Diagnostic stability and evolution of bipolar disorder in clinical practice: a prospective cohort study


Objective: To evaluate the long-term stability of International Classification of Diseases-10th revision bipolar affective disorder (BD) in multiple settings.

Method: A total of 34,368 patients received psychiatric care in the catchment area of a Spanish hospital (1992–2004). The analyzed sample included patients aged ≥18 years who were assessed on ≥10 occasions and received a diagnosis of BD at least once (n = 1153; 71,543 assessments). Prospective and retrospective consistencies and the proportion of subjects who received a BD diagnosis in ≥75% of assessments were calculated. Factors related to diagnostic shift were analyzed with traditional statistical methods and Markov’s models.

Results: Thirty per cent of patients received a BD diagnosis in the first assessment and 38% in the last assessment. Prospective and retrospective consistencies were 49% and 38%. Twenty-three per cent of patients received a BD diagnosis during ≥75% of the assessments.

Conclusion: There was a high prevalence of misdiagnosis and diagnostic shift from other psychiatric disorders to BD. Temporal consistency was lower than in other studies.

Significant outcomes

• The temporal consistency of bipolar affective disorder (BD) was lower than that found in other studies.
• There was a high prevalence of misdiagnosis and diagnostic shift from other psychiatric disorders to BD.
• Patients with a stable diagnosis of BD (‘stable BD’ group) presented some diagnostic fluctuation.

Key words: bipolar disorder; diagnosis; International Classification of Diseases; reproducibility of results; classification


1Department of Psychiatry, Fundacion Jimenez Diaz University Hospital, Autonoma University of Madrid, Madrid, Spain, 2Department of Psychiatry, Ramón y Cajal Hospital, Madrid, Spain, 3Mental Health Center of Centro District, Madrid, Spain, 4Mental Health Center of Arganzuela District, Universidad Complutense de Madrid, Madrid, Spain, 5University of Alcala, Madrid, Spain, 6Department of Signal Theory and Communications, Universidad Carlos III, Madrid, Spain, 7New York State Psychiatric Institute/Columbia University in New York, New York, USA, 8Mental Health Research Center at Eastern State Hospital in Lexington, Kentucky, USA

Accepted for publication November 28, 2006
Limitations

- The limitations are those inherent to a naturalistic study, performed in real world conditions. Structured or semi-structured clinical interviews were not used in this study.
- The clinicians who assigned the diagnoses were not specifically trained to increase inter-rater reliability.
- The prevalence of BD in this psychiatric sample is lower than in other studies.

Introduction

Bipolar affective disorder (BD) is considered a lifelong illness. In theory, the diagnosis of BD, once established, should be stable over time (1). However, this may not always be the case in clinical practice (2). Several factors may affect the diagnostic stability of BD (3): i) manifestations of BD might change over time and overlap with those of other disorders (4); ii) comorbid conditions may alter the clinical appearance or course of BD (5, 6); iii) diagnoses by different observers may be inconsistent (7); and iv) sociodemographic factors may alter the course of the illness, its presenting symptoms, or their perception by clinicians. Therefore, the analysis of factors that influence the diagnostic stability of BD and the likelihood of a diagnostic switch from another disorder to BD and vice versa is relevant for psychiatric research. The stability of BD and related factors has been evaluated in several studies (1–3, 8–15). These studies usually had a small number of evaluations – two or three in most of them (1, 9, 10, 12) – and the follow-up period is <3 years in most of them (1, 9, 12) with some exceptions (2, 3, 10, 11, 13).

Aims of the study

The aim of the present study is an ecologic evaluation of the long-term stability and evolution of the International Classification of Diseases-10th revision (ICD-10) diagnosis of BD in multiple clinical settings in real world conditions.

Material and methods

Patients

The Fundacion Jimenez Diaz, a general hospital in Madrid, Spain, which is a part of the Spanish National Health Services, provides free medical coverage to a catchment area of 280 000 people. From January 1st, 1992 to December 31st, 2004, at this catchment area, 34 368 patients received psychiatric care. There were 449 317 psychiatric assessments in three clinical settings, including visits to out-patient psychiatric facilities (438 622), emergency visits (9101) and admissions to the psychiatric brief hospitalization unit (1594). A subsample was selected (n = 10 025) of patients aged 18 and over who were assessed on at least 10 occasions during the study period. A total of 1153 patients received a diagnosis of BD, according to ICD-10, (16) during at least one evaluation. These 1153 patients had 71 543 psychiatric consultations. The mean duration of follow-up for the patients was 6.2 (SD 3.6) years and the median of visits was 34.

Participants (n = 1153) were assessed in three different clinical settings: in-patient unit (psychiatric brief hospitalization unit, 2000–2004), psychiatric emergency room (2000–2004) and out-patient psychiatric facilities (mental health care centers) within the hospital catchment area of the Fundacion Jimenez Diaz (1992–2004).

Diagnostic procedure

In all settings, diagnoses were assigned after reviewing all available information, including data from medical records, other research assessments, and clinical interviews with the patient and relatives. The psychiatrists who assigned the clinical diagnoses in any of these three settings were not aware of the study in process. The ICD is the diagnostic system of choice in Spain, and psychiatry residents are trained to use ICD-10. Most of the diagnostic psychiatrists were trained psychiatrists with many years of experience, whereas others were supervised residents still in training. Of course, we cannot rule out that some of the diagnostic psychiatrists may favor DSM and use ICD only because they have to.

Diagnostic groups included in the statistical analysis

In addition to BD (ICD-10 F31), we included all blocks from Chapter V of the ICD-10 [Mental and Behavioral Disorders (F00–F99)] (two digit categories, Fx) in the analysis after excluding Disorders...
of Psychological Development (F80–F89). We also included all three (Fxx.) and four digit (Fxx.x) categories with prevalences ≥ 1% in the whole sample.

Data extraction and analysis

**Diagnostic stability** Diagnostic stability through all the evaluations was first calculated according to Schwartz et al. and Baca-García et al. (1, 17) with traditional statistical methods using version 13.0 of spss (SPSS Inc., Chicago, IL, USA). Three measures of stability are presented: i) prospective consistency (the proportion of individuals with a diagnosis of BD at the first evaluation who retained the same diagnosis at the last evaluation, conceptually similar to positive predictive value if the last diagnosis were the gold standard); ii) retrospective consistency (the proportion of individuals with a diagnosis of BD assigned at the last evaluation who had received the same diagnosis at the first evaluation, conceptually similar to sensitivity); and iii) the proportion of subjects who received a diagnosis of BD in at least 75% of the evaluations. The agreement between diagnoses at the first and the last evaluation was calculated by the kappa coefficient, which measures the agreement correcting the effect of chance. We performed this analysis with the joint data from the three clinical settings, to reflect the evolution of diagnoses through the clinical process.

We performed a second statistical analysis of diagnostic stability using Markov’s Models, (18, 19). A first order Markov model represents a process in which evolution only depends on the present state. In other words, all the information from the past that is useful for predicting the future is condensed in each state. Therefore, if the process in state A, it has a Pb probability of moving to state B and a Pc probability of moving to state C. Pb and Pc remain the same regardless of the number of states the process has been through before reaching state A. The values of the transitions from one state to another can be represented as an image in which the color of the pixels reflects the probability of each transition.

**Diagnostic changes** To study the diagnostic switch between BD and other disorders, we analyzed three sets of patients, each of which consisted of two non-overlapping groups of subjects: i) first set: 342 subjects who received a diagnosis of BD at the first evaluation (first diagnosis BD) and 811 subjects who were given any other diagnosis (first diagnosis not BD); ii) second set: 443 subjects who received a diagnosis of BD at the last evaluation (last diagnosis BD) and 710 subjects who did not receive a diagnosis of BD at the last evaluation (last diagnosis not BD); and iii) third set: 266 subjects who received the diagnosis of BD in at least 75% of the evaluations (stable BD) and 887 who did not receive the diagnosis of BD in ≥ 75% of the evaluations (non-stable BD).

**Results**

**Stability of BD**

From the sample with ≥ 10 assessments (n = 1153), 30% (n = 342/1153) received a diagnosis of BD at the first evaluation (first diagnosis BD) and 38% (n = 443/1153) at the last evaluation (last diagnosis BD). Kappa first vs. last evaluation was 0.4 (P < 0.001). Prospective consistency was 49% for BD. Retrospective consistency was 38%.

The ‘stable BD’ patients were 23% of subjects (n = 266/1153) who received the diagnosis of BD during at least 75% of the evaluations. Within this ‘stable BD’ group, 70% (n = 185/266) received the diagnosis of BD at the first psychiatric assessment, providing a kappa of 0.5 (P < 0.001). Within the ‘non-stable BD’ patients, 18% (n = 157/887) received the diagnosis of BD at the first psychiatric evaluation. Within the ‘stable BD’ patients, 79% (n = 211/266) received the diagnosis of BD at the last psychiatric evaluation, providing a kappa of 0.4 (P < 0.001). Within the ‘non-stable BD’ patients, 26% (n = 232/887) received the diagnosis of BD at the last psychiatric evaluation.

In the whole sample (n = 1153), the mean number of evaluations from the first treatment contact within the psychiatric service system to the first time they were diagnosed with BD was 17.9 (31st percentile of the total number of assessments in the 1153-patient sample). The median was 6.0 assessments (18th percentile of the total number of assessments in the 1153-patient sample). The proportion of patients who did not receive the diagnosis of BD until the last evaluation was 2% (n = 20/1153).

Among the 266 ‘stable BD’ patients, the mean number of assessments from the first treatment contact within the psychiatric service system to the first time the patient was diagnosed with BD was 2.1 (seventh percentile of the total number of evaluations). The median was 1.0 assessment (fifth percentile). All 266 ‘stable BD’ patients had received the diagnosis of BD at the 33rd percentile of the total number of evaluations.

Among the 887 ‘non-stable BD’ patients, the mean number of evaluations from the first
treatment contact within the psychiatric service system to the first time the patient was diagnosed with BD was 22.6 (38th percentile of the total number of evaluations). The median was 9.0 (31st percentile). The proportion of patients who received the diagnosis of BD at the last evaluation was 2% (n = 20/887).

The percentile of the total number of evaluations at which the patients were first diagnosed with BD was significantly different in the ‘stable BD’ and ‘non-stable BD’ groups (Mann–Whitney’s U = 43231.5; P < 0.001).

Factors related to diagnostic stability According to the logistic regression model, four variables (gender, age ≥40 years, number of psychiatric assessments, and treatment at out-patient mental health centers) were related to diagnostic stability of BD. However, no association could be found between the following variables and diagnostic stability: marital status, educational level, and socioeconomic level.

Diagnostic frequencies

Diagnostic frequencies for the most common diagnostic groups are presented in Table 1.

The most frequent diagnoses (≥5%) during the study period in the 71 543 assessments of the 1153-patient sample were: paranoid schizophrenia (F20.0; 12%, 8447/71 543); residual schizophrenia (F20.5; 11%, 7576/71 543); BD, current episode mild or moderate depression (F31.3; 11%, 8131/71 543); BD, current episode manic without psychotic symptoms (F31.1; 10%, 7113/71 543); dysthymia (F34.1; 9%, 6314/71 543); and major depressive disorder, recurrent (F33; 7%, 4855/71 543).

The most frequent diagnoses during the study period among all 13 148 assessments of the 266 ‘stable BD’ patients were: BD, current episode mild or moderate depression (F31.3; 30%, 3896/13 148 consultations); BD, current episode manic without psychotic symptoms (F31.1; 29%, 3767/13 148); other BD (F31.8; 15%, 2017/13 148); BD, most recent episode unspecified (F31.9; 9%, 1174/13 148); and BD, most recent episode mixed (F31.6; 7%, 981/13 148).

The most frequent diagnoses during the study period among all 58 395 assessments of the 877 ‘non-stable BD’ patients were: paranoid schizophrenia (F20.0; 14%, 8409/58 395); residual schizophrenia (F20.5; 13%, 7543/58 395); dysthymia (F34.1; 11%, 6205/58 395); major depressive disorder, recurrent (F33; 8%, 4691/58 395); BD, current episode mild or moderate depression (F31.3; 7%, 4235/58 395); and BD, current episode manic without psychotic symptoms (F31.1; 6%, 3346/58 395).

The most frequent diagnoses during the study period among all 17 122 assessments of the 342 ‘first diagnosis BD’ patients were: BD, current episode mild or moderate depression (F31.3; 20%, 3425/17 122); BD, current episode manic without psychotic symptoms (F31.1; 20%, 3430/17 122); other BD (F31.8; 12%, 2100/17 122); dysthymia (F34.1; 6%, 1069/17 122); BD, most recent episode unspecified (F31.9; 6%, 984/17 122); and BD, most recent episode mixed (F31.6; 5%, 909/17 122).

The most frequent diagnoses during the study period among all 54 421 assessments of the 811 ‘first diagnosis not BD’ patients were: paranoid

---

Table 1. Diagnostic frequencies of the most common diagnoses

<table>
<thead>
<tr>
<th>Diagnosis</th>
<th>Whole sample [% (n)]</th>
<th>Stable BD [% (n)]</th>
<th>Non-stable BD [% (n)]</th>
<th>Last diagnosis BD [% (n)]</th>
<th>Last diagnosis not BD [% (n)]</th>
<th>First diagnosis BD [% (n)]</th>
<th>First diagnosis not BD [% (n)]</th>
</tr>
</thead>
<tbody>
<tr>
<td>F20.0 paranoid schizophrenia</td>
<td>11.8 (8447)</td>
<td>0.3 (38)</td>
<td>14.4 (8409)</td>
<td>2.4 (525)</td>
<td>16.0 (7922)</td>
<td>3.3 (561)</td>
<td>14.5 (7886)</td>
</tr>
<tr>
<td>F20.5 residual schizophrenia</td>
<td>10.6 (7576)</td>
<td>0.3 (33)</td>
<td>12.9 (7543)</td>
<td>2.0 (434)</td>
<td>20.0 (7472)</td>
<td>3.3 (558)</td>
<td>12.9 (7018)</td>
</tr>
<tr>
<td>F31.1 BD, current episode manic without psychotic symptoms</td>
<td>9.9 (7133)</td>
<td>28.7 (3376)</td>
<td>5.7 (3346)</td>
<td>20.6 (4969)</td>
<td>5.1 (2524)</td>
<td>20.0 (3430)</td>
<td>6.9 (3683)</td>
</tr>
<tr>
<td>F31.3 BD, current episode mild or moderate depression</td>
<td>11.4 (8131)</td>
<td>28.6 (3986)</td>
<td>7.3 (4235)</td>
<td>21.8 (4817)</td>
<td>6.7 (3314)</td>
<td>20.0 (3425)</td>
<td>8.7 (4706)</td>
</tr>
<tr>
<td>F31.6 BD, most recent episode mixed</td>
<td>2.2 (1551)</td>
<td>7.5 (981)</td>
<td>1.0 (570)</td>
<td>4.6 (1014)</td>
<td>1.1 (537)</td>
<td>5.3 (909)</td>
<td>1.2 (642)</td>
</tr>
<tr>
<td>F31.8 other BD</td>
<td>4.6 (3288)</td>
<td>15.3 (2017)</td>
<td>2.2 (1271)</td>
<td>10.0 (2060)</td>
<td>2.2 (1082)</td>
<td>12.3 (2100)</td>
<td>2.2 (1188)</td>
</tr>
<tr>
<td>F31.9 BD, most recent episode unspecified</td>
<td>3.5 (2518)</td>
<td>8.9 (1174)</td>
<td>2.3 (1344)</td>
<td>6.8 (1504)</td>
<td>2.1 (1014)</td>
<td>5.8 (984)</td>
<td>2.8 (1534)</td>
</tr>
<tr>
<td>F33. major depressive disorder, recurrent</td>
<td>6.7 (4855)</td>
<td>1.0 (136)</td>
<td>8.0 (4691)</td>
<td>7.1 (1577)</td>
<td>6.6 (3250)</td>
<td>3.3 (557)</td>
<td>7.9 (4270)</td>
</tr>
<tr>
<td>F34.1 dysthymia</td>
<td>8.8 (6314)</td>
<td>0.8 (109)</td>
<td>10.6 (6205)</td>
<td>5.8 (1278)</td>
<td>10.2 (5036)</td>
<td>6.3 (1069)</td>
<td>9.6 (5245)</td>
</tr>
<tr>
<td>Total assessments</td>
<td>100.0 (71543)</td>
<td>100.0 (13148)</td>
<td>100.0 (58395)</td>
<td>100.0 (22117)</td>
<td>100.0 (49426)</td>
<td>100.0 (17122)</td>
<td>100.0 (54421)</td>
</tr>
</tbody>
</table>

BD, bipolar affective disorder.

*Stable BD, subjects who received the diagnosis of BD in at least 75% of the evaluations.
†Non-stable BD, subjects who did not receive a diagnosis of BD in at least 75% of the evaluations.
‡Last diagnosis BD, subjects who received a diagnosis of BD at the last evaluation.
§Last diagnosis not BD, subjects who did not receive a diagnosis of BD at the last evaluation.
*First diagnosis BD, subjects who received a diagnosis of BD at the first evaluation.
**First diagnosis not BD, subjects who did not receive a diagnosis of BD at the first evaluation.
schizophrenia (F20.0; 15%, 7886/54 421); residual schizophrenia (F20.5; 13%, 7018/54 421); dysthymia (F34.1; 10%, 5245/54 421); BD, current episode mild or moderate depression (F31.3; 9%, 4706/54 421); major depressive disorder, recurrent (F33; 8%, 4270/54 421); and BD, current episode manic without psychotic symptoms (F31.1; 7%, 3683/54 421).

The most frequent diagnoses during the study period among all 49 426 assessments of the 710 patients were: BD, current episode mild or moderate depression (F31.3; 22%, 4817/22 117); BD, current episode manic without psychotic symptoms (F31.1; 21%, 3683/54 421); and BD, current episode mild or moderate depression (F31.3; 22%, 3987/54 421); BD, current episode mixed; F31.8, other bipolar affective disorders; F30–F39, affective disorders due to use of alcohol; F20–F29, schizophrenia, schizotypal and delusional disorders; F20, schizophrenia; F20.0, paranoid schizophrenia; F20.5, residual schizophrenia; F22, persistent delusional disorders; F23, acute and transient psychotic disorders; F25, schizoaffective disorders; F30–F39, mood (affective) disorders; F30, manic episode; F31.1, bipolar affective disorder, current episode manic without psychotic symptoms; F31.3, bipolar affective disorder, current episode manic without psychotic symptoms; F31.1, bipolar affective disorder, current episode manic without psychotic symptoms (F31.1; 5%, 2524/49 426).

**Diagnostic changes**

*Markov’s models* Four Markov’s models were calculated. The first model, which included the whole sample, is represented in Fig. 1. The y-axis represents ‘prior’ diagnostic states, and the x-axis represents ‘next’ diagnostic states. Each pixel in Figs 1–4 represents the probability of a transition between the ‘prior’ diagnostic stage (i.e. F10) and the ‘next’ diagnostic stage (i.e. F20) in this case, switch from previous diagnosis of ‘mental and behavioral disorders due to use of alcohol’ to next diagnosis of ‘schizophrenia’). The probability of each transition is represented by a colour gradient from dark blue (the lowest probability) to dark red (the highest probability). This model shows that the most probable transitions in the whole sample were within the same diagnostic block. This means that diagnoses remained quite stable over time within the same diagnostic group (i.e. a diagnostic change between a ‘prior’ diagnostic stage of F20.5 and a ‘next’ diagnostic stage of F20.0 – a diagnostic stage within the same diagnostic block – has high probability, whereas a diagnostic change between a ‘prior’ diagnostic stage of F20.5 and a ‘next’ diagnostic stage of F31.1 – a diagnostic stage from a different diagnostic block – has low probability). The highest probabilities are distributed on the diagonal of Fig. 1.

The second model (Fig. 2) included patients who had received the diagnosis of BD in at least 75% of the evaluations (‘stable BD’ group). This is the most interesting model, as the ‘stable BD’ group includes all patients who have consistently been assigned the diagnosis of BD by the clinicians who have assessed them. The most probable transitions across diagnostic blocks were from other diagnoses to BD, but not from BD to other diagnoses. This indicates that patients who receive a stable diagnosis of BD have previously received other psychiatric diagnoses, but once they receive a stable diagnosis of BD they do not switch to any other diagnostic block. The most probable transitions across diagnostic blocks were:

![Fig. 1. Probability of transitions between prior diagnoses and next diagnoses in the whole sample (n = 1153).](image-url)
i) from F10–F19, mental and behavioral disorders due to use of psychoactive substances, to F31.1 BD, current episode manic without psychotic symptoms;
ii) from F20, schizophrenia, to F31.1 BD, current episode manic without psychotic symptoms;
iii) from F23, acute and transient psychotic disorders, to F31.9 BD, most recent episode unspecified;
iv) from F42, obsessive–compulsive disorder, to F31.1 BD, current episode manic without psychotic symptoms;
v) from F60–F69, disorders of adult personality and behavior, to F31.1 BD, current episode manic without psychotic symptoms;
vi) from F90–F98, behavioral and emotional disorders with onset usually occurring in childhood and adolescence, to F31.8, other BD;
vii) from F90–F98, behavioral and emotional disorders with onset usually occurring in childhood and adolescence, to F41, other anxiety disorders.

The third model (Fig. 3) included patients who had received the diagnosis of BD at the last evaluation (last diagnosis BD). The results are similar to the first model and the most probable transitions are within the same diagnosis.

The fourth model (Fig. 4) included patients who had received the diagnosis of BD at the first evaluation (first diagnosis BD). The results are similar to the first model and the most probable transitions are within the same diagnosis. In this model, we included time between stage changes (<1 month or ≥1 month) in the analysis.

Discussion

Summary of principal findings

Only 30% of subjects received the diagnosis of BD at the first evaluation, whereas 70% got the diagnosis during later contacts. Kessing (13) obtained similar results and pointed out that these figures are consistent with the high prevalence of misdiagnosis of 48% (20) and 69% (21) found in naturalistic investigations using self-administered questionnaires on contact to doctors in general.

Patients with a stable diagnosis of BD (‘stable BD’ group) presented some diagnostic fluctuation involving the typical diagnoses included in the differential diagnosis of BD. The disorders that presented the highest probability of transition to BD according to the second Markov’s model were:

Fig. 2. Probability of transitions between prior diagnoses and next diagnoses in the 266 ‘stable bipolar affective disorder (BD)’ patients (patients who have received the diagnosis of BD in ≥75% of the evaluations).

Fig. 3. Probability of transitions between prior diagnoses and next diagnoses of the 443 ‘last diagnosis bipolar affective disorder (BD)’ patients (patients who received the diagnosis of BD at the last evaluation).

Fig. 4. Probability of transitions between prior diagnoses and next diagnoses of the 342 ‘first diagnosis bipolar affective disorder (BD)’ patients (patients who received the diagnosis of BD at the first evaluation) including time between stages.
Diagnostic stability of bipolar disorder

The prevalence of BD in this psychiatric sample is specifically trained to increase inter-rater reliability. Clinicians who assigned the diagnoses were blind to study procedures. Other published studies have used semistructured interviews and other diagnostic instruments not used ordinarily in clinical practice. The results of our study may more accurately reflect the real use of diagnostic classifications in psychiatric practice and may be more useful in estimating the clinical utility of current psychiatric classification systems.

Our study has limitations. The limitations are those inherent to a naturalistic study performed in real world conditions. Structured or semistructured clinical interviews were not used in this study for the assessment of ICD-10 BD. Psychiatrists used ICD criteria to classify the patients. Moreover, the clinicians who assigned the diagnoses were not specifically trained to increase inter-rater reliability. The prevalence of BD in this psychiatric sample is lower than that found in other studies performed on psychiatric populations. This may be related to the fact that the Spanish psychiatric services are easily accessible by individuals in the community.

Strengths and weaknesses in relation to other studies, discussing important differences in results

Other authors have reported the rates of consistency that are much higher than the ones found in the present study (1, 3, 12, 13, 15). However, most studies that have evaluated the stability of BD have shorter follow-up periods than the present study and have focused on a single clinical setting (mainly the in-patient setting). Schwartz et al. (1) reported that the rates of consistency for some diagnoses decreased as the follow-up period increased. The retrospective consistency of BD was 85% when comparing 6- and 24-month diagnoses, but lowered to 73% when comparing baseline and 24-month diagnoses. However, compared with Schwartz’s data, the retrospective consistency of BD across clinical settings in this study (38%) is strikingly low. A structured interview, the Structured Clinical Interview for DSM-III-R (SCID) provided DSM-III-R psychiatric diagnoses in the study by Schwartz et al. (1). Perhaps the use of semistructured interviews would have enhanced reliability and therefore stability. However, given the large number of assessments in this study, it is possible that this instability reflects poor validity of psychiatric diagnostic categories as currently conceived.

Meaning of the study: possible implications

Follow-up studies including the evidence of diagnostic stability and diagnostic consistency over time have traditionally been proposed to test the validity of psychiatric diagnoses (22–25). Diagnostic changes over time may reflect the evolution of an illness, the emergence of new information, or unreliability of measurement (1).

The relative lack of stability in diagnoses over time in the present study may be due to the evolution of the illness or reflect the inherent weaknesses in clinical assessment. The temporal consistency of BD in our study was lower than that found in other longitudinal studies. The relative lack of diagnostic stability over time is striking given that there is likely to be a bias toward maintaining the same diagnosis over time. Psychiatrists treating the patients in this study often had access to past records and diagnoses, and may have been inclined to keep the previous diagnosis rather than assign a different one.

In spite of the limitations of the present study, the low probability of transitions between bipolar disorder and psychotic disorder in stable and non-stable bipolar patients might reflect the fact that in clinicians’ minds, bipolar disorder and psychotic disorder are two independent disorders following the classic Kraepelinian dichotomy. The current popular view among researchers is that bipolar and psychotic disorders might not be discrete ‘disease entities’ but dimensions of continuous variations.
Baca-Garcia et al.

(26). New studies with bigger samples including the whole spectrum of psychotic and affective disorders are needed to further explore these views.

The results of the present investigation raise worrisome concerns regarding the validity of the results of epidemiologic, clinical, and pharmacologic psychiatric research, particularly, in studies of chronic disorders with short follow-up periods that may not allow enough time to reach the right diagnosis or in studies that do not take setting into account. Diagnostic stability is relevant for estimating prevalence and incidence. Whereas incidence is a measure of risk, prevalence is influenced by episode duration (prognosis) and by mortality. Ideally, it would be possible to classify associations that are observed in epidemiologic prevalence data according to their main determinants: incidence and episode duration (27).

Acknowledgements

These patients were studied with no external support.

References


Results from the Vienna follow-up study. Psychopathology 1991;24:328–335.