4.91)
No	3337	86.2%	72	2.2%	944	28.3%	4.52 (5.76)
Hospital stay (days, median, IQR)	5 (1–10)		6 (2–14)		8 (3–14)		

[a]Column percentages.
[b]Row percentages.

4 | Discussion

The study found that 30% of patients aged 11–35 with a first psychiatric diagnosis in a specialized ward were readmitted for a psychiatric disorder within 10 years. Among those readmitted, the concordance between the first and last diagnosis was 57.8%, with a weighted kappa of 0.45. Repeated stability (the same diagnosis in 75% of admissions) was 31.2%. Schizophrenia, other functional psychoses, substance use disorders, and eating disorders showed the highest diagnostic concordance. Intermediate concordance was observed for schizophrenia, schizoaffective disorder, bipolar disorder, and personality disorder, while conditions like depression, anxiety, dissociative, and somatic symptom disorders had low concordance. Delusional and schizophreniform disorders had the lowest concordance and kappa values. The study identified key diagnostic changes: Bipolar and personality disorder diagnoses increased over time; Some schizophrenia diagnoses converted into bipolar or personality disorders, and vice versa; Anxiety, dissociative, and somatoform disorders showed general instability, often shifting

to personality disorders, depression, or schizophrenia. Factors linked to diagnostic stability included male gender, older age, initial hospitalization in a University Hospital, absence of co-morbidity, short hospital stays at baseline, and a short interval between the first and last hospitalization.

4.1 | Comparison With Other Studies

The study found a relatively low rate of re-hospitalisation for a psychiatric diagnosis of 29.6%. Some of the patients who were not readmitted to the hospital may have been taken into care by the community services and did not need to be admitted again. This observation is justified by the mental health care in Italy that relies both on hospitals and community centres; in this model, hospital and community care are under unique management, guaranteeing continuity of care. A large international study (Schwartz et al. 2000) found a readmission rate of 32% for Italy but included the first 30 days post discharge while we excluded the rehospitalizations in the 30 days after the first

TABLE 3 | Repeated diagnostic stability by diagnostic category.

	Subjects		Repeated stability*				
Diagnostic categories	N.	%	N.	%	95% CI		
Total	1145	100.0%	357	31.2	28.5	—	33.9
1. Schizophrenia spectrum and other functional psychoses	476	41.6%	113	23.7	19.9	—	27.6
1.1 schizophrenia	157	13.7%	78	49.7	41.9	—	57.5
1.2 schizophreniform disorder	86	7.5%	7	8.1	2.4	—	13.9
1.3 schizoaffective disorder	54	4.7%	8	14.8	5.3	—	24.3
1.4 delusional disorder (paranoia)	22	1.9%	3	13.6	0	—	28.0
1.5 brief psychotic disorder or other non-specified reactive psychoses	19	1.7%	2	10.5	0	—	24.3
1.6 psychoses NOS	27	2.4%	3	11.1	0	—	23.0
2. Bipolar disorders	170	14.8%	85	50.0	42.5	—	57.5
3. Depression	101	8.8%	25	24.8	16.3	—	33.2
4. Anxiety. Dissociative and somatoform disorders	93	8.1%	16	17.2	9.5	—	24.9
6. Personality and behavioural disorders	175	15.3%	58	33.1	26.2	—	40.1
7. Alcoholism and substance use disorders	96	8.4%	45	46.9	36.9	—	56.9
9. Eating disorders	26	2.3%	14	53.8	34.7	—	73.0
12. Behavioural disorders	6	0.5%	1	16.7	0	—	46.5

Note: 'Repeated stability': the same diagnosis is repeated in at least 75% of hospitalisations in the 10 years follow-up:2/2, 2/3, 3/4 of hospitalisations, etc.
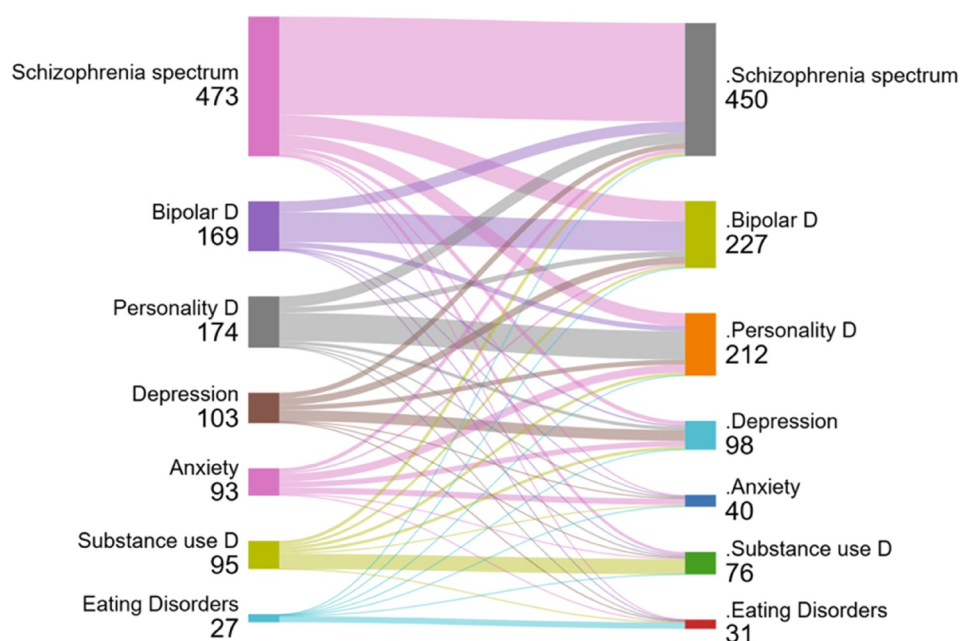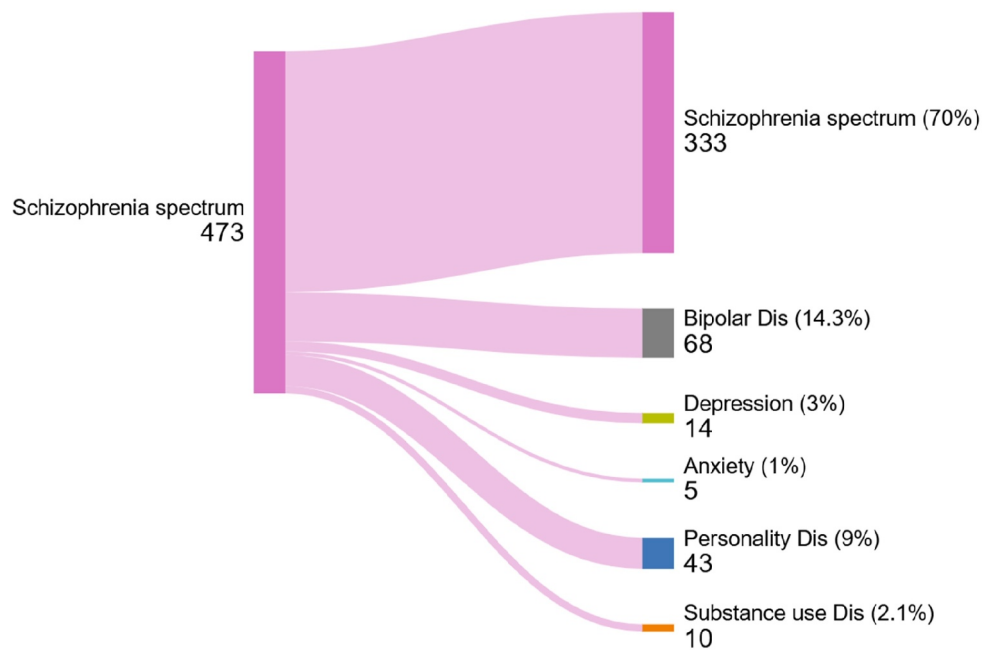

FIGURE 1 | Alluvial graph for all the conditions. Changes of the diagnostic categories from the first to the last hospitalisation. The numbers indicate the number of patients from one specific category to the others.

admission. Our finding agrees with a systematic review of First Episode Psychosis (FEP) remission rates (Katschnig et al. 2019). The overall remission rate in individuals with FEP was 58%, with higher remission rates in studies from more recent years; the recovery rate was 38%, and the authors suggested that remission and recovery rates may be more favourable than previously thought, at least for FEP. A large proportion (2.6%) of those recruited to the study died during the 10-year follow-up, compared with an expected 1.3% in the general population of the same age, based on national statistics. This excess mortality

was exceptionally high for personality disorders (4%). These results confirm the literature data on high mortality in psychiatric patients (Lally et al. 2017). It cannot be excluded, however, that an unknown proportion of the original patients were lost to follow-up due to migration outside the region, especially those of foreign citizenship.

It should be noted that 'schizophrenia spectrum and other functional psychoses' is a broad category, making it logical for concordance to be higher compared to individual subcategories

Schizophrenia spectrum disorder
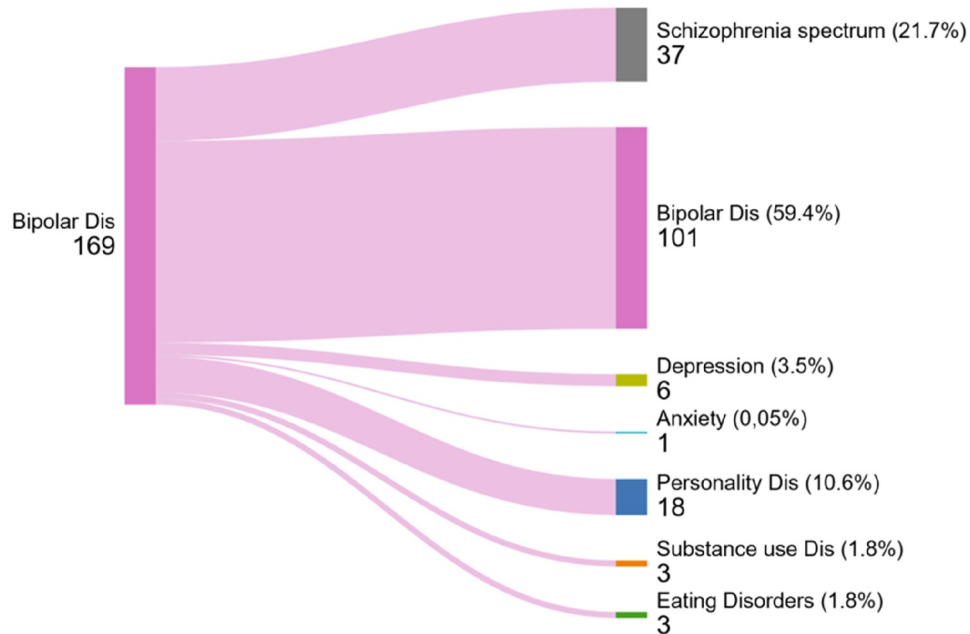

Bipolar Disorder

FIGURE 2 | Alluvial graph for schizophrenia spectrum and bipolar disorders. Changes in the diagnostic categories from the first to the last hospitalisation. The numbers indicate the number of patients from one specific category to the others.

(e.g., schizophrenia, schizophreniform disorder, etc.). In general, the estimated prospective concordance values of this study (schizophrenic spectrum 70%, schizophrenia 58%, bipolar disorders 59.4%) are lower than those reported in systematic reviews for psychosis and bipolar disorder (Oakley et al. 2018; Fusar-Poli et al. 2016; Palomar-Ciria et al. 2019; F. B. Cegla-Schvartzman et al. 2019). In the case of functional psychoses, Fusar-Poli reported a high concordance rate between the first and last diagnosis in the following instances: schizophrenia 90%, affective spectrum psychosis 84%, and schizoaffective disorder

TABLE 4 | 'Prospective concordance' and 'Retrospective concordance' of the diagnosis at the first and the last hospitalisation.

Diagnostic categories	Prospective concordance			Retrospective concordance				
	First hospital admissions	Concordant	%	Last admissions	Concordant	%	Kappa	95% CI
Total	1145	662	57.8	1145	662	57.8	0.47	0.42–0.52
1. Schizophrenia spectrum and other functional psychoses	476	333	70.0	452	333	74	0.53	0.47–0.58
1.1 schizophrenia	157	91	58.0	205	91	44.4	0.41	0.34–0.48
1.2 schizophreniform disorder	86	11	12.8	72	11	15.3	0.08	0.0–0.15
1.3 schizoaffective disorder	54	19	35.2	35	19	54.3	0.41	0.27–0.54
1.4 delusional disorder (paranoia)	22	5	22.7	30	5	16.7	0.17	0.03–0.32
1.5 brief psychotic disorder or other non-specified reactive psychoses	19	1	5.3	7	1	14.3	0.03	0.0–0.12
1.6 psychosis NOS	27	7	25.9	51	7	13.7	0.16	0.04–0.28
2. Bipolar disorders	170	101	59.4	227	101	44.5	0.4	0.33–0.47
3. Depression	101	37	36.6	98	37	37.8	0.31	0.22–0.40
4. Anxiety, dissociative and somatoform disorders	93	21	22.6	40	21	52.5	0.28	0.18–0.38
6. Personality and behavioural disorders	175	96	54.9	216	96	44.4	0.38	0.31–0.45
7. Alcoholism and substance use disorders	96	52	54.2	76	52	68.4	0.57	0.48–0.66
9. Eating disorders	26	21	80.8	29	21	72.4	0.76	0.63–0.89
12. Behavioural disorders	6	1	54.9	1	1	100.0	0.29	0.0–0.72

72%. In the systematic reviews on bipolar disorder by Cano-Ruiz et al. (2020) and F. B. Cegla-Schvartzman et al. (2019), the prospective concordance was 77% and 60%, respectively. However, our investigation differs from other studies, being a population-based cohort and not limited to specific diagnostic facilities. In contrast, many papers in the meta-analyses were related to clinical units, often research units. For comparison, see the detailed study by Cano-Ruiz et al. (2020), who carried out structured interviews at different times in a research centre and obtained high stability values. In addition, the time interval considered in the literature was a few years, whereas it was 10 years in our case.

We found low stability for some conditions. The average stability found for schizoaffective disorder is consistent with the suggestions from Wood et al. (2021), who see schizoaffective disorder (SAD) as a chronic psychotic disorder often used as a diagnosis for individuals who have a mix of mood and psychotic symptoms whose diagnosis is uncertain. This may represent a midpoint on a continuum between bipolar disorder and schizophrenia. The same is true for NOS psychosis. Miller and Black (2019) explored some aspects related to 'unspecified psychosis', commonly used if there is inadequate information for diagnosing a specific psychotic disorder, finding lower diagnostic stability of unspecified psychosis compared to other psychotic disorders. Commonly, NOS psychosis is associated with psychiatric comorbidities, and further studies are needed to clarify the role of comorbidities.

Depression has a lower stability than other conditions, as indicated in population-based studies, such as the Danish National Register study (Taş et al. 2019). In this study, in about 56% of patients, the initial diagnosis of depressive disorder changed during follow-up, mainly to the schizophrenic spectrum (16%) but also to personality disorders (9%), neurotic, stress-related and somatoform disorders (8%) and bipolar disorder (8%). These results are very similar to those reported in our study and indicate that the diagnosis of depressive disorders has low stability over time in clinical practice.

Anxiety disorders represent another case of low stability. Kessing (2005) assessed the longitudinal stability of anxiety disorders in 447 subjects with this diagnosis at 2, 4, and 6-year follow-ups. Rates of change were high across all anxiety disorder diagnoses, ranging from 21.1% for a social anxiety disorder to 46.3% for panic disorder with agoraphobia at 6 years of follow-up. The data indicate that anxiety disorder diagnoses are unstable over time as the symptoms are transdiagnostic dimensions, and a misdiagnosis is likely.

Among the factors responsible for the increased stability we found: male gender, a frequent finding in the literature (Hovenkamp-Hermelink et al. 2016), adulthood when the diagnostic characteristics are more stabilised, the first diagnosis made in a University Hospital as probably linked to the greater use of standardised instruments and appropriate training in university facilities. Finally, it is clear that a longer

TABLE 5 | Results of the univariate and multivariate analyses of the factors related to diagnostic instability (OR, Odds ratio, 95% Confidence Intervals, CI).

Variable	Subjects N	Univariate analysis OR	95% CI		p-value	Multivariate analysis OR	95% CI		p-value
Characteristics at the first hospitalisation									
Gender									
Males	630	0.82	0.64	1.06	0.13	0.75	0.58	0.98	0.04
Females	515	1.00	—	—	—	1.00	—	—	—
Age groups (years)									
11–19	235	1.22	0.85	1.74	0.29	1.42	0.93	2.17	0.11
20–24	268	0.90	0.65	1.26	0.54	0.95	0.67	1.35	0.77
25–29	257	1.06	0.76	1.50	0.73	1.03	0.72	1.47	0.89
30–35	385	1.00	—	—	—	1.00	—	—	—
Citizenship									
Italian	1010	1.00	—	—	—	1.00	—	—	—
Foreign	135	0.89	0.60	1.30	0.53	1.01	0.67	1.50	0.98
Education									
No education	47	0.62	0.33	1.17	0.14	0.53	0.27	1.06	0.07
Primary school	85	0.77	0.48	1.26	0.30	0.86	0.52	1.42	0.56
Middle school	583	1.00	—	—	—	1.00	—	—	—
High school	385	1.20	0.89	1.60	0.23	1.18	0.87	1.60	0.28
University degree	41	0.61	0.32	1.16	0.13	0.71	0.36	1.39	0.32
Hospital facility									
Main hospital	201	1.32	0.92	1.88	0.13	1.31	0.91	1.90	0.14
University hospital	158	0.62	0.43	0.88	0.01	0.62	0.42	0.90	0.01
Research hospital	55	1.07	0.53	2.16	0.84	1.59	0.72	3.53	0.25
Local hospital	731	1.00	—	—	—	1.00	—	—	—
Psychiatric comorbidity									
Yes	201	1.27	0.90	1.78	0.18	1.35	0.95	1.94	0.10
No	944	1.00	—	—	—	1.00	—	—	—
Duration of hospital stay (days)									
< 6	407	1.04	0.78	1.40	0.77	1.06	0.78	1.43	0.72
6–15	491	1.00	—	—	—	1.00	—	—	—
> 15	247	0.72	0.52	0.99	0.05	0.66	0.47	0.93	0.02
Lag time first and last hospitalisation (years)									
< 2	334	0.33	0.24	0.45	< 0.0001	0.31	0.23	0.43	< 0.0001
2–5	363	0.64	0.46	0.88	0.01	0.63	0.45	0.86	0.00
> 6	448	1.00	—	—	—	1.00	—	—	—

stay in the first admission is associated with a disorder of higher severity and, therefore, more stability. We found that the shorter the lag time between the first and the last observation, the higher the stability. This finding justifies the different results of our study in comparison with other investigations with a shorter follow-up.

The interpretation of observational data is always complex, making it difficult to fully understand the reasons behind changes in psychiatric diagnoses over time. It is evident that there are no biomarkers for functional psychiatric disorders, and studies on the cross-sectional interrater reliability of psychiatric diagnoses have yielded relatively poor results (Laursen et al. 2020). Some psychiatric diagnoses can only be established by assessing the longitudinal course—this is particularly true for bipolar disorder and, to some extent, personality disorders—which may explain the increase in these diagnoses over time.

4.2 | Strengths and Limitations of the Study

This study has several strengths that should be emphasised. It is a population-based study that did not select patients and reflects the conditions of the diagnoses in an entire Italian region. It was not conducted in a single diagnostic centre, like many of the published studies, but it included all the hospital centres in the Lazio Region during the period considered. The data are based on consolidated information systems, and the study is very similar to longitudinal population studies based on registers, as in Denmark and Spain. The results are, therefore, easily generalisable.

The study selected adolescents and young adults to better focus on first diagnoses and to avoid making comparisons of populations with a long previous psychiatric history. The follow-up was 10 years, long enough to assess diagnostic stability, while many available studies limited the follow-up to a few years. The analysis methodology was standardised, and various stability indices (prospective, retrospective, repeated) and the kappa index were calculated. Diagnostic trajectories were identified and illustrated for each diagnostic category using the 'alluvial graph' methodology.

The study has limitations that need to be considered. A few subjects were observed for some conditions that did not allow a complete assessment. The entire diagnostic trajectory was not assessed, only the first and last admission, limiting the descriptive possibility. Additionally, we could not identify the doctors who made the diagnoses in different time periods. It is evident that if a patient sees the same doctor, the likelihood of the diagnosis being confirmed increases. Given the complexity of the study, little attention was paid to comorbidity diagnoses, which can, in many ways, be crucial in conditioning stability. Limited attention was also paid to factors associated with instability. Future studies must better describe the personal, social, and structural factors that are most responsible for instability and trajectory change. Obviously, a significant limitation is that only categorical diagnoses and not dimensional approaches to the individual's psychopathology were considered.

Finally, it should be recognized that the commonly used categorical system for recording data in public health portrays only one aspect of the way the diagnosis is made. There are inconsistencies between international diagnostic classification manuals, clinicians' work, and database registration needs. In Italy, for example, psychiatrists commonly use the DSM-5 to guide their diagnoses. At the same time, the data recording system of hospitalisations at the national level is still based on the ICD-9 revision, which dates back to the 70s. In contrast, nowadays, the ICD-11 revision is available.

4.3 | Implications of the Results

This work aims to contribute to population studies on the stability of psychiatric diagnosis and its trajectories. Insights into the various disorders' diagnostic conversions and developmental trajectories are helpful for clinical risk assessment, improving prevention and early intervention protocols in adolescents and young adults and treatment in the transition from adolescence to adulthood.

It is commonly accepted at an epidemiological level that the more stable the diagnosis, the more its construct describes a discrete psychopathological condition. On the other hand, diagnostic variability may reflect clinical variations in the patient's course and methodological artefacts such as variations in information, poor reliability in measurement methods, the interviewer's judgement, or the diagnostic reference manual. Using administrative data, as we have done, does not permit us to disentangle the single factors underlying the diagnostic instability of a psychiatric disorder.

Diagnostic stability over the years is a significant factor in the definition of homogeneous diagnostic groups; the homogeneity and reliability of a diagnostic group is the *conditio sinae qua non* for research into the classification, biology, genetics, prognosis, and treatment of mental disorders. The methodologies adopted for the stability studies are rarely superimposable, and often, the type of data collection and the type of data differ substantially. The methods and settings of the diagnosis are important (one or more assessors, structured interview or clinical judgement, the different diagnostic manual of reference, etc.) as well as the number of diagnoses for the same patient, the period between successive diagnostic moments, the centre in which the patient is admitted or contacted.

5 | Conclusions

Our population-based approach with extended follow-up estimated overall diagnostic stability lower than that reported in the literature. Several factors were related to diagnostic stability, and we found various trajectories that helped to characterise different evolutionary patterns. From a clinical point of view, the low predictive value of the first psychiatric diagnosis suggests a cautious attitude in communicating a categorical 'conclusive' diagnosis to a young adult.

Author Contributions

Anna Forastiere: conceptualization, data curation, methodology, visualization, writing – original draft, writing – reviews and editing. **Enrico Calandrini:** data curation, formal analysis, writing – original draft. **Giulia Cesaroni:** conceptualization, data curation, formal analysis, writing – original draft. **Anna Maria Bargagli:** writing – original draft. **Nera Agabiti:** conceptualization, data curation, methodology, writing – original draft, writing – reviews and editing.

according to the national privacy law (national legislative decree on privacy policy n. 196/30 June 2003). Individuals cannot be identified directly or through identifiers, and results are shown in aggregate form.

Conflicts of Interest

The authors declare no conflicts of interest.

Data Availability Statement

Data related to the findings reported in our manuscript are available to all interested researchers upon request because of stringent legal restrictions regarding privacy policy on personal information in Italy (national legislative decree on privacy policy n. 196/30 June 2003). For this reason, our dataset cannot be made available on public data deposition. All interested researchers can contact the following persons to request the data: Nera Agabiti, Department of Epidemiology, Lazio Regional Health Service, Rome, Italy, E-mail: n.agabiti@deplazio.it.

Permission to Reproduce Material From Other Sources

The authors have nothing to report.

References

Aboraya, A., E. Rankin, C. France, A. El-Missiry, and C. John. January 2006. "The Reliability of Psychiatric Diagnosis Revisited: The Clinician's Guide to Improve the Reliability of Psychiatric Diagnosis." *Psychiatry* 3, no. 1: 41–50: PMID: 21103149; PMCID: PMC2990547.

Baca-Garcia, E., M. M. Perez-Rodriguez, I. Basurte-Villamor, et al. June 2007. "Diagnostic Stability and Evolution of Bipolar Disorder in Clinical Practice: A Prospective Cohort Study." *Acta Psychiatrica Scandinavica* 115, no. 6: 473–480: PMID: 17498159. https://doi.org/10.1111/j.1600-0447.2006.00984.x.

Cano-Ruiz, P., P. Sanmartin-Salinas, A. Gómez-Peinado, C. Calero-Mora, and L. Gutiérrez-Rojas. January 2020. "Diagnostic Stability in Bipolar Disorder: A Systematic Review." *Actas Españolas de Psiquiatria* 48, no. 1: 28–35: Epub 2020 Jan 1. PMID: 32297649.

Cawkwell, P. B., K. W. Bolton, R. Karmacharya, D. Öngür, and A. K. Shinn. December 2020. "Two-Year Diagnostic Stability in a Real-World Sample of Individuals With Early Psychosis." *Early Interv Psychiatry* 14, no. 6: 751–754: Epub 2020 Feb 10. PMID: 32043313; PMCID: PMC7774998. https://doi.org/10.1111/eip.12930.

Cegla-Schvartzman, F., S. Ovejero, J. López-Castroma, et al. September 2021. "Diagnostic Stability in Bipolar Disorder: A Follow-Up Study in 130,000 Patient-Years." *Journal of Clinical Psychiatry* 82, no. 6: 20m13764: PMID: 34529901. https://doi.org/10.4088/JCP.20m13764.

Cegla-Schvartzman, F. B., S. Ovejero, J. López-Castroman, and E. Baca-García. January, February 2019. "Diagnostic Stability in Bipolar Disorder: A Narrative Review." *Harvard Review of Psychiatry* 27, no. 1: 3–14: PMID: 30358663. https://doi.org/10.1097/HRP.0000000000000187.

Fichter, M. M., and N. Quadflieg. November 2007. "Long-term Stability of Eating Disorder Diagnoses."supplement, *International Journal of Eating Disorders* 40, no. Sl: S61–S66: PMID: 17702021. https://doi.org/10.1002/eat.20443.

Fusar-Poli, P., M. Cappucciati, G. Rutigliano, et al. November 2016. "Diagnostic Stability of ICD/DSM First Episode Psychosis Diagnoses: Meta-Analysis." *Schizophrenia Bulletin* 42, no. 6: 1395–1406: Epub 2016 Mar 15. PMID: 26980142; PMCID: PMC5049518. https://doi.org/10.1093/schbul/sbw020.

Høj Jørgensen, T. S., M. Osler, M. B. Jorgensen, and A. Jorgensen. January 2023. "Mapping Diagnostic Trajectories From the First Hospital Diagnosis of a Psychiatric Disorder: A Danish Nationwide Cohort Study Using Sequence Analysis." *Lancet Psychiatry* 10, no. 1: 12–20: Epub 2022 Nov 27. PMID: 36450298. https://doi.org/10.1016/S2215-0366(22)00367-4.

Hovenkamp-Hermelink, J. H., H. Riese, D. C. van der Veen, N. M. Batelaan, B. W. Penninx, and R. A. Schoevers. January 2016. "Low Stability of Diagnostic Classifications of Anxiety Disorders over Time: A Six-Year Follow-Up of the NESDA Study." *Journal of Affective Disorders* 190: 310–315: Epub 2015 Oct 28. PMID: 26544613. https://doi.org/10.1016/j.jad.2015.10.035.

Katschnig, H., C. Straßmayr, F. Endel, et al. and CEPHOS-LINK study group. November 2019. "Using National Electronic Health Care Registries for Comparing the Risk of Psychiatric Re-Hospitalisation in Six European Countries: Opportunities and Limitations." *Health Policy* 123, no. 11: 1028–1035: Epub 2019 Jul 20. PMID: 31405616. https://doi.org/10.1016/j.healthpol.2019.07.006.

Kessing, L. V. January, February 2005. "Diagnostic Stability in Depressive Disorder as According to ICD-10 in Clinical Practice." *Psychopathology* 38, no. 1: 32–37: Epub 2005 Feb 15. PMID: 15714010. https://doi.org/10.1159/000083968.

Kjaer, J. N., R. Biskin, C. Vestergaard, L. N. Gustafsson, and P. Munk-Jørgensen. August 2016. "The Clinical Trajectory of Patients With Borderline Personality Disorder." *Personality and Mental Health* 10, no. 3: 181–190: Epub 2016 Apr 29. PMID: 27126217. https://doi.org/10.1002/pmh.1337.

Lally, J., O. Ajnakina, B. Stubbs, et al. December 2017. "Remission and Recovery From First-Episode Psychosis in Adults: Systematic Review and Meta-Analysis of Long-Term Outcome Studies." *British Journal of Psychiatry* 211, no. 6: 350–358: Epub 2017 Oct 5. PMID: 28982659. https://doi.org/10.1192/bjp.bp.117.201475.

Laursen, M. F., R. W. Licht, C. U. Correll, et al. May 2020. "Diagnostic Stability in Children and Adolescents With Bipolar Disorder, a Nationwide Register-Based Study." *Int J Bipolar Disord* 8, no. 1: 14: PMID: 32372109; PMCID: PMC7200954. https://doi.org/10.1186/s40345-020-0179-3.

Miller, J. N., and D. W. Black. February 2019. "Schizoaffective Disorder: A Review." *Annals of Clinical Psychiatry* 31, no. 1: 47–53: PMID: 30699217.

Milos, G., A. Spindler, U. Schnyder, and C. G. Fairburn. December 2005. "Instability of Eating Disorder Diagnoses: Prospective Study." *British Journal of Psychiatry* 187, no. 6: 573–578: PMID: 16319411; PMCID: PMC2710504. https://doi.org/10.1192/bjp.187.6.573.

Oakley, P., S. Kisely, A. Baxter, et al. July 2018. "Increased Mortality Among People With Schizophrenia and Other Non-Affective Psychotic Disorders in the Community: A Systematic Review and Meta-Analysis." *Journal of Psychiatric Research* 102: 245–253: Epub 2018 Apr 26. PMID: 29723811. https://doi.org/10.1016/j.jpsychires.2018.04.019.

Palomar-Ciria, N., F. Cegla-Schvartzman, J. D. Lopez-Morinigo, H. J. Bello, S. Ovejero, and E. Baca-García. 2019. "Diagnostic Stability of Schizophrenia: A Systematic Review." *Psychiatry Research* 279: 306–314. https://doi.org/10.1016/j.psychres.2019.04.020.

Peralta, D., L. Janda, E. García de Jalón, et al. and SEGPEPs Group. November 2023. "Long-Term Diagnostic Stability, Predictors of Diagnostic Change, and Time Until Diagnostic Change of First-Episode Psychosis: A 21-Year Follow-Up Study." *Psychological Medicine* 21: 1–10: Epub ahead of print. PMID: 37987188. https://doi.org/10.1017/S0033291723003173.

Schwartz, J. E., S. Fennig, M. Tanenberg-Karant, et al. June 2000. "Congruence of Diagnoses 2 Years After a First-Admission Diagnosis of Psychosis." *Archives of General Psychiatry* 57, no. 6: 593–600: PMID: 10839338. https://doi.org/10.1001/archpsyc.57.6.593.

Taş, H. İ, M. Çelik, and K. Altinbaş. March 2019. "Evaluation of Four-Year Stability of Unspecified Psychosis." *Noro Psikiyatr Ars* 56, no. 1: 47–51: Epub 2018 Sep 17. PMID: 30911237; PMCID: PMC6427075. https://doi.org/10.29399/npa.22903.

Wood, A. J., A. R. Carroll, A. K. Shinn, D. Ongur, and K. E. Lewandowski. October 2021. "Diagnostic Stability of Primary Psychotic

Disorders in a Research Sample." *Frontiers in Psychiatry* 12: 734272: PMID: 34777044; PMCID: PMC8580873. https://doi.org/10.3389/fpsyt.2021.734272.

Supporting Information

Additional supporting information can be found online in the Supporting Information section.