**UNIVERSIDAD AUTÓNOMA DE MADRID**


**Facultad de Medicina**

**Departamento de Psiquiatría**

# ESTABILIDAD DIAGNÓSTICA DEL TRASTORNO BIPOLAR

## TESIS DOCTORAL

## JORGE LÓPEZ CASTROMÁN

**Mención de doctorado europeo**

**Director:**

**DR. ENRIQUE BACA GARCÍA**

**Codirector:**

**DR. IGNACIO BASURTE VILLAMOR**

**UNIVERSIDAD AUTÓNOMA DE MADRID**


**Medical Faculty**

**Psychiatry Department**

# DIAGNOSTIC STABILITY OF BIPOLAR DISORDER

## DOCTORAL THESIS

## JORGE LÓPEZ CASTROMÁN

**"Doctor Europeus" mention**

**Primary supervisor:**

**DR. ENRIQUE BACA-GARCÍA**

**Secondary supervisor:**

**DR. IGNACIO BASURTE VILLAMOR**

**Para Ana,**

**que me ha acompañado y sufrido,**

**y sigue guardando el secreto del amor.**

# AGRADECIMIENTOS

Muchas personas e instituciones han colaborado directamente en la elaboración de esta tesis y muchas otras lo han hecho de forma indirecta durante mi formación. A todos ellos, mi gratitud.

Al Dr. Enrique Baca García, director de esta Tesis Doctoral, por su impulso, su apoyo y sus consejos. Su favor inestimable me acompaña para mi suerte desde que comencé la especialidad de psiquiatría. Espero seguir aprendiendo de él.

Al Dr. Ignacio Basurte Villamor, codirector, sus aportaciones han sido esenciales para el desarrollo metodológico del presente trabajo y su ironía me aconseja a menudo con acierto.

Al Dr. Juan José Carballo y la Dra. Mercedes Pérez, con los que he compartido el duro trabajo y espero compartir la grata recompensa del éxito.

A los Profesores Gilles Bertschy (Suiza) y Manuel Esteves (Portugal), que con su revisión experta de esta tesis han ayudado amablemente a que se llevase a cabo.

Al Departamento de Teoría de la Señal de la Universidad Carlos III, colaboradores imprescindibles en el análisis estadístico no convencional.

A los compañeros y la dirección de los equipos de salud mental en donde se llevó a cabo este estudio, sin su asistencia nada de esto hubiera sido posible.

A los miembros del Tribunal de Tesis, ante los que me siento honrado de presentar mi trabajo. Confío en que mis méritos estén a la altura de su prestigio.

A todos los compañeros que desde que inicie mi formación, me han transmitido la dedicación a la profesión; en el Centro de Salud Mental de Centro, en el Hospital 12 de Octubre, en el Centro de día del Hospital Puerta de Hierro, en el Servicio de Psiquiatría Infantil del Hospital General Universitario Gregorio Marañón y en el Centro de Terapia Breve de Jonction en Ginebra.

Muy especialmente a todos los compañeros del Servicio de Psiquiatría de la Fundación Jiménez Díaz en Madrid. A mis residentes “mayores” y “menores” con los que afortunadamente he llegado a coincidir. En su compañía he vivido y convivido la psiquiatría.

A mis amigos, que a pesar de no entender qué es el trastorno bipolar, siguen pensando que no tengo ni idea de lo que he escrito.

A mi familia, a la que tanto extraño desde la lejanía. A mis hermanos, que me han guiado, y a mis sobrinos, que traen la ilusión del futuro.

A mis padres, que son y serán siempre el pilar y el martillo de mis inquietudes y esfuerzos. Mis logros son fruto de su labor y de su cariño.

# ÍNDEX

# INTRODUCCIÓN

En un artículo pionero publicado en 1970, Robins y Guze mencionan la estabilidad diagnóstica como uno de los criterios necesarios para verificar la presencia de un síndrome psiquiátrico y la relacionan por primera vez con la validez predictiva de los diagnósticos en psiquiatría (1). Posteriormente la estabilidad diagnóstica ha sido definida como la medida en la que un diagnóstico es confirmado en evaluaciones consecutivas (2;3). En ausencia de correlatos biológicos objetivables de la enfermedad, la estabilidad del diagnóstico en el tiempo representa la mejor prueba para validar diagnósticos psiquiátricos y sirve en gran medida para predecir el curso de un trastorno (4). Por este motivo el estudio de la estabilidad diagnóstica puede fundamentar el manejo terapéutico de los pacientes. Por el contrario la ausencia de estabilidad en un diagnóstico puede generar graves repercusiones.

Existen múltiples factores que pueden provocar inestabilidad en los diagnósticos psiquiátricos. Spitzer (5;6) ha señalado entre las fuentes de falta de fiabilidad (fuentes de varianza) que conllevan desacuerdos en el diagnóstico entre clínicos las siguientes categorías:

(i) el propio paciente (cambios en el paciente o varianza propia del paciente),

(ii) el episodio concreto de la enfermedad que es valorado o varianza dependiente de la ocasión, por ejemplo las diferencias que se encuentran clínicamente en diferentes episodios de un mismo paciente con trastorno bipolar,

(iii) la varianza en la información, la aparición de nuevos datos durante el seguimiento o las diferencias que se producen en función del escenario clínico o de los informantes,

(iv) la varianza en la observación (interpretaciones diferentes del mismo cuadro por distintos especialistas),

(v) los criterios utilizados en la valoración o varianza en los criterios, por ejemplo el uso por dos observadores distintos de diferentes criterios para diagnosticar un delirio.

La propia evolución de las enfermedades mentales, y en concreto la evolución del trastorno bipolar, hace difícil a menudo diferenciar la entidad clínica de los cuadros y se relaciona directamente con las dos primeras causas mencionadas al producir cambios en el paciente y en el cuadro clínico. Asimismo, la falta de información completa sobre el curso de la enfermedad y posibles errores diagnósticos previos pueden también dar lugar a un sesgo en la valoración. Se han propuesto distintos métodos para potenciar la estabilidad de un diagnóstico aunque ninguno de ellos asegura la fiabilidad del resultado. Estos métodos incluirían la evaluación u observación longitudinal (7;8), y también estudios diagnósticos avanzados de tipo genético (9), la monitorización de la respuesta al tratamiento (10) o la evaluación de los efectos de la enfermedad sobre la función psicosocial (11).

La inestabilidad diagnóstica generará inevitablemente dudas en el abordaje y tratamiento de los trastornos mentales, y las potenciales consecuencias negativas para los pacientes son difíciles de cuantificar al carecer de datos concluyentes. Los estudios enfocados a delimitar la estabilidad diagnóstica de los trastornos psiquiátricos en poblaciones adultas son relativamente escasos, aunque su número ha aumentado de forma importante en los últimos años (2;4;7;12-44). Cabe destacar que entre ellos sólo unos

pocos se han orientado de forma exclusiva hacia el estudio del trastorno bipolar (7;13-16), mientras que un número mucho mayor ha estudiado la evolución de los primeros episodios psicóticos (2;4;17-24). Por otro lado, un número considerable de estudios se ha centrado en la evolución en la etapa adulta de diagnósticos otorgados en la infancia (3;45-61).

Es importante resaltar que este estudio forma parte de un proyecto más amplio. La generación de una extensa base de datos que ambiciona reflejar las actuaciones psiquiátricas en toda la región de Madrid ha posibilitado hasta el momento la realización de múltiples trabajos de investigación epidemiológica (12;13;62-64) y facilitará de cara al futuro la producción de muchos otros. En el presente caso, las características de este registro nos han permitido estudiar la estabilidad diagnóstica del trastorno bipolar en base a un número mucho mayor de evaluaciones y de escenarios que en la mayoría de los estudios previos (13;65).

El trastorno bipolar es considerado una enfermedad crónica y como tal, una vez establecido, el diagnóstico debería ser estable. Sin embargo en la práctica clínica encontramos con frecuencia que esto no es así (7). Las dificultades diagnósticas parecen especialmente relevantes al inicio de la enfermedad y a menudo el diagnóstico se asegura únicamente con el seguimiento de los pacientes a largo plazo. En muchas ocasiones se producen cambios en su clasificación durante la evolución del trastorno, sobre todo hacia diagnósticos del espectro de la esquizofrenia (14;17), y se han descrito diversos factores que pueden estar relacionados con la inestabilidad de este diagnóstico:

1. Las manifestaciones clínicas de la enfermedad pueden variar a lo largo del tiempo y solaparse con las de otros trastornos psiquiátricos (66).
2. Los trastornos comórbidos también pueden alterar la apariencia clínica de la enfermedad (10;67).

3. Los diagnósticos realizados por distintos observadores podrían ser inconsistentes (34).
4. Los factores sociodemográficos pueden también alterar el curso del trastorno, sus síntomas iniciales o la percepción de los clínicos (17).

Por este motivo el análisis de la evolución diagnóstica del trastorno bipolar y la posibilidad de un cambio diagnóstico desde otro trastorno hacia el trastorno bipolar o en sentido inverso es relevante para la investigación psiquiátrica. También resulta relevante la búsqueda de los factores implicados en su inestabilidad. Hasta la actualidad existen escasas investigaciones que hayan evaluado el impacto de la estabilidad diagnóstica del trastorno bipolar en la continuidad y persistencia del mismo a lo largo del tiempo (68); sólo algunos estudios han investigado hasta el momento el cambio diagnóstico en el trastorno bipolar o las relaciones entre los distintos diagnósticos en su evolución (2;7;14;17;23). En general se utiliza en mayor medida la clasificación DSM-IV, en la búsqueda bibliográfica encontramos sólo dos autores que analizan la estabilidad del diagnóstico de trastorno bipolar utilizando criterios CIE-10 para la clasificación de los pacientes (14;19). La estabilidad diagnóstica del trastorno bipolar encontrada en estos estudios ha sido moderada o alta, sin embargo, muchos de estos resultados se encuentran limitados por el empleo de un número escaso de evaluaciones – generalmente dos o tres- (2;16-18;22-24;35;61), así como por la corta duración de seguimiento, inferior a 3 años en la mayoría de los trabajos (2;17;18;21-24;61), con algunas excepciones (7;8;14;16;35). Estas limitaciones condicionan la generalización de los resultados de las investigaciones previas y sugieren la necesidad de desarrollar nuevos estudios.

Por otra parte, varias razones sugieren que las investigaciones sobre la estabilidad y el cambio diagnóstico del trastorno bipolar podrían adquirir una notable relevancia:

1. Las dificultades diagnósticas complican el estudio epidemiológico sobre este trastorno, pero en términos generales se acepta una prevalencia de entre el 1-2% de la población (69;70), independientemente del grupo étnico. Sin embargo, los estudios poblacionales más recientes han mostrado cifras más elevadas de prevalencia-vida del trastorno bipolar tipo I con respecto a estudios previos (71-75), y la mayoría de expertos en este trastorno coinciden en afirmar que los datos actuales probablemente subestimen la prevalencia real de la enfermedad, que podría acercarse a cifras en torno al 5% si se incluyen los trastornos del espectro bipolar (76-78).
2. Numerosos estudios señalan que los errores en el diagnóstico del trastorno bipolar son frecuentes (79-83). La prevalencia de errores diagnósticos en la evaluación inicial, en su mayor parte debida a la confusión con episodios depresivos unipolares, puede situarse entre un 48% y un 69% de acuerdo con los datos de las investigaciones realizadas por la National Depressive and Manic-Depressive Association (79;80). Un estudio señala que el periodo de tiempo medio desde el inicio de la enfermedad hasta iniciar un tratamiento de mantenimiento con litio se sitúa en 8.38 años en una muestra de pacientes con trastorno bipolar (84).
3. Diversos trabajos han señalado un mejor pronóstico en aquellos pacientes en los que el trastorno bipolar es diagnosticado precozmente debido a varias razones: alteraciones de la vida diaria, beneficios del inicio temprano del tratamiento, efectos deletéreos del tratamiento inadecuado y riesgo de suicidio (83). Algunos autores han propuesto el empleo de instrumentos de cribado en la evaluación inicial de los pacientes con síntomas depresivos para prevenir estos problemas a través del diagnóstico precoz (76;77). Entre

las consecuencias clínicas del infradiagnóstico de trastorno bipolar se señalan: suicidio, abuso de sustancias y complicaciones iatrogénicas. En conjunto los pacientes sufren una notable disminución de la calidad de vida, importantes costes económicos (76) y riesgo de suicidio, que en los pacientes bipolares puede ser mayor que el de cualquier otro trastorno psiquiátrico (85;86).

4. De acuerdo con el informe de la Organización Mundial de la Salud en el año 2001, entre las principales causas de años vividos con discapacidad (ylds) en todas las edades el trastorno bipolar figura como novena y en el grupo de edad entre los 15 y los 44 años figura quinta en todo el mundo (87). Asimismo el trastorno bipolar ha sido clasificado tercero entre las enfermedades mentales como causa de gravamen por enfermedad en las economías de mercado, después de la depresión unipolar y de la esquizofrenia. El trastorno bipolar a menudo se relaciona con disputas familiares, problemas con el sistema de justicia y problemas en el lugar de trabajo (69;88). Constituye una enfermedad crónica severa y supone un gran coste a la sociedad.
5. La comorbilidad médica y psiquiátrica es especialmente frecuente en pacientes diagnosticados de trastorno bipolar (89). Se suele situar la prevalencia de comorbilidad en pacientes con trastorno bipolar de tipo I por encima del 50% (90). Especialmente comunes son los trastornos por abuso de sustancias (48-71% de los pacientes) y los trastornos de ansiedad (42-93% de los pacientes) (83).

Los datos anteriores indican la importancia de profundizar en el estudio del trastorno bipolar y las notables repercusiones que puede tener un mejor entendimiento

de la enfermedad. El objetivo de este trabajo es describir la evolución del diagnóstico de trastorno bipolar en condiciones reales. El proceso diagnóstico de este trastorno se enfrenta a múltiples escollos y por el momento el trastorno bipolar continúa siendo una entidad cuya definición y delimitación es controvertida a pesar de su importante prevalencia en la población general. La estabilidad de un diagnóstico se ha relacionado con la validez del mismo y numerosas evidencias sugieren que puede condicionar el adecuado abordaje y tratamiento de la enfermedad. El estudio de la aplicación real de las clasificaciones nosológicas en la práctica clínica puede clarificar los límites diagnósticos de la enfermedad. La evaluación ecológica en múltiples escenarios clínicos y condiciones reales de la estabilidad y la evolución a largo plazo del diagnóstico de trastorno bipolar según la Clasificación Internacional de Enfermedades en su 10ª revisión puede contribuir a este proceso.

# INTRODUCTION

In a novel paper published in 1970, Robins and Guze mention the diagnostic stability as one of the necessary criteria to verify the presence of a psychiatric syndrome and establish for the first time a relationship with the predictive value of psychiatric diagnoses (1). Subsequently the diagnostic stability has been defined as the measure in which a diagnosis is confirmed in consecutive assessments (2;3). In the absence of objective biological symptomatology of the disorder, the diagnostic stability over time represents the best proof to validate psychiatric diagnoses and to a great extent it can be used to predict the course of the disorder (4). For this reason the study of diagnostic stability can serve as a basis for the therapeutic handling of the patients. On the contrary the absence of stability in a diagnosis may generate serious impact.

There are many factors that may cause instability in psychiatric diagnoses. Spitzer (5;6) has signalled among the sources of unreliability (sources of variance) that lead to diagnostic disagreement among clinicians the following categories:

(i) the person itself (changes in the patient or subject variance);

(ii) the particular episode of the disease or occasions variance (e.g., differences found in the clinical picture of a patient with bipolar disorder between two episodes of the disease);

(iii) the information variance, for instance through the apparition of new data along the follow-up or differences across settings and informants;

(iv) the observation variance (different interpretations of the same profile by diverse clinicians);

(v) the criteria used for the assessment or criterion variance (e.g., use of different criteria between two observers to diagnose a delusion).

The evolution of the mental diseases itself, and specifically that of the bipolar disorder, often makes the differentiation of the clinical profiles problematic; and the it is directly related with the two initial above-mentioned reasons both being a source of changes in the patient and in the clinical profile. Likewise, the lack of the whole information concerning the course of an illness and the possible previous diagnostic errors may as well cause the apparition of an evaluation bias. Different methods have been proposed to increase the stability of a diagnosis though none of them assure the fiability of the result. These methods would include the longitudinal evaluation or observation (7;8), and also advanced diagnostic procedures like genetic links (9), the monitoring of treatment response (10) or the evaluation of the illness effects on the psychosocial functioning (11).

Diagnostic instability questions the approach to mental disorders and the treatment applied on them. The potential negative consequences for the patients are difficult to quantify in the absence of concluding data. There is a limited number of studies focused on diagnostic stability of psychiatric disorders in adult populations, though its number has been largely increased in the last years (2;4;7;12-44). It is worth pointing out that just some of these studies were focused on bipolar disorder (7;13-16), while a larger number was devoted to study the evolution of first-psychotic episodes (2;4;17-24). On the other hand considerable research has been done on evolution of childhood diagnoses into adulthood (3;45-61).

It is important to emphasize that this study is part of a broader project. The production of a vast database that aspires to include psychiatric interventions in the whole region of Madrid has made possible the accomplishment of many epidemiological research studies (12;13;62-64) and will encourage towards the future the apparition of many others. In our case, register characteristics have helped us to study the diagnostic stability of

bipolar disorder on account of a larger number of assessments and settings than most of the previous studies (13;65).

Bipolar disorder is considered a lifelong illness and as such, once established, diagnosis should be stable. Nevertheless in clinical practice we frequently find that this is not the case (7). Diagnostic difficulties seem to be especially relevant near the onset of the illness, and quite often the only way of ensuring diagnosis is the long-term follow-up of the patients. Along the evolution of the disorder changes in its classification are common, especially towards schizophrenia spectrum diagnoses (14;17), and many different factors have been reported to decrease diagnostic stability of bipolar disorders:

1. Clinical signs of the illness may vary along time and overlap other psychiatric disorders (66).
2. Comorbid disorders may as well change the clinical appearance of the disease (10;67).
3. Diagnoses made by different observers may be inconsistent (34).
4. Sociodemographic factors may also disturb the course of the disorder, its initial symptoms or the perception of the practitioners (17).

For this reason the analysis of diagnostic evolution of bipolar disorder and the possibility of a diagnostic change from another disorder towards bipolar disorder or vice versa, is relevant for psychiatric research. As it is relevant the investigation of factors related with its instability. Till now, little research has been done to evaluate the impact of diagnostic stability on the continuity and persistence of bipolar disorder (68); few studies have investigated till the moment diagnostic change in bipolar disorder or the relationship among the different diagnoses along the evolution (2;7;14;17;23). Generally DSM-IV classification is more used, in the bibliographic search we found

only two authors trying to establish diagnostic stability of bipolar disorder using ICD-10 criteria for the classification of patients (14;19). Bipolar disorder stability in these studies has been found moderate or high. Nevertheless, many of the results are limited by the use of a scant number of assessments –commonly two or three- (2;16-18;22-24;35;61), and by the short duration of follow-up not, being less than 3 years in most of them (2;17;18;21-24;61), with some exceptions (7;8;14;16;35). These facts limit the generalisation of previous research results and suggest the need for the development of new studies that may make up for those limitations.

Moreover, some reasons suggest that research on stability and diagnostic change of bipolar disorder may acquire remarkable relevancy:

1. Diagnostic difficulties complicate epidemiological study of this disorder, but in general terms accepted prevalence is between 1-2% of the population, independently of the ethnic group (69;70). However recent poblational studies have shown higher rates of life-prevalence for bipolar I disorder than the previous investigations (71-75), and most of the experts on this disorder agree that actual data probably underestimate the real prevalence of the illness, that could approach 5% including bipolar spectrum disorders (76-78).
2. Misdiagnosis of bipolar disorder is frequent according to several studies (79-83). Prevalence of misdiagnosis in the initial assessment, mostly out of the confusion with unipolar depression episodes, may range from 48% to 69% according to research data of the National Depressive and Manic-Depressive Association (79;80). One study on the average time from illness onset to maintenance with lithium treatment shows a mean period of 8.38 years in a simple of bipolar patients (84).

3. Several studies have signalled that early recognition of bipolar disorder improve the prognosis of the patients accounting to: daily life disturbances, benefits of treatment early onset, detrimental effects of unsuitable treatment and suicide risk (83). Some authors have proposed the use of screening tests in the initial evaluation of patients with depressive symptomatology to prevent those problems through early diagnosis (76;77). Among the clinical consequences of under recognised bipolar disorder are described: suicide, substance abuse and iatrogenic complications. Moreover, patients suffer a remarkable diminution in their quality of life and important socio-economic costs (76) and may have a higher risk of suicide than patients with any other psychiatric or medical illness (85;86).
4. According to the World Health Organisation report 2001, among the top-ten leading causes of years of life lived with disability (ylds) worldwide, in all ages bipolar disorder is ranked ninth and in 15–44-year-olds is ranked fifth (87). Moreover, bipolar disorder is ranked third among mental illnesses as the source of disease burden in market economies, after unipolar depression and schizophrenia. Bipolar disorder is frequently combined with family discord, problems with the justice system and workplace problems (69;88). As a chronic and severe psychiatric disorder it is extremely expensive to society.
5. Medical and psychiatric comorbidity are especially frequent in bipolar patients (89). Comorbidity rates for type I bipolar disorder are often placed over 50% (90). Especially common are substance abuse disorders (48-71% of the patients) and anxiety disorders (42-93% of the patients) (83).

The above-mentioned data points out the importance of further investigation on bipolar disorder and the prominent impact that may bring about a better understanding of the illness. The aim of this work is to describe thoroughly the evolution of bipolar disorder diagnosis in real conditions. Diagnostic process of this disorder faces many pitfalls and at present continues to be an entity whose definition and delimitation remains controversial in spite of its important prevalence in the general population. The stability of a diagnosis has been related to the validity of it, and several studies suggest that it may condition the right approach and treatment of the illness. Study of the realistic application of nosologic classifications in clinical practice may help clarify the diagnostic limits of the illness. Ecological evaluation in multiple clinical settings and realistic conditions of stability and long-term evolution of bipolar disorder diagnosis according to International Disease Classification, 10th revision (ICD-10) may contribute to this process.

# BACKGROUND AND SIGNIFICANCE

The issue of the "gold standard" as a basement for diagnostic procedures is particularly problematic in psychiatry, where diagnoses are defined by their manifestations rather than by direct biological markers (91). Moreover, categorial limits of the psychiatric nosologies are frequently diffuse (92) and diagnostic crossover is not scarce. Research advances in psychiatric genetics have raised the expectation of a major breakthrough, however, the extent to which genetic findings can resettle psychiatric nosology is in doubt (91;93) and recent reports on the liability to bipolar disorder has shown evidence of a large number of genes influencing the apparition of the illness (94). Krishnan in 2005 describes the following criteria for defining disease (95):

a. The condition should be one that leads to a risk for adverse outcomes, either mortality or functional impairment.
b. If an identifiable characteristic (environmental factor, pathology or genetic factors) can be clearly defined that characteristic should separate the entity on at least one of the following criteria from similar entities:
    i. clinical symptoms,
    ii. course and outcome,
    iii. familial pattern and,
    iv. treatment response.
c. If no identifiable characteristic can be clearly defined and the defining feature is therefore nominalistic; then these nominalist features should separate from similar entities on at least one of the following:
    i. course and outcome,
    ii. familial pattern,

iii. treatment response.

Thus, whereas in most medical fields there is an identifiable characteristic that can separate clinical entities and ultimately tisular lesions seen in pathological examinations are used as the gold standard and allow the investigators to establish consistently the presence of a particular disorder (96), in the current state of knowledge a number of different instruments must be used in psychiatry to guarantee the consistence of a diagnosis:

a. Best estimate diagnosis: Considered the most valid method for diagnosing psychiatric disorders (96-99), best estimate diagnosis relies on the agreement between a number of experts; it is a diagnosis based on personal interview, family history from family informants, and medical records. Data from two different studies has shown that the diagnoses based exclusively on clinical data and using the best available expertise and a multilevel evaluation to arrive at a consensus best estimate diagnosis can only reach a kappa for agreement of 0.69 (97;99).
b. Longitudinal evaluation performed by an expert, using all data available, was as well conceived as a standard for validating psychiatric diagnoses (the longitudinal, expert, all data [LEAD] procedure) (100;101).

Importance of longitudinal evaluation and the use of multiple sources of information can be illustrated in the Health 2000 Study (102). This paper demonstrates using Best Estimate Diagnosis how lifetime prevalence of psychotic disorders may be underestimated in general population surveys. We expect the present study to follow a similar path in providing useful information to show the importance of longitudinal diagnosis in bipolar disorder under real world conditions.

## *CONCEPT OF STABILITY*

Diagnostic stability is the measure of the degree to which a diagnosis remains the same at subsequent assessments of the patient and constitutes a longitudinal validation of the original baseline diagnosis. It is based on agreement of diagnoses over time and is irrespective of cross-sectional diagnosis at a single point of follow-up (2). Traditionally evolution and prognosis of psychiatric disorders have played an important role as classification criteria since the beginning of XIX century. Kahlbaum made the difference between acute and chronic dementia and Kraepelin considered that evolution and outcome of the disorders was an essential feature for diagnosis (103;104). Follow-up studies including evidence of diagnostic stability and diagnostic consistency over time have been proposed since to test the validity of psychiatric diagnoses (105):

1. Robins and Guze in 1970 listed five criteria for establishing the validity of psychiatric diagnoses: 1) clinical description, 2) laboratory studies, 3) delimitation from other disorders, 4) follow-up studies including evidence of diagnostic stability, and 5) family studies (1). In their article relationship between diagnostic stability and predictive validity was proposed for the first time.
2. Kendler in 1980 revised this schema distinguishing between 1) antecedent validators, 2) concurrent validators, and 3) predictive validators that included diagnostic consistency over time, together with rates of relapse and recovery and response to treatment (106).
3. Andreasen in 1995 proposed to sum up the use of newer methods of validation that may link symptoms and diagnoses to their neural substrates with the clinical and epidemiological approach that included validators such as the

characteristic course of the illness and the follow-up studies to determine outcome (107).

4. First et al in 2004 defined diagnostic validity as a complex multifaceted construct that includes a number of different types of validity: 1) Face validity, meaning to describe accurately the disorder; 2) Descriptive validity, whether the features of a category serve to distinguish it from the others; 3) Predictive validity, the extent to which a diagnosis predicts future clinical course or diagnostic stability; 4) External or construct validity, association with expected external validators (108). This new definition attained to amend the weakness of previous validity criteria, that implicitly assumed psychiatric disorders to be discrete entities (109).

## *IMPORTANCE OF DIAGNOSTIC STABILITY*

Until this moment definitions for psychiatric diagnoses are based on expert opinion rather than on their biological basis, and the modest knowledge base regarding underlying aetiologies has hindered the use of aetiological factors in psychiatric classification systems (12). It is assumed that the higher the diagnostic stability, the more likely is to reflect a consistent psychopathological or pathophysiological process. Being that the main clinical purpose of diagnosis, as a formulation, is to furnish the informational basis for planning and conducting clinical care (110), stability of a diagnosis gives a relevant base not only for prediction of the course and outcome of a disorder but also for effective planning and provision of treatment. The best approach for evaluating the natural history of psychiatric disorders is through longitudinal studies, especially prospective ones, methodologically preferred due to their inherent prevention of recall bias to occur.

The introduction of explicit diagnostic criteria in diverse rule-based classification systems, like DSM-III and ICD-10, has deeply affected psychiatric practice (111). This standard frame of reference has permitted to achieve better diagnostic agreement and to improve statistical reporting and analysis (109). Classifications have progressively contributed to a better recognition and diagnosis of psychiatric disorders. The availability of longitudinal data, however, may cause significant fluctuations in diagnostic stability as changes in clinical presentation are seen (95). Thus, evolving longitudinal observations should lead to periodic updating of the comprehensive diagnostic formulation (112), and yet, despite the inherent problems derived from criteria based on cross-sectional observations, our diagnostic system relies on stable diagnoses (113). Accounting for the potentially harmful consequences of unsuitable treatment options or clinical interventions, the study of diagnostic stability remains an essential issue in psychiatry.

## Stability estimate

Diagnostic stability has been operationalized in different ways by the researchers (2). Most longitudinal studies use data on the course of illness and symptom patterns over time to confirm or question the original diagnosis. In our study we will use the most common approach, examining stability of bipolar disorder on the light of multiple evaluations along the evolution of the illness. To our knowledge only one study have considered a diagnosis stable only if the confirmation criteria are fulfilled at the time of follow-up (66). This one study, that addressed the short-term temporal stability of psychotic disorders in a first-admission sample, found that 80% of patients initially classified as bipolar and major depressive disorder had no symptoms and received no diagnosis at the 9 and 18 months reassessments.

Positive concordance rates describe the percentage of those diagnosed with a disorder at one time that manifest the same disorder at subsequent assessments (present–present). However, there is a deficiency in positive concordance rates, as they fail to account for the introduction of new cases of a disorder. The kappa coefficient is able to correct this problem by accounting for positive and negative concordance rates, as well as rates of discordant cases (114;115). In this way, Kappa provides a more comprehensive estimate of stability and corrects for agreement due to chance and is commonly preferred to concordance rates. Nonetheless, kappa may produce a misleading estimate of stability if reported alone; it is grossly reduced by high incidences of new cases and excessively magnified by high negative concordance rates. Accounting for these pitfalls, it is important to examine both concordance rates and kappa estimates together to establish diagnostic stability accurately (116).

## *THE BURDEN OF BIPOLAR DISORDER*

Bipolar affective disorder is an ICD–10 mental disorder characterised by at least two episodes involving clinically significant disturbed mood, energy and activity (117). It is a severe lifelong illness characterized by unpredictable recurrent manic (or hypomanic) and depressive episodes and a high mortality rate (118).

### Epidemiology

The prevalence rate, long debated, is at present accepted to be between 1-2% of the general population, though bipolar spectrum is not included in this estimation (69;119). Findings reported by the recent multi-center European study ESEMeD reveal lower frequencies, under 1% (120;121). NEMESIS Study suggested a prevalence rate for bipolar disorder of 1.9% (72). ECA study (Epidemiologic Catchment Area) life-long prevalence of bipolar disorder was found to be 1.3% (0.8% for type I and 0.5% for type

II) (122). NCS study (National Comorbidity Survey) placed the life-long prevalence for mania and hypomania at 1.6% (71). In Spain two authors have studied the prevalence of bipolar disorder in small samples, Vasquez-Barquero et al. found a rate of 0.08 for manic-depressive psychosis manic type, and Canals et al. found a prevalence of 2.4% for hypomania (ICD-10) (123-125). Overall high variability in the results, from 0.2% to 3.3%, is observed when including studies performed on smaller and clinical samples (71-75;81;126). This variability has been explained in relation to methodological differences, comparing the use of structured interviews in population-based surveys, to the use of another kind of instruments (78;102;127). Multiple sources of information are essential for accurate estimation of lifetime prevalence of bipolar disorder (102).

Last decades have seen an enlargement of the concept of bipolar disorder and a significant increase of its importance (118;128). It is widely admitted that bipolar II disorder, cyclothymic disorder and other forms of bipolar disorder could be up to at least 5% of the general population (126;129-131), and that the conception of the disorder is changing towards a disease continuum. Some studies have considered the possibility of a continuation between borderline personality and bipolar-II disorder (132). In a prospective investigation of the hereditary nature of schizophrenia and bipolar disorders (the Iowa study) the prevalence rates of mania among relatives of patients affected by bipolar disorders, calculated on the basis of diagnostic interviews, was 1.9%; however, when additional sources of information such as clinical records and cross-interviewing of relatives were considered, the rate increased to 5.3% (133). A twenty-year longitudinal community study carried out in Zurich (134) found that depressives with a subthreshold hypomanic syndrome were similar to bipolar II disorders in terms of positive family history for mania, course, comorbidity and

treatment rates. Furthermore, sub-threshold manic symptoms in adolescence have been described as highly predictive of subsequent onset of a manic episode (135;136).

## Misdiagnosis

The onset of bipolar disorder usually occurs before the age of 25 years but it is frequently treacherous presenting a confusing picture for both clinician and patient (83). Many patients experience behavioural mood problems long before the first episode of the disease, while many others present major episodes of the illness without a clear prodrome. Some typical features of bipolar depression have been described to distinguish it from unipolar depression: atypical depression, abrupt onset and end, highly recurrent pattern, positive family history, early age of onset (137;138). However, the current systems of classification use the same single set of symptoms to describe both types of depression and the reason for seeking help in bipolar disorder is often the presence of depressive symptomatology, rather than manic or hypomanic symptoms, which could explain a substantial proportion of misdiagnosis in the first consultations. Stigma and misunderstanding of the illness may explain as well a significant delay. Its worth pointing out that depressive episodes are the commonest cause of morbidity and, indeed, of death by suicide (139).

The problematic diagnosis of bipolar disorder has been depicted in several studies. According to research data of the National Depressive and Manic-Depressive Association (NDMDA) the prevalence of misdiagnosis in the initial assessment of bipolar patients ranges from 48% (Lish et al 1994) to 69% (Hirschfeld et al 2003) (79;80), mostly out of the confusion with unipolar depression episodes. Specifical data from 2000 NDMDA survey showed that 31% of bipolar disorder patients have been misdiagnosed as unipolar depression and in nearly half of them (49%) the condition was neither recognised nor diagnosed (79;83). Ghaemi et al, 1999, presented evidence of an

underdiagnosis of bipolar disorder in clinical samples. 40% of the patients had received the wrong diagnosis of major depression (81). Similar results are reported by Angst et al, 2002, 25% to 50 % of major depression cases in their sample were actually affected by bipolar disorder (82). Hirschfeld et al, 2003, reported an average time to bipolar disorder diagnosis over 10 years in the third part of the sample and that those who were misdiagnosed consulted a mean of 4 physicians prior to receiving the correct diagnosis (79). One study on the average time from illness onset to maintenance with lithium treatment shows a mean period of 8.38 years in a sample of bipolar patients (84), this figure is corroborated by other studies reporting that treatment for bipolar disorder is not started until up to 10 years after onset (80;140). The mean therapeutic gap, proportion of individuals with psychiatric disorders that remain untreated although effective treatments exists, was calculated to be 50.2% in the population worldwide for bipolar disorder (141).

Misdiagnosis is a major factor leading to a poor outcome for patients. It frequently complicates attempts at effective management of bipolar disorder and also plays a major role in the economic burden of the illness (119). Emerging evidence suggests that early intervention results in a more favourable outcome for bipolar disorder (79); at the same time a delay in initiation of mood-stabilizing therapy for patients with bipolar disorder has been associated with increased healthcare costs (142). Accurate diagnosis and treatment may also be protective against the functional impairment associated with bipolar disorder (143).

## Cost of bipolar disorder

According to the World Health Organization 2001 report bipolar disorder is the fifth in the top 10 causes of disability worldwide in the 15 to 44 year age group, and the ninth in all ages (87). It has been ranked seventh among the world-wide causes of non-

fatal disease burden. Das Gupta & Guest placed the cost of bipolar disorder in 2002 for the UK to £2 billion; Wyatt & Henter estimated in 1991 the total 1-year cost of approximately 2 million prevalent cases in the USA to $45 billion (126;144). These studies did not include bipolar spectrum disorders.

Following the most frequent terminology in cost-of-illness studies, direct costs include all those directly produced by medical attention, while indirect costs derive from the incapacity level, the effect on work productivity, effects on social welfare and costs related with justice processes (145). Direct healthcare costs derived from the illness were significantly superior in patients not initially diagnosed as bipolar disorder, which did not receive mood stabilizer treatment after a first depressive episode following the California Medicaid Program (142). Nevertheless the majority of the cost derived from bipolar disorder is accounted for by the indirect costs in relation to decreased functional capacity and lost work (88). Bipolar disorder is associated with high rates of unemployment, job-related difficulties, and interpersonal stress (69). In one survey 88% of the respondents reported occupational difficulties (79).

The treatment of bipolar illness could be enhanced by public health efforts to promote early diagnosis and treatment (80). In a recent study based on ESEMeD, Fernández et al. have found that only one third of the mental health treatment in Spain met minimal adequacy criteria (146), while the World Health Organization (WHO) through the programme entitled CHOosing Interventions that are Cost-Effective (WHO–CHOICE) reveals that, assuming 50% population coverage, clinical interventions have the potential to reduce the current burden of bipolar disorder by 10-33% (144).

## Comorbidity

An array of illness patterns occur in bipolar disorder, including rapid cycling, mixed states and an extensive comorbidity that often complicates diagnosis and treatment and contributes to the cost of the illness (88;89). Comorbidity rates have been calculated over 50% in a number of studies (147-151). Vieta et al. reported comorbidity rates of 31% in a sample of Spanish bipolar patients in 2001, the difference explained by the author on the selection of euthymic patients in psychiatric primary care outpatients settings, thus avoiding the risk of mistaking acute affective symptomatology as symptomatology provided by a second illness (90).

Among the most frequent comorbidity of bipolar disorder is included:

1. Substance and alcohol abuse. Estimates of comorbid drug abuse range from 14% to 60% following the review made by Cassidy et al (152), with most authors pointing out rates bigger than 30%. McElroy et al describe 42% comorbid substance misuse in a sample with 288 bipolar patients (151). In the ECA study bipolar I disorder patients were found to be more than 3 times as likely to have alcohol abuse or dependence and about 7 times more likely to have drug abuse or dependence than the general population (153).
2. Anxiety disorders. High rates for comorbid lifetime panic disorder (21%) and comorbid lifetime obsessive-compulsive disorder (21%) were found in bipolar I and II patients included in the ECA survey (154;155). McElroy et al found 42% of comorbid anxiety disorders (151). In the NCS study 92% of patients with bipolar I disorder also met criteria for a lifetime anxiety disorder (148).
3. Personality disorders. Usual rates of comorbid personality disorders in patients affected by bipolar disorder are around 30% when evaluating

patients during euthymic intervals (156-159). Cluster B (dramatic/emotionally erratic) and cluster C (fearful/avoidant) personality disorders are the most likely to be codiagnosed with bipolar disorder (156;160;161). Personality disorder traits predict poorer medication compliance among bipolar adults (162;163) and the absence of social supports that buffer against relapse (161).

4. Attention deficit hyperactivity disorder (ADHD). Symptomatology both of ADHD and bipolar disorder frequently overlaps (138;164). This may mislead the diagnostic process and subsequent treatment (psychoestimulants may induce mania or rapid cycling). Rates of comorbidity with ADHD range from 60% to 90% (165).
5. Suicide. In the fields of clinical practice and prevention, it is often underlined that bipolar disorders represent a devastating risk factor for both suicide attempts and suicide itself (134;166). Suicide rates, averaging 0.4% per year in men and women diagnosed with bipolar disorder, are 20-fold higher than in the general population (86). Bipolar disorder has shown a strong relationship to a history of suicide attempts (29.2%) relative to unipolar disorder (15.9%) and other Axis I disorders (4.2%) (85).

## *BIPOLAR DISORDER STABILITY*

### Boundaries of bipolar disorder

In the original article by Blacker and Tsuang published in 1992, bipolar disorder was claimed to be one of the most robust diagnostic entities in psychiatry though the article identifies a number of "contested boundaries" of the disorder raising the question on whether improved diagnostic criteria may be necessary for a number of research and clinical purposes (10). Among the reasons that could demonstrate the distinct

phenomenology of bipolar disorder the following are mentioned: its occurrence across history (167-169) and cultures (70), its patterns of inheritance (133;134), and its clear disturbance of physiologic function (11).

The heterogeneity in the expression and progression of the clinical manifestations of the bipolar disorder complicates substantially the correct and early diagnosis of the disease. Diagnosis of bipolar patients, specially during the early acute phase of the disease, is frequently hampered by impediments such as instability of symptoms, patient's and family's denial, inconsistencies in retrospective information, concomitant substance abuse or personality disorders, unclear relationship between affective and psychotic symptoms and symptomatology overlap with other axis I disorders (113). Some of the common entities that may hinder the correct diagnosis of bipolar disorder are:

1. Schizophrenia. A major cause of inaccurate diagnosis of bipolar disorder is the confusion with schizophrenia (9;31;133). The patient may not be clearly classifiable when present symptomatology may be simultaneously similar to bipolar disorder's and schizophrenia's usual symptoms (7). Schizoaffective disorder forms a buffer zone between both diagnoses including the cases in which psychotic symptoms are not clearly linked to the affective episodes (10).
2. Unipolar depression (170). Due to the high prevalence of unipolar depression, its distinction with bipolar depression becomes extremely important. Failure on the recognition of potential bipolars (or false unipolars) (11) is one of the main pitfalls to determine the true prevalence of bipolar disorder . Bipolar II disorder plays a boundary role, it has been identified by many researchers with the course of illness and epidemiological

characteristics of bipolar disorder type I (75). Difficulties in assessing hypomania limit the application of this diagnosis (15;130).

3. Personality disorders. Confusion with personality disorders essentially concerns borderline personality disorder (171). There are contradictory findings regarding the relationship borderline personality disorder and bipolar disorder (172;173). The symptomatic overlap with bipolar II disorder has been related with mood instability and impulsivity, both diagnostic criteria of borderline personality disorder (132).
4. Drug abuse may also lead to misdiagnosis of bipolar disorder disguising affective syndromes in substance abusers (174). Both false positives, due to patients intoxicated with stimulants that may be erroneously considered maniac (150), and false negatives, due to masking, can appear (10).
5. Childhood syndromes, including mania and depression of early onset, should be considered (3;55;136;175-178). An association has been found between onset at early age and increased switching from unipolar to bipolar disorder (54). ADHD may as well induce misdiagnosis of bipolar disorder, due to the substantial overlap in the symptomatology (164;179;180).
6. Cyclothimia. The mood cycling in bipolar II disorder may be difficult to distinguish from cyclothymic temperamental disorder (129;171). Some authors propose subthreshold mood lability of a cyclothymic nature to be the common thread that links the bipolar spectrum (181).

## Literature on bipolar disorder stability

Notwithstanding the difficulties in diagnosis, most studies to date suggest moderate to high levels of temporal diagnostic stability of bipolar disorder (2;7;8;17;19;22-24;31). As far as we know only two studies have investigated the

stability of ICD-10 diagnosis of bipolar disorder (14;19) and only one of them has done it under realistic clinical conditions (14). Most of the investigations till the moment are limited by the use of few assessment points – two or three in most of them − and short follow-up periods (see INTRODUCTION, pages 12/19), raising concerns about the generalizability of results and suggesting the need for the development of new studies capable of overcoming such limitations. As the assessments are usually conducted at two remote points in time, a time gap between them not controlled for diagnostic status and temporary changes is present in the majority. On the other hand most of the previous studies did not analyze the factors related to diagnostic change nor did they report the subsequent diagnoses of patients who were not given a specific diagnosis at entry into the study.

We will briefly describe epidemiological and clinic-based studies that have evaluated diagnostic stability of bipolar disorder using prospective positive rates and/or kappa estimates. The following studies are focused exclusively on the stability of bipolar disorder diagnosis:

1. Kessing (14) investigated the diagnostic stability of the ICD-10 diagnosis of mania/bipolar disorder in a sample of 4116 patients. Data was obtained from the Danish Psychiatric Central Research Register, a nation-wide database that included registration of all psychiatric hospitalizations (1994-2002) and information on patients in public psychiatric ambulatories and community psychiatry centers (1995-2002). Subjects had got at least one diagnosis of manic episode (F30) or bipolar disorder (F31) along the study. Follow-up was divided into 10 contact periods. Chi-square test and Mann-Whitney test for two independent groups were used for the statistical analysis. 85.4% of individuals with main initial diagnosis of mania/bipolar disorder (N=2315)

got the diagnosis of bipolar disorder at the end of the second contact period, this proportion decreasing continuously till the tenth contact period (68,8%). On the contrary the number of subjects within this group with main diagnosis schizophrenia increased from 4.1% in the second period to 12.9% in the last, and substance abuse from 1.7% to 7.5%. Initial diagnosis different from mania/bipolar disorder was most commonly in the affective spectrum (40.7%), acute and transient psychotic disorders (15.6%), adjustment disorder (10.4%) or substance abuse (9.2%). Results showed that only 56.2% of the subjects obtained the diagnosis of bipolar disorder or mania at the first contact and that approximately 30% of those who were initially diagnosed eventually changed their diagnosis during the follow-up. Stability of bipolar disorder was thus studied on the base of the initial assessment. The study also found that female and younger patients had an increased risk of delay in the diagnosis of bipolar disorder.

2. Chen et al (7) reviewed the records of 936 patients with at least 4 hospitalizations within 7 years to assess the diagnostic change from bipolar disorder to other mental disorders. The hospital database contained longitudinal information on the diagnoses of these patients, assigned following DSM-III-R criteria. The set of patients was divided in two groups regarding the initial diagnosis of bipolar disorder or any other mental disorder to study the diagnostic flow. A subset of 443 patients with initial and subsequent diagnoses of bipolar disorder and/or schizophrenia was used to study specifically the flow between bipolar disorder and schizophrenia. To compare the rates of diagnostic changes to and from bipolar disorder chi-square or Fisher's exact tests were used. The results showed that only 60%

of the subjects completing the study period with a bipolar disorder diagnosis started the study with the same diagnosis. The most frequent diagnosis change from bipolar disorder was found to be schizophrenia (70.1%) though only 24.8% of those who changed to bipolar disorder changed from schizophrenia. The study found as well that more women than men changed diagnosis to bipolar disorder and that African-Americans were more likely to change from bipolar disorder to schizophrenia. As the stability was evaluated in readmission populations to the same facility; the reliability of these clinical diagnoses is limited by the inherent bias in sampling rehospitalized patients (17).

3. Weeke (16) published in 1984 a study based on the Danish Psychiatric Register. This database contains information concerning the patients admitted to Danish Psychiatric institutions. He studied the evolution of patients that had been admitted between April 1970 and March 1972 for the first time, and that had at least one more admission before March 1977, being classified as manic depressive following the ICD-8th revision in any of the admissions. 3062 individuals fulfilled the requirements. After the observation period 623 persons (20% of the register sample) were retrospectively classified as bipolar. Nevertheless in his study manic-depressive diagnoses included both unipolar and bipolar patients and his results regarding stability of manic-depressive diagnosis become misleading, the main finding of the study being that manic-depressive diagnosis is more stable among bipolar than among unipolar patients.

Some other prospective studies examined the stability of bipolar disorder diagnosis after a psychotic episode:

1. Schwartz et al (17) conducted a prospective epidemiological study on a cohort of 547 adults living in Suffolk County (New York). Patients enrolled were reassessed 6 and 24 months after a first-admission diagnosis of psychosis. Diagnoses were assigned by clinical consensus and following DSM-IV criteria, psychiatrists blind to previous research diagnoses. Information came from the Structure Clinical Interview for DSM-III-R administered at baseline and at 6- and 24- month follow-up plus medical records. The analysis of diagnostic stability was based on crossed-tabulation of diagnostic categories between the assessments and established two measures of stability: prospective consistency and retrospective consistency. Prospective consistency was the proportion of individuals in a category at 6-months who retain the same category of diagnosis in the 24-month assessment, and would correspond to positive predictive value taking the 24-month diagnosis to be the gold standard. Retrospective consistency conveying the proportion of subjects in a 24- month category that previously received the same diagnosis would represent sensitivity. The prospective consistency of bipolar disorder was high: 83% and, following the author, sustained the distinct nature of the disorder. Retrospective consistency of 73% was found for bipolar disorder. The study was focused on factors associated with the diagnostic shift to schizophrenia.
2. Amin et al (19) evaluated the stability of first-episode psychosis comparing ICD-10 and DSM-III-R systems. The study followed a cohort of 168 subjects with first-episode psychosis assigning to each of them a consensus onset diagnosis. After a three year follow-up, a longitudinal consensus diagnosis was decided, blind to the onset diagnoses. Diagnoses were based

on research interviews. Stability was measured by the positive predictive values (PPVs) of onset diagnoses, being considered the most common and accurate system on diagnostic stability. The study discusses different measures of stability, signalling other valuable measures like sensitivity, specificity and number of additional patients needed to prevent a false positive. Kappa statistic is used to calculate the concordance between onset and follow-up diagnoses and shows moderate agreement in the overall results but is not specifically calculated for bipolar disorder. The results of the study showed that only 78% of the patients with initial DSM-III-R diagnosis of bipolar disorder got the same diagnosis at 3-year follow-up, while 91% of the patients with an initial ICD-10 diagnosis of bipolar disorder (F30-31) got the same diagnosis at reassessment. However, only 21 patients where initially classified as manic psychosis according to ICD-10 which reduces the validity of this finding. Measures of diagnostic stability showed similar data between classification systems, though a trend for lower sensitivity in DSM-III-R compared to ICD-10 was found.

3. Fennig et al (2) presented an epidemiological study on the short-term stability of schizophrenic and other psychotic disorders. 278 first-admission subjects made up the sample. A best estimate diagnosis was made at baseline and after 6 months using the Structured Clinical Interview for DSM-III-R. Two psychiatrists examined the reasons for changes in diagnosis. Affective psychoses were relatively stable over the 6-month period, 86.5% of the patients keeping the same diagnostic category. This study addressed also the underlying cause for diagnostic change, classified in 4 different possibilities: 1) symptoms during the interval, that explained 43% of the changes; 2) new

interpretation of the original data, i.e. the diagnostic process itself, that explained up to 34.3% of the changes; 3) new information from other sources and 4) from the subject, that together were responsible for 22.1% of the changes. Among the findings of the study, the stability for bipolar disorder with psychosis was higher than the rates reported in the previous literature for hospitalised subjects with bipolar disorder.

4. A study by Rufino et al (18) tried to evaluate the stability of first psychotic episodes diagnosis in the emergency context. The sample included 59 patients assessed initially in the psychiatric emergency unit (with admission and discharge emergency diagnosis) and followed during a period of at least 12 months after the first evaluation. During the emergency admission severity scales were applied and the structured clinical interview for DSM-IV (SCID) was used on the follow-up. The agreement between diagnoses was calculated by kappa coefficient. SCID diagnoses after the follow-up determined four emergency diagnostic groups, namely: brief psychotic disorder, schizophrenia, manic index episode and depressive index episode. Manic episode diagnosis showed high levels of specificity (100%) but moderate levels of sensitivity (61.5%). A similar pattern was observed in the diagnosis of depressive episodes (specificity=77, 8%, sensitivity=98.0%).
5. Veen et al (23) conducted a study of a Dutch population-based psychosis incidence cohort. They tried to establish the diagnostic stability considered as the proportion of patients who received a follow-up diagnosis in the same main category as in the incidence study, and then focused on the diagnostic change to and from schizophrenic disorder. The incidence cohort consisted of 181 patients recruited in psychiatric and medical consultations. The

subjects went through a diagnostic interview and assessment instruments. Two and a half years later 168 participants were considered to have sufficient information available on which to base a second assessment. Psychotic mood disorders, including major depression and bipolar disorder, showed 67% of diagnostic stability.

6. Addington et al (24) examined diagnostic stability in a sample of 228 individuals who completed the one-year follow-up assessment after being admitted with a first episode of psychosis to a specialised program (the Calgary Early Psychosis Program). Subjects were excluded if they had previous history of affective psychosis. At initial assessment Structured Clinical Interview for DSM-IV was used. Diagnostic stability existed if the information gathered along the one-year follow-up confirmed the base-line diagnosis. Prospective and retrospective consistency was used as measure of diagnostic stability. The overall consistency of diagnosis over one year was 68% with an increase to 89% when schizophreniform was excluded. Only 4% (n=10) developed an affective disorder including bipolar disorder (n=5) and major depression (n=5).
7. Schimmelmann et al (22) assessed the diagnostic stability of psychotic disorders from 6 weeks to 18 months after initiation of treatment in a first-episode psychosis sample. Subjects were admitted in the Early Psychosis Prevention and Intervention Centre (EPPIC) in Australia from 1998 to 2000. Data were collected from patients' medical records (MRs) using a standardized questionnaire. Four hundred ninety-two subjects were analyzed. The same diagnosis was made at baseline (< or = 6 weeks after admission into EPPIC) and 18 months for 69.9% of the patients. Among the most

consistent diagnoses was bipolar disorder (83.2%). They concluded that it is necessary a longitudinally based diagnostic process, especially in schizophreniform disorder and bipolar disorder.

| Author/year | Sample | Study design | Methods | Instruments/Measures | Results |
|---|---|---|---|---|---|
| Kessing 2004 | N=4116 subjects (all ages) | Retrospective epidemiological study based on Danish Psychiatric Central Research Register. 9 years period | - Sample: Subjects with one diagnosis of manic episode (F30) or bipolar affective. disorder (F31)<br>- 10 contact periods of assessment.<br>- Diagnostic change quantified from initial diagnosis. | ICD-10 | 56,2% diagnosed as mania/bipolar disorder at first contact, of them, 30% changed their diagnosis during the follow-up |
| Chen 1998 | N=936 adults | Prospective longitudinal study. 7 years observation period. | - Sample: Subjects with at least 4 admissions to inpatient unit.<br>- Subsample of patients with initial and final diagnosis schizophrenia or bipolar disorder. | DSM-III-R | Prospective consistency 60% for bipolar disorder. Diagnostic change from bipolar disorder mostly schizophrenia: 70.1%. |
| Schwartz 2000 | N=547 adults ≥18 years | Prospective longitudinal. Suffolk County Mental Health Project. | - Sample: First-admission patients with psychosis.<br>-Baseline assessment.<br>- 6$^{th}$ and 24$^{th}$ months reassessments. | DSM-IV Structured Clinical Interview for DSM-III (SCID)<br>Scale for the assessment of negative/positive symptoms (SANS/SAPS)<br>Brief Psychiatric Rating Scale (BPRS) | Prospective consistency: 83%<br>Retrospective consistency: 73%<br>for bipolar disorder. |

| Author/year | Sample | Study design | Methods | Instruments/Measures | Results |
|---|---|---|---|---|---|
| Rufino 2005 | N=59 ≥16 years | Prospective longitudinal study. | - Sample: First episode of psychotic disorder.<br>- Emergency setting.<br>- 12-month minimal follow-up period. | DSM-IV<br>Structured clinical interview for DSM-IV axis I Disorders<br>Brief Psychiatric Rating Scale<br>Young Mania Rating Scale<br>Hamilton Rating Scale for Depression | Kappa=0.25 between admission emergency diagnosis and longitudinal diagnosis.<br>High levels of specificity in manic episode (100%) but moderate levels of sensitivity (61.5%) |
| Amin 1999 | N=168 subjects | Prospective cohort. 3-year follow-up. | -Sample: First-episode psychosis.<br>- Positive predictive values (PPV).<br>-Consensus diagnosis. | DSM-III-R<br>Schedules for clinical assessment in neuropsychiatry<br>Broad rating schedule<br>Disability assessment schedule<br>Scale for the assessment of negative symptoms | PPV: 78% initial DSM-IV bipolar disorder patients, 91% initial ICD-10 bipolar disorder patients (F30-31) |
| Fennig 1994 | N=278 patients, 15-60 years | Prospective longitudinal study. Suffolk County Mental Health Project. | -Sample: First-admission patients with psychosis.<br>-Consensus diagnosis.<br>-Assessment at baseline and after 6-month follow-up. | DSM-III-R<br>Structured clinical interview for DSM-III-R | Prospective consistency: 85.7%, retrospective consistency: 81.9% for bipolar disorder with psychotic features. |
| Veen 2004 | N=181 patients, 15-54 years | Dutch-population based incidence cohort. | - Sample: first consultation suspected psychotic disorder.<br>- Rediagnosed 30 months after first contact. | DSM–IV<br>Comprehensive assessment of symptoms and history<br>Retrospective assessment of the onset of schizophrenia | Psychotic mood disorders showed 67% of consistency. |

| Author/year | Sample | Study design | Methods | Instruments/Measures | Results |
|---|---|---|---|---|---|
| Weeke 1984 | N=3062, adults age not specified. | Epidemiological retrospective study. Danish Psychiatric Central Research Register. 7 years observation period. | - Sample: admitted patients with at least one manic-depressive diagnosis.<br>- Reassessed in second admission. | ICD-8 | 20% of the sample was retrospectively classified as bipolar. |
| Addington 2006 | 228 individuals, 16-50 years | Prospective longitudinal study. Calgary Early Psychosis Program | - Sample: First episode of psychosis.<br>- Reassessed at one-year follow-up. | DSM-IV Structured Clinical Interview for DSM-IV | Overall consistency: 68%. Only 2% developed a bipolar disorder. |
| Schimmelmann 2005 | 492 subjects | Prospective longitudinal study. Early Psychosis Prevention and Intervention Centre (EPPIC), Australia | - Sample: First-episode psychosis admitted patients.<br>-Reassessed after 18 months. | DSM-IV<br>Clinical Global Impressions-Severity of Illness scale<br>Global Assessment of Functioning score | 69.9% of overall consistency, 83.2% for bipolar disorder |

**Table 1. Studies on diagnostic stability and bipolar disorder**

## Future trends

Recent studies have suggested that permanent structural brain changes may be associated with bipolar disorder (143;182-184). Euthymic bipolar patients have shown diminished activation in response to the affective stimuli in both cortical and subcortical brain regions when compared with healthy subjects and are perhaps constrained in their ability to engage affective processing (183). Psychosocial function has been found to be compromised by the mood-state-related cognitive deficits in both bipolar depression and hypomania (184). Individuals diagnosed with bipolar disorder in both the acute and euthymic phases of illness display deficits on a range of neuropsychological tasks, and correlations between experienced number of affective episodes and task performance are commonly reported (143). Deficits of attention, learning and memory, and executive function have been asserted to be present (182). Furthermore, neuropsychological investigation of juveniles with bipolar disorder suggests that the same abnormalities present in adult bipolar sufferers may also be present in children (185;186).

These findings have fostered the existing interest in the neuropsychological profile of individuals with bipolar disorder. Evolution of bipolar disorder may thus be comparable to the models for neurodegenerative disorder causing progressive neuronal degeneration and resulting in a certain level of disability. Some studies have signalled the future use of biomarkers as the way to ensure the early detection of preclinical and clinical disease and provide the opportunity to start preventive therapy. Biomarkers (through genetics, clinical manifestations, neuroimaging or biochemistry) may help to identify at-risk groups, accelerate and enhance the accuracy of diagnoses and favour the development of drug treatments (187). Candidate brain function endophenotypes include attention deficits, deficits in verbal learning and memory, cognitive deficits after

tryptophan depletion, circadian rhythm instability, and dysmodulation of motivation and reward (188).

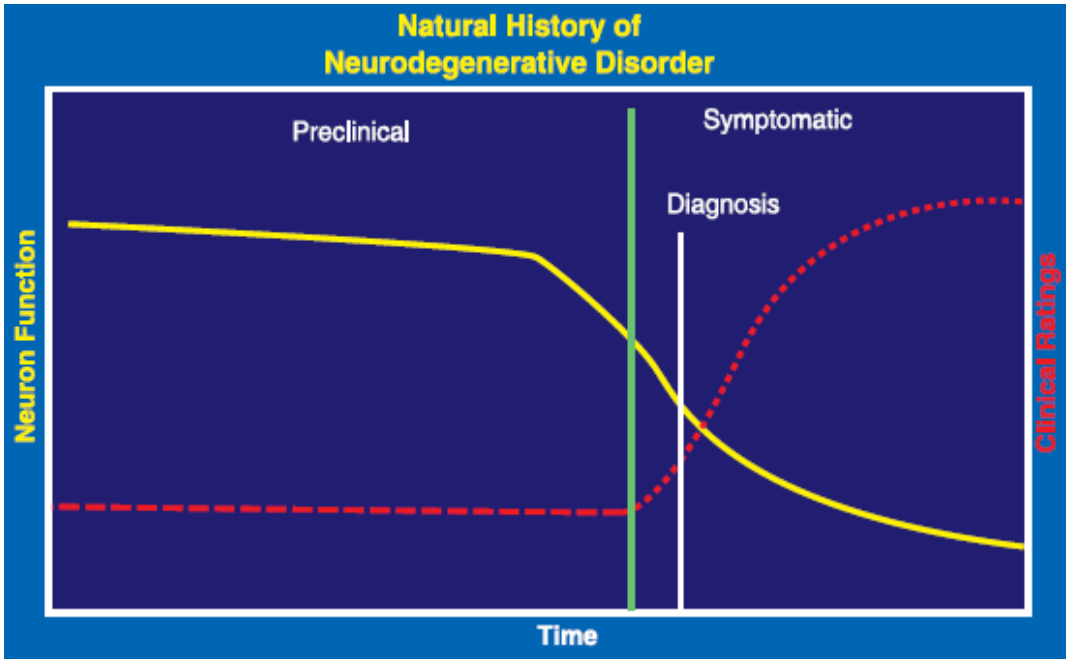



**Figure 1. Model for the progression of loss of neuronal function in neurodegenerative disorders. Source: DeKosky 2003.**

Evidence in literature suggest that substantial impairment is present in patients with bipolar disorder, even when they have not experienced more than two depressive or manic symptoms for a relatively long period, and when the patients have a relatively high level of education and do not abuse substances (143;189;190). The early detection of the illness becomes especially important in the view of these findings (185). Improvement in diagnostic procedures through the longitudinal assessment of the patients might play a crucial role correcting actual deficiencies in the detection and treatment of the disorder and may facilitate the practical application of the biological markers proposed. The wisdom on factors related to diagnostic change, usual diagnostic pathways and clinical consequences of diagnostic instability is still far from perfect.

## *Summary*

The importance of longitudinal evaluation has been often highlighted to ascertain the validity of diagnoses in psychiatry. Stability reflects the agreement of diagnoses over time and can be determined through different instruments. Bipolar disorder generates an important burden and yet there is a high prevalence of misdiagnosis that may contribute extensively to increase the costs. Summing up, a comparatively small number of studies have been purposefully designed to revise the relevance of diagnostic stability both in the particular case of bipolar disorder and applied to general psychiatric conditions. Findings from the existent epidemiological studies on bipolar disorder and from a small amount of studies conducted on the evolution of psychotic diagnoses have shown a moderate to high consistency of diagnosis of bipolar disorder. While these previous studies have provided detailed information about the diagnostic stability of the bipolar disorders, in general terms they are limited by few assessment points and a short follow-up. In addition, the use of fixed predetermined time intervals between assessment points may have contributed to the occurrence of recall bias.

Given the paucity information regarding diagnostic stability of bipolar disorders, we aimed to evaluate their long-term stability in a large sample of adult population who were evaluated at multiple time points, at least ten, in psychiatric clinical settings. This study provided a unique opportunity to shed light on the question of how stable do bipolar disorder diagnoses remain over time and how the degree of diagnostic stability may influence the clinical practice so determining the burden of this disease. We hypothesized that the stability over time of bipolar disorders could differ when assessed at multiple points and during a longer period of time, compared with the previous studies. In addition we intended to analyze the degree of diagnosis change to and from bipolar disorder over time and the entities most commonly included in this process. We

used three differences indices of diagnostic stability in the analysis, including a statistical model (Markov's model) that allowed us to a closer study of the conditions of diagnostic change in bipolar disorder.

# OBJETIVES

The aim of the present study is to carry out an ecological evaluation in multiple clinical settings (inpatient unit, psychiatric emergency room and outpatient clinic) of the long-term stability of bipolar affective disorder according to the International Classification of Diseases-10th edition.

The study contributes to ascertain the temporal consistency of bipolar disorder and the usual diagnostic changes occurring along the course of the illness, so establishing the base for future investigations on the evolution of diagnoses and the reasons for diagnostic changes.

# HIPOTHESIS

1. The degree in which a patient is consistently classified as having a bipolar disorder along the follow-up is an important marker for the validity of the diagnosis itself.
2. Temporal consistency of bipolar disorder may have been overestimated by previous studies. Studies containing multiple assessment points, different settings and larger samples are necessary to verify previous results.
3. 75% of coincidence in assessments along the follow-up could represent a suitable cut-off point to determine the existence of diagnostic stability.
4. Prospective and retrospective consistency of a diagnosis (between first and last diagnosis) are useful to evaluate the initial and final degree of misdiagnosis of a disorder.

# MATERIAL AND METHODS

## *Source of Data*

Beginning in 1986, public mental health centers in the province of Madrid, Spain, have recorded all psychiatric visits in a regional registry ('Registro Acumulativo de Casos de la Comunidad de Madrid'). From 1986 to 1992, diagnoses were coded according to the *International Classification of Diseases, Ninth Revision* (ICD–9) (World Health Organization (WHO), 1978). Since 1992, diagnoses were coded according to *International Classification of Diseases, Tenth Revision* (ICD-10) (WHO, 1992). Individual service users are reliably identified in the database used for our analyses because each patient is given an identifying number (a numeric code is used to ensure patient anonymity), which remains the same throughout all contacts with psychiatric services within the study area. To ensure that no patient had been assigned more than one identifier, we reviewed all the cases in the database and removed any duplicates we found. We defined duplicates as 'patients with identical first name, family name, gender and year of birth'; 'patients with identical first name, family name, gender and street address', or 'patients with identical first name, family name, gender and hospital/ambulatory record number'. We deleted any cases with significant suspicion of duplication. A unique identifying number (12) assigned to individual service users ensured patient anonymity and remained unchanged throughout all medical contacts.

### Data extraction

We extracted regional registry data regarding all psychiatric visits to all public psychiatric clinics belonging to the catchment area of Fundación Jiménez Díaz. From January 1$^{st}$, 1992, to December 31$^{st}$, 2004, all psychiatric visits in the area of the Fundación Jiménez Díaz, a general hospital in Madrid, Spain, have been recorded. This

hospital provides coverage to a catchment area of 300000 people and is a part of the Spanish National Health Services, which are financed by taxes to supply free of charge health care for all Spanish citizens and legal immigrants. In the period of study 34368 patients received psychiatric care to a total of 449317 psychiatric assessments.

## Description of the catchment area

The population that belongs to the catchment area of our study is placed in the central quarters of Madrid. Two districts with six basic zones each are included (Table 2). The demographic characteristics of this population are described in Table 3 thoroughly.

It is worth to note the high rate of immigrants, attaining over 26% of the total population in Centro district and about 15.5% in the case of Arganzuela district, compared with a 15.9% in the region of Madrid (191;192). These figures exceed markedly the 9.3% rate reported for the whole country at the same moment (193).

| Sanitary area in Madrid | Districts in the catchment area of Fundación Jiménez Díaz | Basic zones |
|---|---|---|
| 7 | 7.1 CENTRO | 7.1.1 Cortes<br>7.1.2 Justicia<br>7.1.3 Universidad<br>7.1.4 Palacio<br>7.1.5 Embajadores-1<br>7.1.6 Embajadores-2 |
| 11 | 11.2 ARGANZUELA | 11.2.1 Imperial<br>11.2.2 Acacias<br>11.2.3 Chopera<br>11.2.4 Palos de Moguer<br>11.2.5 Delicias-1<br>11.2.6 Delicias-2 |

**Table 2. Districts and zones included in the catchment area**

| Characteristics | | Madrid | Centro | Arganzuela |
|---|---|---|---|---|
| Area (Ha.) | | 60.708,69 | 523,08 | 648,10 |
| Density (inh./ha.) | | 53 | 270 | 231 |
| Population 1-1-2007 | | 3.187.062 | 141.396 | 149.577 |
| Age | 0 to 14 | 411.537 | 12.709 | 18.447 |
| | 15 to 64 | 2.177.603 | 104.352 | 105.935 |
| | 65 to 74 | 296.384 | 10.151 | 11.484 |
| | Over 75 | 301.493 | 14.183 | 13.706 |
| Nationality | Spain | 2.680.830 | 104.131 | 126.382 |
| | Foreign | 505.572 | 37.232 | 23.164 |
| | Foreigners % | 15,9 | 26,3 | 15,49 |
| Education (2001 Census) | Educational qualification | | 111.594 | 115.486 |
| | Illiterate | 42.863 | 1.506 | 943 |
| | No studies | 261.892 | 10.087 | 8.401 |
| | Primary education | 436.885 | 19.466 | 18.791 |
| | Secondary education | 552.111 | 23.113 | 24.021 |
| | Occupational training (FP) | 217.663 | 8077 | 9460 |
| | Degree | 227.215 | 10.133 | 12.794 |
| | Bachelor's degree | 357.518 | 18.719 | 18.848 |
| | Doctored | 33.534 | 2.115 | 1.497 |
| Economic activities (2001 Census) | Men over 16 | 1.169.870 | 51.140 | 52.012 |
| | Active | 796.904 | 36.917 | 36.175 |
| | Occupied | 712.498 | 31.684 | 32.614 |
| | Unemployed | 84.406 | 5.233 | 3.561 |
| | Unoccupied | 372.966 | 14.223 | 15.837 |
| | Women over 16 | 1.366.786 | 60.454 | 63.474 |
| | Active | 672.717 | 32.429 | 34.147 |
| | Occupied | 574.890 | 27.490 | 29.774 |
| | Unemployed | 97.827 | 4.939 | 4.373 |
| | Unoccupied | 694.069 | 28.025 | 29.327 |
| Natural increase (2005) | | 5.351 | -80 | 216 |
| Familiar income per capita in 2000 | Euros | 12.768 | 12.393 | 13.179 |
| | Index | 100 | 97,06 | 103,22 |

**Table 3. Demographic characteristics of the population**

In the following figures the districts included in the catchment area can be graphically seen (source: www.munimadrid.es). A smaller figure on the lower right corner shows its position in a map of Madrid.

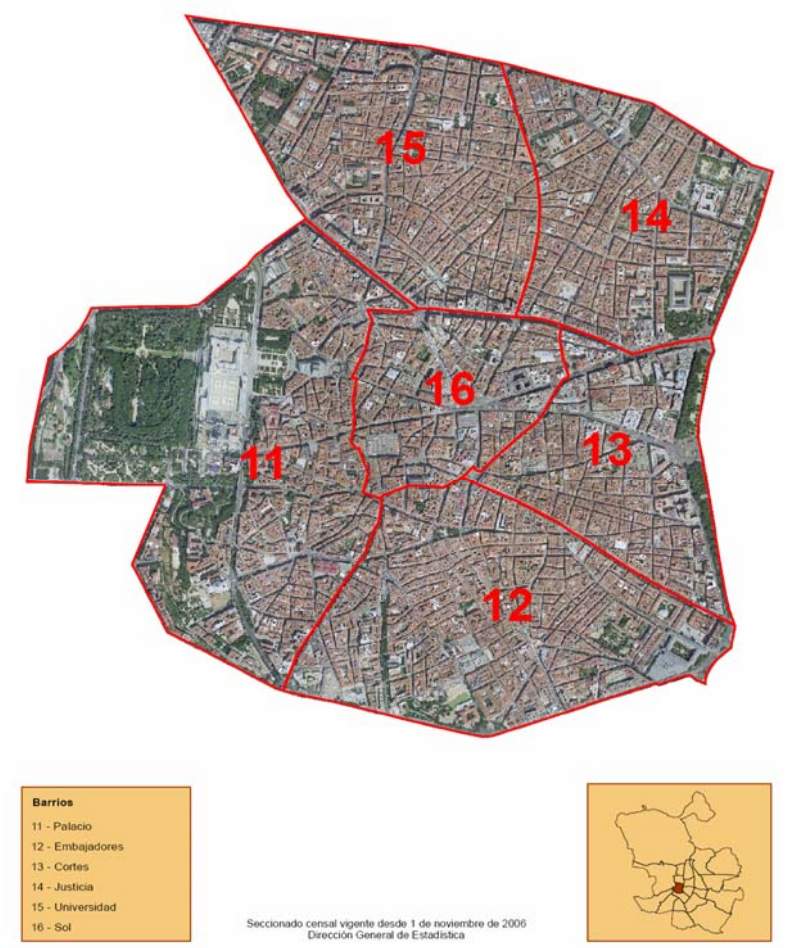



**Figure 2. Aerial view of Centro district and its basic zones**


**Figure 3. Aerial view of Arganzuela district and its basic zones**

## Setting

The assessments took place in three different clinical settings:

1. Outpatient psychiatric facilities (mental health care centers) within the catchment area of the Fundación Jiménez Díaz from 1992 to 2004.
2. Emergency room consultations from 2000 to 2004.
3. Inpatient unit (psychiatric brief hospitalization unit) from 2000 to 2004.

| | Study period | Psychiatric consultations in the study period |
|---|---|---|
| Outpatient facilities | 1992-2004 | 438622 |
| Emergency room | 2000-2004 | 9101 |
| Inpatient unit | 2000-2004 | 1594 |

**Table 4. Settings and number of psychiatric consultations.**

## Sistema de información

Se partió del registro de pacientes de los centros de salud mental del distrito de Arganzuela (área 11) y del distrito Centro (área 7) y de los registros de consultas externas, urgencias y de hospitalización de la Fundación Jiménez Díaz. Estos registros informatizados recogen el conjunto mínimo básico de datos (CMBD) definido por la Comunidad de Madrid (ver Anexo 1). La base de datos resultante contiene la información asistencial de 150.000 pacientes y cerca de 2 millones de actos médicos. Las fuentes fundamentales utilizadas en este trabajo fueron:

- Registro acumulativo de casos atendidos en los CSM de Centro (desde el 1/1/1992 hasta el 31/12/2004) y Arganzuela (desde el 1/1/1992 hasta el 31/12/2004). Este registro recoge el CMBD definido por la Comunidad de Madrid (ver Anexo I).

- Registro de urgencias de la Fundación Jiménez Díaz (desde el 1/1/2000 hasta el 31/12/2004). La información recogida en esta base de datos de filiación del paciente sin incluir datos sociodemográficos, ni diagnósticos, ni clínicos. Este registro se ha utilizado fundamentalmente para servir de soporte al registro desarrollado por el Servicio de Psiquiatría.
- Registro de hospitalización de la Fundación Jiménez Díaz (desde el 1/1/2000 hasta el 31/12/2004). Este registro está orientado a la facturación y recoge el CMBD con la intención de asignar grupos relacionados de diagnósticos (GRD) a los pacientes. El software para el cálculo es el 3Mv.
- Registros elaborados por el Servicio de Psiquiatría. El Servicio de Psiquiatría desarrolló su propio sistema de información paralelo al sistema de información de la Fundación Jiménez Díaz para validar la información recogida por el sistema general y recoger información adicional que no se refleja en el CMBD. Al detectarse algunas discrepancias se mantuvo.
- Registro de hospitalización de la Unidad de de Hospitalización Breve de la Fundación Jiménez Díaz, desde octubre de 2002 se protocoliza.
- Registro de las urgencias de la Fundación Jiménez Díaz atendidas por Psiquiatría (desde el 1/1/2000 hasta el 31/12/2004). Este registro comienza con la apertura de la urgencia psiquiátrica en la FJD. Desde octubre 2002 este registro se protocolizó.

Se diseñó una base de datos relacional con el programa File Maker v6.0 para integrar estos registros. El diseño de la base de datos relacional puede verse en la Figure 4. A cada paciente se le asignó una clave de identificación numérica y a cada asistencia se le asignó un número correlativo.

El procedimiento de fusión exigió un proceso de unificación de las bases de datos, depuración de posibles casos y asistencias repetidas y validación posterior, para ello se desarrollaron diversas rutinas de programación que permitirán actualizar periódicamente el sistema con los datos procedentes de las distintas fuentes. Se recodificaron los diagnósticos de la CIE-9 MC utilizada por el CMBD a la CIE-10 empleando las tablas de conversión entre la CIE-9 y la CIE-10, según criterios de la Organización Mundial de la Salud (194).

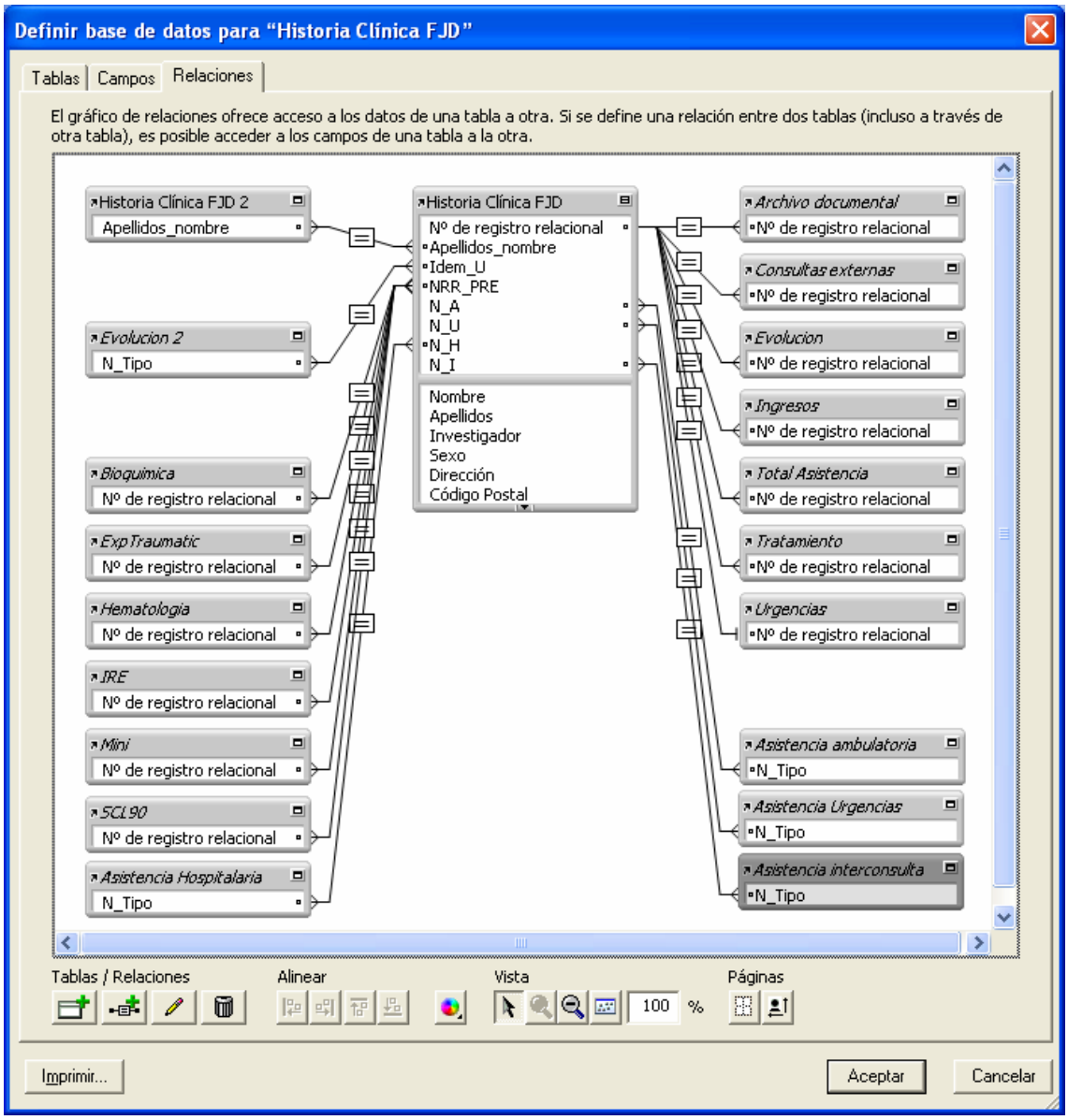



**Figure 4. Diseño relacional de la base de datos FileMaker. Fuente: Tesis del Dr. Ignacio Basurte Villamor.**

Paralelamente se desarrolló un interface para la introducción y consulta de datos (Figure 5 y Figure 6). La bondad de este sistema permite fusionar casi cualquier base de datos desarrollada para la asistencia sanitaria, basándose en una metodología similar a la descrita o que esté diseñada bajo los requisitos del CMBD.

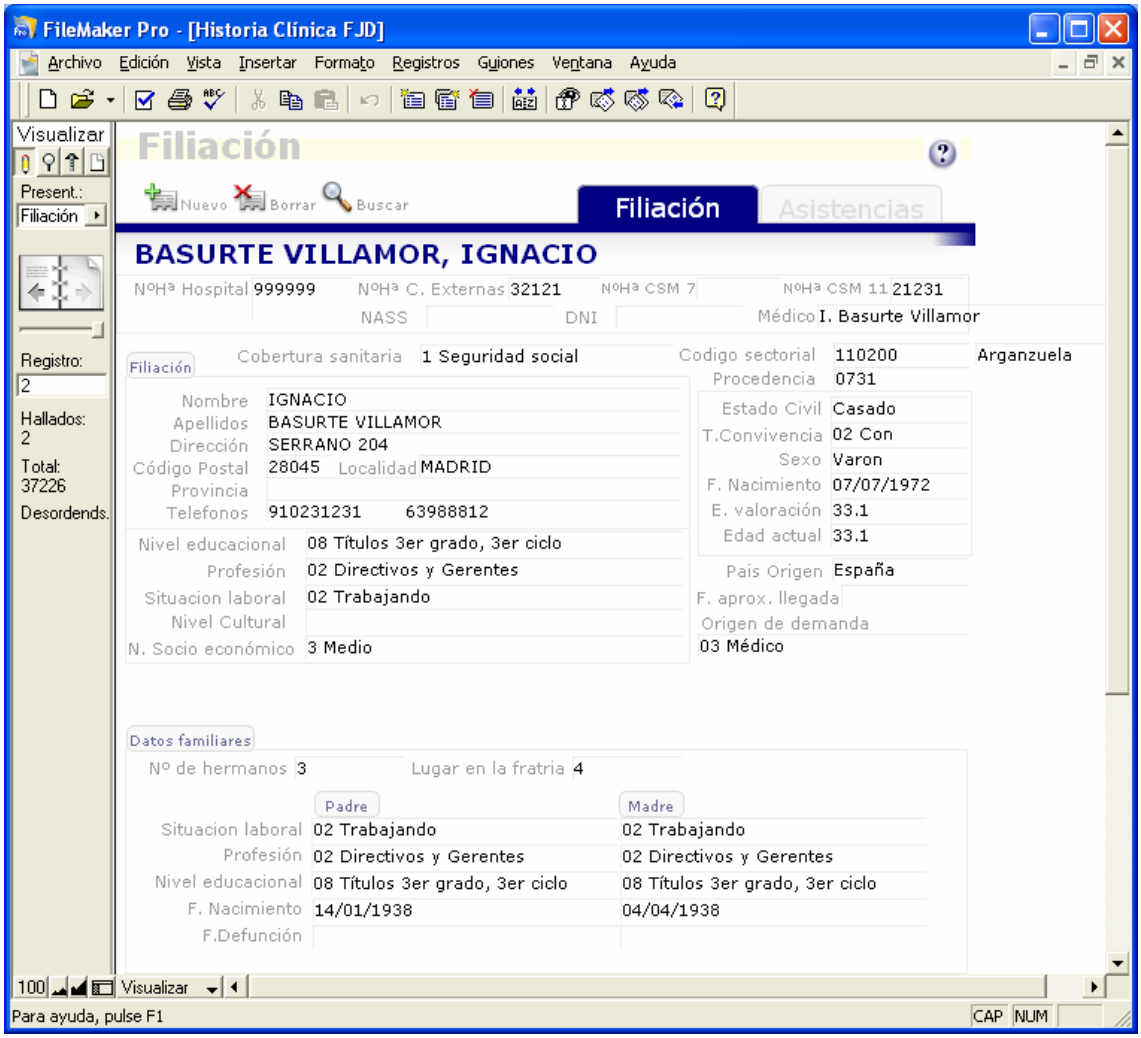



**Figure 5. Interface gráfico para introducción y consulta de datos (Filiación). Fuente: Tesis del Dr. Ignacio Basurte Villamor.**

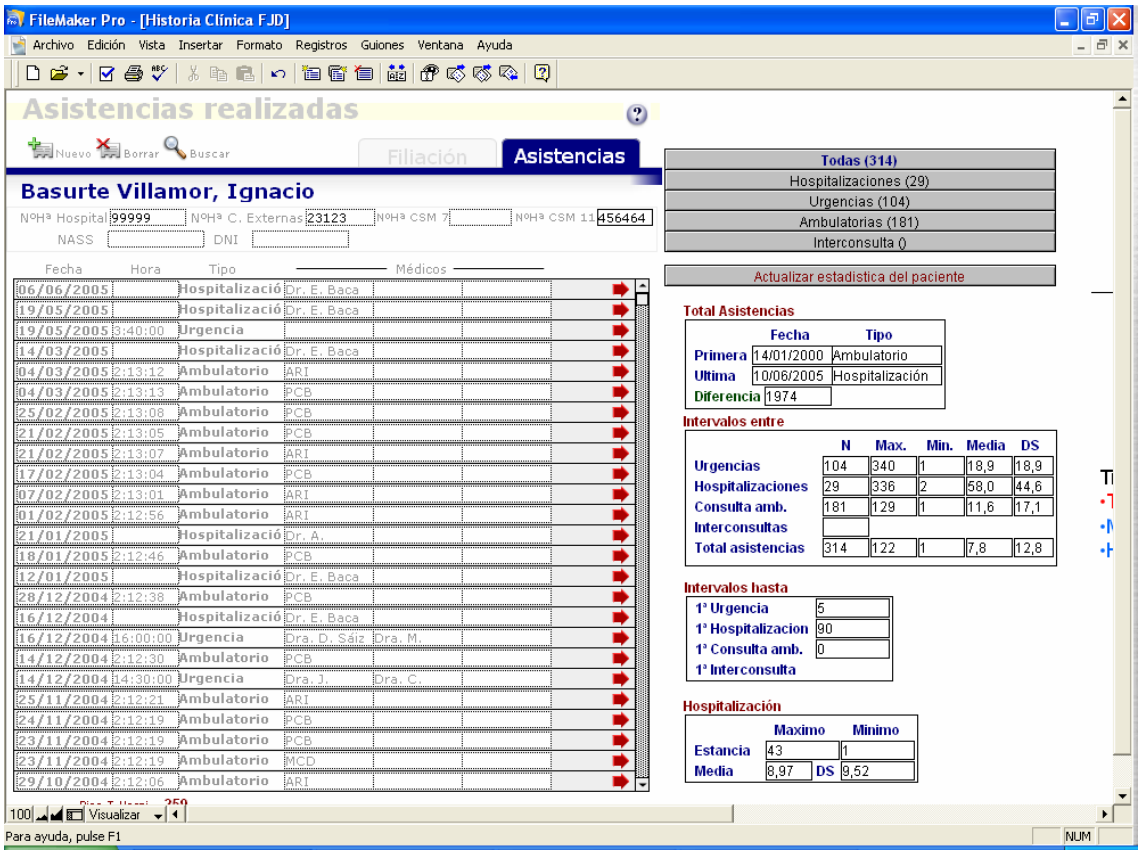



**Figure 6. Interface gráfico para introducción y consulta de datos (Evolución). Fuente: Tesis del Dr. Ignacio Basurte Villamor.**

## Participants

Participants were selected from a subsample of patients over 18 years of age that were assessed on at least 10 visits during the study period (N=10025). 10 assessments were agreed as a guarantee for a follow-up period of no less than half year, given that the interval between each consultation would be 18 days. We expected that this measure might reduce the risk of the assessments being made by different practitioners or that the briefness of the follow-up could interfere with the diagnostic process. A total of 1153 patients received a diagnosis of bipolar disorder according to ICD-10, in at least one evaluation. These 1153 patients had 71543 psychiatric consultations. The mean duration of follow-up for these patients was 6.2 years (SD 3.6) and the median of visits was 34. The study did not require informed conformity of the patients given that no additional intervention was accomplished on them and the anonymity has been

guaranteed by the use of a numeric codification system (see Sistema de información, page 48). The Institutional Review Board of the Fundación Jiménez Díaz Hospital approved the study.

## Diagnostic procedure

In all settings diagnostic procedure was similar. After reviewing all available information, including data from medical records, other research assessments, and clinical interviews with the patient and relatives, diagnoses were assigned. Psychiatrists who assigned the clinical diagnoses in any of these three settings were blind to the study in process. Most of the diagnostic psychiatrists were trained psychiatrists with many years of experience in Mental Health Care Centres, whereas others were supervised residents still in training. The ICD is the diagnostic system of choice in Spain though most psychiatrists have good knowledge of DSM-IV system; psychiatry residents are trained to use ICD-10. Obviously, we cannot discard that some of the diagnostic psychiatrists would favour DSM and use ICD only because they have to.

## Variables

Diagnoses were made by treating psychiatrists/psychologists according to ICD-9 or ICD-10, depending on the assessment date. Treating clinicians had standard clinical training in diagnostic assessment and were hired by the National Mental Health System. Responsible psychiatrists/psychologists had an extended experience evaluating and treating patients with at least a required 4-year-time minimum experience. Psychiatrists/psychologists recorded a maximum of 2 diagnoses per patient per visit for administrative purposes and were blind to the study process.

**Diagnostic groups included in the statistical analysis**

In addition to bipolar disorder (ICD-10 F31) and manic episode (ICD-10 F30), 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 prevalence ≥ 1% in the whole sample.

The following table lists the bipolar disorder ICD-10 diagnoses that were included, diagnostic criteria can be seen in Appendix 3: ICD-10 manic episode and bipolar disorder diagnostic criteria, page 115.

| ICD-10 Psychiatric Diagnosis Code | ICD-10 Psychiatric Diagnosis |
|---|---|
| F30 | Manic episode |
| F30.0 | Hypomania |
| F30.1 | Mania without psychotic symptoms |
| F30.2 | Mania with psychotic symptoms |
| F30.20 | With mood-congruent psychotic symptoms |
| F30.21 | With mood-incongruent psychotic symptoms |
| F30.8 | Other manic episodes |
| F30.9 | Manic episode, unspecified |
| F31 | Bipolar affective disorder |
| F31.0 | Bipolar affective disorder, current episode hypomanic |
| F31.1 | Bipolar affective disorder, current episode manic without psychotic symptoms |
| F31.2 | Bipolar affective disorder, current episode manic with psychotic symptoms |
| F31.20 | With mood congruent psychotic symptoms |
| F31.21 | With mood incongruent psychotic symptoms |
| F31.3 | Bipolar affective disorder, current episode mild or moderate depression |
| F31.30 | Without somatic syndrome |
| F31.31 | With somatic syndrome |
| F31.4 | Bipolar affective disorder, current episode severe depression without psychotic symptoms |
| F31.5 | Bipolar affective disorder, current episode severe depression with psychotic symptoms |
| F31.50 | With mood-congruent psychotic symptoms |
| F31.51 | With mood incongruent psychotic symptoms |
| F31.6 | Bipolar affective disorder, current episode mixed |
| F31.7 | Bipolar affective disorder, currently in remission |
| F31.8 | Other bipolar affective disorders |
| F31.9 | Bipolar affective disorder, unspecified |

**Table 5. Bipolar disorder ICD-10 psychiatric diagnoses included in the analysis**

## *Analytic strategy*

### Diagnostic stability

Through all the evaluations diagnostic stability was calculated according to Schwartz et al. and Baca-Garcia et al. (12;17) with traditional statistical methods using version 13.0 of Spss (SPSS Inc., Chicago, IL, USA). Three complementary indices of diagnostic stability were used to increase the consistency of our results:

**1. Temporal consistency**

Temporal consistency is the presence or absence of a particular disorder at two different time points (58). We made use of three different measures of temporal consistency for bipolar disorder (17). The first, "prospective consistency", is the proportion of individuals in a category at the first evaluation who remain in the same category at their last evaluation. This would correspond to positive predictive value if the last diagnosis were the gold standard. It is clinically useful because it indicates the extent to which a diagnosis given at the initial evaluation will be present at the last evaluation, thus directing clinical treatment.

The second, "retrospective consistency", is the proportion of individuals with a diagnosis assigned at the last evaluation that had received the same diagnosis at the first evaluation. This is conceptually similar to sensitivity and as with prospective consistency high values indicate good temporal consistency of the diagnosis. Thus, if a diagnosis made by a clinician at the last evaluation -when more information has become available- coincides with the diagnosis given at the initial evaluation, it could be argued that the initial clinical presentation was adequately captured and diagnosed.

However, prospective and retrospective consistency rates fail to account for the fact that new cases may develop after initial presentation and other cases may remit (58), which is corrected by the use of the third measure of temporal consistency, the

kappa coefficient (195). The kappa coefficient is the agreement between diagnoses at first and last evaluations and measures the agreement correcting the effect of chance. We adopted the guidelines for the interpretation of kappa coefficients from Altman (116): <0.20, poor agreement; 0.21-0.40, fair agreement; 0.41-0.60, moderate agreement; 0.61-0.80 good agreement; and 0.81-1.00 very good agreement.

**2. Diagnostic constancy:**

Because prospective and retrospective consistency and the kappa coefficient rely only on two evaluations, they often fail to reflect the diagnostic process through multiple evaluations that is more characteristic of routine clinical practice (12). To capture this process, we also measured the proportion of patients who received the same diagnosis in at least 75% of the evaluations. From a clinical perspective, this measure would better assess the stability of the diagnoses throughout successive clinical encounters than the diagnostic information obtained at two distant time points (up to 13 years in our study). Subjects who received bipolar disorder diagnoses in at least 75% of evaluations were categorized as having a constant bipolar disorder.

**3. Probability of diagnostic change:**

We used First-order Markov Models to discern what diagnoses are more likely to be made in a next visit for patients previously diagnosed with a bipolar disorder (See Appendix 2: Markov Models, page 113 ) (196;197).

A First-order Markov Model represents a process in which the future is independent from the past and depends only on the current state (in this case, the current diagnosis). For making a prediction at time *t*, the relevant information is the state at time *t* (in this case, the diagnosis at time *t*) and no further information on how the process developed before time *t* is needed. The Markov model calculates the probabilities of

diagnostic change from one given diagnosis to the following diagnosis. See Hougaard P 2000 (198).

Markov model results can be interpreted to mean that: a) subjects who have received a diagnosis with high transition probability to the same diagnosis in Markov Models would have a high likelihood of receiving the same diagnosis in the next visit; conversely b) subjects who have received a diagnosis with low transition probability towards the same diagnosis in Markov Models would have a low likelihood of receiving the same diagnosis in the next visit.

## Misdiagnosis and comorbidity

The issues of misdiagnosis and comorbidity have been approached simultaneously in the present study. The frequent comorbidity in bipolar disorder (74;151;171;199) such as anxiety disorders, personality disorders and substance abuse adds a noteworthy complication to its accurate diagnosis, and explains a large proportion of its misdiagnosis (see Misdiagnosis, page 29). It should be stressed out that as a consequence of our endeavour to explore changes in the main diagnostic picture over time only main diagnoses were included in the study. Comorbid illnesses according to diagnostic guidelines should be recorded as auxiliary diagnoses when they are independent of the primary illness. Comorbidity denotes the joint occurrence of more somatic or psychiatric disorders with different pathophysiology in a single person, either simultaneously or on a lifetime basis (200). The use of auxiliary diagnoses in Spain is scarce but the inclusion of those figures might have altered partially our results on comorbidity.

Acknowledging this disadvantage, we investigated:

(i) the prevalence of principal psychiatric disorders in the sample from the total number of assessments;

(ii) the relationship between the frequency of psychiatric disorders and the constancy of bipolar disorder diagnosis comparing the number of specific diagnoses different to bipolar disorder in the 'stable' and 'not stable' subsamples.

## *Statistical analysis*

Wald's method (116) served us to compare the temporal consistency measures of bipolar disorder diagnoses and to calculate confidence intervals for each measure of temporal consistency (Statistical Package for the Social Sciences, version 14.0). We conservatively considered two confidence intervals that share a boundary or do not overlap to be significantly different from one another. We also compared the prevalence of different psychiatric diagnoses between those with and without a constant bipolar disorder diagnosis using Fisher´s Exact Test. To compare the subjects in the sample in regard to diagnostic constancy and gender we used $Chi^2$ tests. All these comparisons were performed two-tailed. The statistical analyses were conducted in two steps to search determinants of instability: univariate analyses followed by a multivariate analysis using logistic regression. Significance was assessed with chi square tests. The significant independent variables were then selected and introduced in logistic regression analyses with stability versus instability as the dependent variable. The Hosmer-Lemeshow goodness-of-fit test was also used. The clinical variables were thus introduced as independent variables in the univariate analyses, and significant variables were then introduced into a logistic regression model. This analysis was designed to determine which variables could help clinical psychiatrists determine when patients would be instable.

# RESULTS

## *Characteristics of the sample*

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.

The distribution of the sample by sex and age at first evaluation is represented in figure 5.

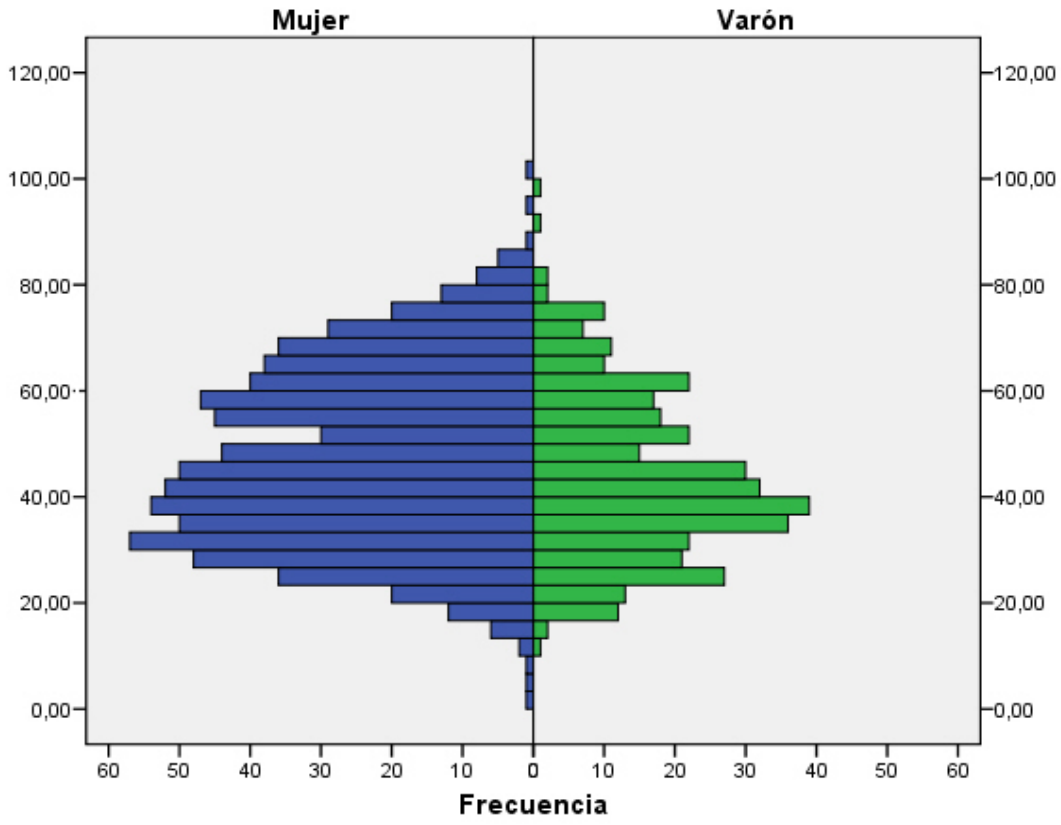



**Figure 7. Sex distribution of the sample by age at first evaluation**

## *Diagnostic frequencies*

During the study period there were 71 543 assessments of the 1153-patient sample. The most frequent diagnoses (≥ 5%) in order of importance along the study period and including the whole sample were:

1. Paranoid schizophrenia (F20.0; 12%, 8447/71 543);
2. Bipolar disorder (BD), current episode mild or moderate depression (F31.3; 11%, 8131/71 543);
3. Residual schizophrenia (F20.5; 11%, 7576/71 543);
4. BD, current episode manic without psychotic symptoms (F31.1; 10%, 7113/71 543);
5. Dysthymia (F34.1; 9%, 6314/71 543);
6. Major depressive disorder, recurrent (F33; 7%, 4855/71 543).

The 266 ‘stable BD’ patients showed a different spectrum of diagnoses along the study period, most of them included in the bipolar disorder categories. There were 13 148 assessments and the most frequent diagnoses are arranged here in order of importance:

1. BD, current episode mild or moderate depression (F31.3; 30%, 3896/13 148 consultations);
2. BD, current episode manic without psychotic symptoms (F31.1; 29%, 3767/13 148);
3. Other BD (F31.8; 15%, 2017/13 148); BD, most recent episode unspecified (F31.9; 9%, 1174/13 148);
4. BD, most recent episode mixed (F31.6; 7%, 981/13 148).

Diagnoses in the schizophrenic category appeared often in the 877 ‘non stable BD’. The most frequent diagnoses during the study period among all 58 395 assessments of these patients were:

1. Paranoid schizophrenia (F20.0; 14%, 8409/58 395);
2. Residual schizophrenia (F20.5; 13%, 7543/58 395);
3. Dysthymia (F34.1; 11%, 6205/58 395);
4. Major depressive disorder, recurrent (F33; 8%, 4691/58 395);
5. BD, current episode mild or moderate depression (F31.3; 7%, 4235/58 395);
6. BD, current episode manic without psychotic symptoms (F31.1; 6%, 3346/58 395).

Patients with ‘first diagnosis BD’ kept along the study period mostly diagnoses of bipolarity. Among all 342 patients included in this group 17 122 assessments were made, and the most frequent diagnoses were:

1. BD, current episode manic without psychotic symptoms (F31.1; 20%, 3430/17 122);
2. BD, current episode mild or moderate depression (F31.3; 20%, 3425/17 122);
3. Other BD (F31.8; 12%, 2100/17 122);
4. Dysthymia (F34.1; 6%, 1069/17 122); BD, most recent episode unspecified (F31.9; 6%, 984/17 122);
5. BD, most recent episode mixed (F31.6; 5%, 909/17 122).

On the contrary many of the most frequent diagnoses during the study period among all 54 421 assessments of the 811 ‘first diagnosis not BD’ patients were included in the F2 category:

1. Paranoid schizophrenia (F20.0; 15%, 7886/54 421);
2. Residual schizophrenia (F20.5; 13%, 7018/54 421);

3. Dysthymia (F34.1; 10%, 5245/54 421);
4. BD, current episode mild or moderate depression (F31.3; 9%, 4706/54 421);
5. Major depressive disorder, recurrent (F33; 8%, 4270/54 421);
6. BD, current episode manic without psychotic symptoms (F31.1; 7%, 3683/54 421).

'Last diagnosis BD' group is formed of 443 patients that had among all 22 117 assessments during the study period. The most frequent diagnoses in this group were again in the categories of:

1. BD, current episode mild or moderate depression (F31.3; 22%, 4817/22 117);
2. BD, current episode manic without psychotic symptoms (F31.1; 21%, 4589/22 117);
3. Other BD (F31.8; 10%, 2206/22 117); major depressive disorder, recurrent (F33; 7%, 1577/22 117);
4. BD, most recent episode unspecified (F31.9; 7%, 1504/22 117);
5. Dysthymia (F34.1; 6%, 1278/22 117).

Finally the group of 'last diagnosis not BD' included 710 patients and the most frequent diagnoses during the study period among all 49 426 assessments for these patients were:

1. Paranoid schizophrenia (F20.0; 16%, 7922/49 426);
2. Residual schizophrenia (F20.5; 14%, 7142/49 426);
3. Dysthymia (F34.1; 10%, 5036/49 426);
4. BD, current episode mild or moderate depression (F31.3; 7%, 3314/49 426);
5. Major depressive disorder, recurrent (F33; 7%, 3250/49 426);
6. BD, current episode manic without psychotic symptoms (F31.1; 5%, 2524/49 426).

| Diagnosis | Whole sample % (N) | Stable BD[a] % (N) | Non-stable BD[b] % (N) | Last diagnosis BD[c] % (N) | Last diagnosis not BD[d] % (N) | First diagnosis BD[e] % (N) | First diagnosis not BD[f] % (N) |
|---|---|---|---|---|---|---|---|
| F20.0 Paranoid schizophrenia | 11.8% (8447) | 0.3% (38) | 14.4% (8409) | 2.4% (525) | 16.0% (7922) | 3.3% (561) | 14.5% (7886) |
| F20.5 Residual schizophrenia | 10.6% (7576) | 0.3% (33) | 12.9% (7543) | 2.0% (434) | 14.5% (7142) | 3.3% (558) | 12.9% (7018) |
| F31.1 BD, current episode manic without psychotic symptoms | 9.9% (7113) | 28.7% (3767) | 5.7% (3346) | 20.8% (4589) | 5.1% (2524) | 20.0% (3430) | 6.8% (3683) |
| F31.3 BD, current episode mild or moderate depression | 11.4% (8131) | 29.6% (3896) | 7.3% (4235) | 21.8% (4817) | 6.7% (3314) | 20.0% (3425) | 8.7% (4706) |
| F31.6 BD, most recent episode mixed | 2.2% (1551) | 7.5% (981) | 1.0% (570) | 4.6% (1014) | 1.1% (537) | 5.3% (909) | 1.2% (642) |
| F31.8 Other BD | 4.6% (3288) | 15.3% (2017) | 2.2% (1271) | 10.0% (2206) | 2.2% (1082) | 12.3% (2100) | 2.2% (1188) |
| F31.9 BD, most recent episode unspecified | 3.5% (2518) | 8.9% (1174) | 2.3% (1344) | 6.8% (1504) | 2.1% (1014) | 5.8% (984) | 2.8% (1534) |
| F33. Major depressive disorder, recurrent | 6.7% (4855) | 1.0% (136) | 8.0% (4691) | 7.1% (1577) | 6.6% (3250) | 3.3% (557) | 7.9% (4270) |
| F34.1 Dysthymia | 8.8% (6314) | 0.8% (109) | 10.6% (6205) | 5.8% (1278) | 10.2% (5036) | 6.3% (1069) | 9.6% (5245) |
| Total assessments | 100.0% (71543) | 100.0% (13148) | 100.0% (58395) | 100.0% (22117) | 100.0% (49426) | 100.0% (17122) | 100.0% (54421) |

**Table 6. Diagnostic frequencies of the most common diagnoses**

[a]Stable BD= subjects who received the diagnosis of BD in at least 75% of the evaluations

[b]Non-stable BD= subjects who did not receive the diagnosis of BD in at least 75% of the evaluations

[c]Last diagnosis BD= subjects who received a diagnosis of BD at the last evaluation

[d]Last diagnosis not BD= subjects who did not receive a diagnosis of BD at the last evaluation

[e]First diagnosis BD= subjects who received a diagnosis of BD at the first evaluation

[f]First diagnosis not BD= subjects who did not receive a diagnosis of BD at the first evaluation

| | |
|---|---|
| Bipolar disorder | 31,6% |
| Schizophrenia | 23,9% |
| Anxiety disorders | 9,1% |
| Chronic depression | 6,8% |
| Personality disorders | 3,7% |
| Substance abuse | 2,2% |
| Psychosomatic disorders | 0,8% |
| Depressive episode | 0,5% |

**Table 7. Diagnostic frequencies grouped by Axis I categories.**

### Number of assessments

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

## *Temporal consistency of bipolar disorder diagnoses*

### Overall sample

In the first visit 342 patients were diagnosed a bipolar disorder with a prospective consistency of 49.4%. In the last visit 443 patients were diagnosed a bipolar disorder with a retrospective consistency of 38.1%. Remarkably the Kappa value was low between first and last diagnosis, Kappa=0.40. Origin of data (out-patient clinics, in-

patient unit or urgencies) as well as length of follow-up produced no significant difference.

| | | CI % |
|---|---|---|
| First diagnosis BD – n | 342 | |
| Prospective consistency | 49.4% | 44.1-54.7 |
| Last diagnosis BD – n | 443 | |
| Retrospective consistency | 38.1% | 33.7-42.7 |
| Kappa value[1] | 46.7 | 41.4-52.0 |

**Table 8. Temporal consistency of ICD-10 bipolar disorder**
[1]Kappa (κ) statistics are significant (P<0.001).

Over the follow-up we found a great variability in the diagnostic categories. There is a wide range of categories that could cause confusion in the diagnosis of bipolar disorder. We discovered this to be especially clear with schizophrenia spectrum (F2), diagnosis that appears in one of every four visits to the psychiatrists of the patients included in this study. In a lower degree there are three other diagnostic categories that may contribute to misdiagnosis of bipolar disorder: anxiety disorders (F4), personality disorders (F6) and substance abuse disorders. Nevertheless very few patients reached diagnosis stability criteria of our study for schizophrenia or any other psychiatric disorder excepting bipolar disorder.

Figure 8 shows the F3 category (affective disorders) broken down in the different subcategories, so that one can observe that numerous patients were diagnosed in the F33 category (depressive recurrent disorder) and even with some frequency they were classified in this manner also in their last assessment. Patients included in this category did not achieve stability apart from a few cases. The figure shows additionally

the graphic prevalence of diagnoses in other categories. F2 category prevalence corresponds basically to the diagnosis of paranoid schizophrenia.

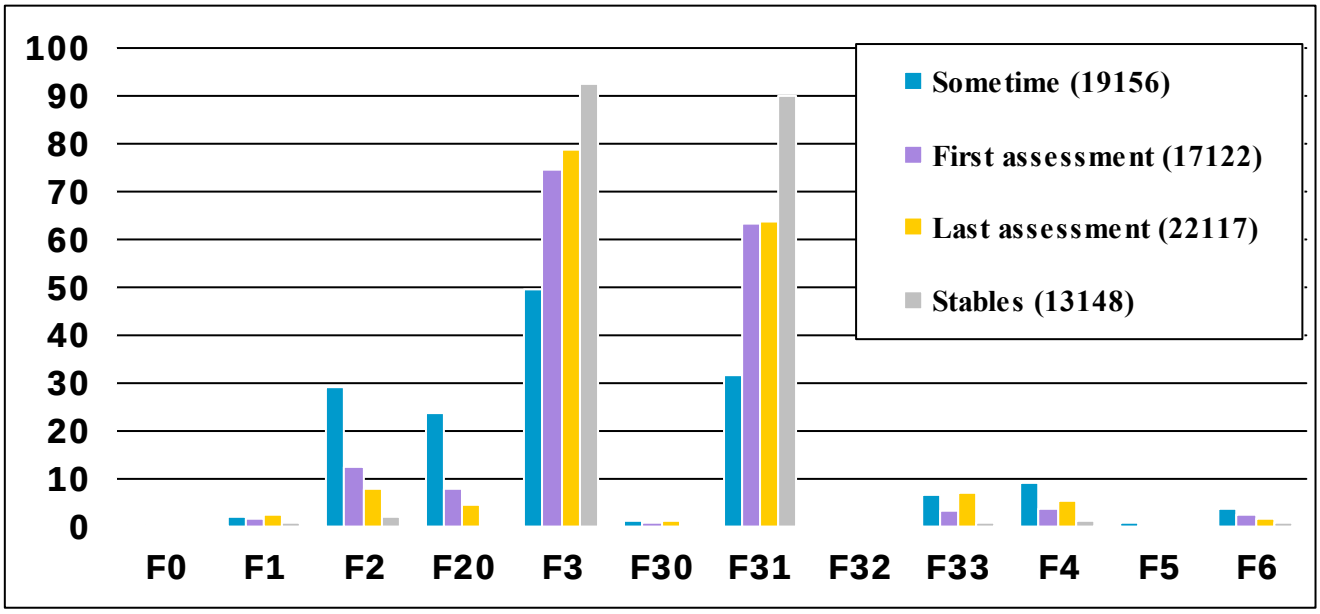



**Figure 8. Diagnostic frequencies of the most common diagnoses**

## *Diagnostic constancy of bipolar disorder diagnoses*

### Overall sample

From the total sample of 1153 patients that were at least in one occasion diagnosed as bipolar disorder, we found that 23.1% (n = 266/1153) confirmed the diagnosis in 75% of the assessments. This analysis was performed with the joint data from the three clinical settings, intending to reflect the evolution of diagnoses through the clinical process. Out of 342 initial diagnosis of bipolar disorder, only 158 kept that diagnosis in at least 75% of the assessments along the follow-up period, meaning a 46.1% of the 266 patients considered stables (n = 158/342, IC95%: 40.9-50.1). 108 patients were found to be stable but were not diagnosed in the first assessment and would constitute an initial error (31%, n = 108/266; IC 95%: 26.6-36.5). In the first consultation 184 out of 342 patients thought to have a bipolar disorder obtained later on

at least a 25% of diagnosis different from F30/F31 and could be considered as initial over diagnosis or false positives (FP), corresponding to 53% of initial diagnosis (n = 184/342; IC95%: 48.5-59.0).

In the last assessment the number of bipolar disorder diagnosis increased steeply, 234 out of 443 subjects included in a F30/F31 category in the last consultation correspond to the 266 stable patients along the follow-up (88%). Nevertheless 209 patients (47.2%, n = 209/443; IC95% 42.5-51.8) finally diagnosed as bipolar disorder in this last visit did not keep stability criteria in their evolution and could be considered final over diagnosis (FP). We consider that 32 patients not diagnosed in this assessment would constitute the final diagnostic error (12%, n = 32/266; IC95%: 8.1-15.9%).

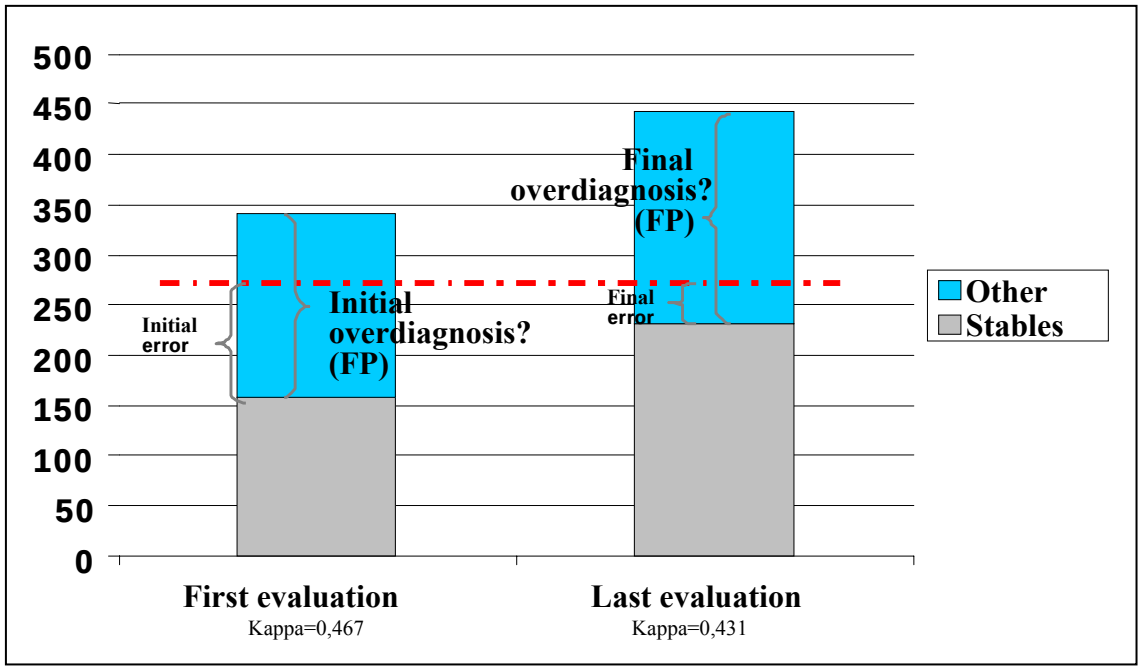



**Figure 9. Comparison between bipolar disorder diagnosis in first and last consultation. Red line represents the number of patients with at least 75% bipolar disorder diagnosis along the follow-up.**

Figure 9 shows the difference between first and last evaluation, the number of 'stable' patients (with over 75% of assessments with a bipolar disorder diagnosis) and the existence of diagnostic errors in these evaluations. 'Initial overdiagnosis' is presented as the excess of bipolar disorder diagnosis that was given to patients found to

be ‘stable’ along the study period. ‘Initial error’ represents the extra number of ‘stable’ patients that should have been diagnosed in the first consultation to reach the number of patients found to be ‘stable’ along the study. ‘Final overdiagnosis’ are the patients who were given a diagnosis of bipolar disorder in their last consultation not reaching the criteria for ‘stability’. ‘Final error’ shows the gap between the ‘stable’ patients that confirmed their diagnosis in the last consultation and the total number of ‘stable’ patients along the study.

Among the 266 bipolar ‘stable’ 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 bipolar disorder was 2.1 (seventh percentile of the total number of evaluations). The median was 1.0 assessment (fifth percentile). All 266 ‘stable’ bipolar patients had received the diagnosis of bipolar disorder at the 33$^{rd}$ percentile of the total number of evaluations.

Among the 887 ‘non stable’ bipolar 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 bipolar disorder was 22.6 (38$^{th}$ percentile of the total number of evaluations). The median was 9.0 (31$^{st}$ percentile). The proportion of patients who received the diagnosis of bipolar disorder 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 bipolar disorder was significantly different in the ‘stable’ and ‘non stable’ bipolar disorder groups (Mann-Whitney’s $U = 43231.5$; $P < 0.001$).

## Factors related with stability

In the first step of statistical analyses, the univariate analyses, odds ratios (ORs) and 95 percent confidence intervals (CIs) were calculated with a dichotomous

dependent variable (stability versus non stability). Significance assessed with chi square tests showed no association with the following variables: marital status, type of cohabitation, socioeconomic level and educational level. The significant independent variables were then selected and introduced in logistic regression analyses with stability versus instability as the dependent variable. 4 variables remained significant, as can be seen in Table 9: gender, age ≥40 years, number of psychiatric consultations, and outpatient mental health centre.

| **Variables included in the logistic regression** | | **Wald** | **df** | **p** | **OR** | **C.I. 95,0%** | |
|---|---|---|---|---|---|---|---|
| | | | | | | Inferior | Superior |
| **Gender** | **Male/ female** | 5,796 | 1 | ,016 | 1,482 | 1,076 | 2,041 |
| **Edad** | **<40/>39** | 2,880 | 1 | ,090 | 1,295 | ,961 | 1,745 |
| **Number of assessments** | | 19,641 | 4 | ,001 | | | |
| | **>78/1-16** | 1,891 | 1 | ,169 | 1,374 | ,873 | 2,163 |
| | **>78/17-26** | 9,650 | 1 | ,002 | 2,019 | 1,296 | 3,146 |
| | **>78/27-43** | ,181 | 1 | ,671 | ,903 | ,562 | 1,448 |
| | **>78/44-78** | ,200 | 1 | ,655 | ,897 | ,557 | 1,445 |
| **Mental Health Care Centre** | **Cent1/Cent2** | 24,082 | 1 | ,000 | 2,065 | 1,546 | 2,758 |
| | **Constant** | 84,727 | 1 | ,000 | ,119 | | |

**Table 9. Significant high scores in a logistic regression to measure the factors involved in instability of bipolar disorder.**

Hosmer and Lemeshow test $\chi^2 = 7,847$, df = 8, p = 0,449.

These results show that the main risk of instability was associated with the outpatient health centre and in a lower degree with the number of assessments (compared with less than 26 visits). Males showed as well an increased risk of instability compared to females (See Diagnostic stability, page 90).

## Sample stratified by sex

The sample was composed of 771 females and 382 males. A significant gender difference was observed on the diagnostic constancy of bipolar disorder diagnosis (two tailed Fisher Exact Test = 0.026).

| | | Non stable | Stable | Total |
|---|---|---|---|---|
| Females | Number | 578 | 193 | 771 |
| | % | 65,2% | 72,6% | 66,9% |
| Males | Number | 309 | 73 | 382 |
| | % | 34,8% | 27,4% | 33,1% |

**Table 10. Diagnostic constancy of ICD-10 bipolar disorder by gender**

## Misdiagnosis with other affective disorders

One of the most frequent causes of confusion in the diagnosis of bipolar disorder, as seen before, are the diverse diagnoses in F3 category, i.e. the affective disorders not considered to be bipolar (see Misdiagnosis, page 29).

Table 11 details the frequencies of these diagnoses depending both on the gender and the stability of bipolar disorder. A positive association is found between instability and the number of diagnoses of affective disorders different to bipolar disorder.

| F3 diagnosis not bipolar | N | % on the total | Non stable | % Non stable | Stable | % Stable | FET[1] |
|---|---|---|---|---|---|---|---|
| Total | 683 | 59,2% | 608 | 68,5% | 75 | 28,2% | 0,000 |
| Men | 206 | 53,9% | 185 | 61,3% | 21 | 26,3% | 0,000 |
| Women | 477 | 61,9% | 423 | 72,3% | 54 | 29,0% | 0,000 |

**Table 11. Frequencies of F3 not bipolar diagnoses and gender differences**

[1] Fisher Exact Test (2-tailed)

Figure 10 represents graphically the proportion of depression diagnoses in the sample of 1153 patients with diagnosis of bipolar disorder. Just about 58% of patients present in not as much as 10% of the assessments a diagnosis of depression.

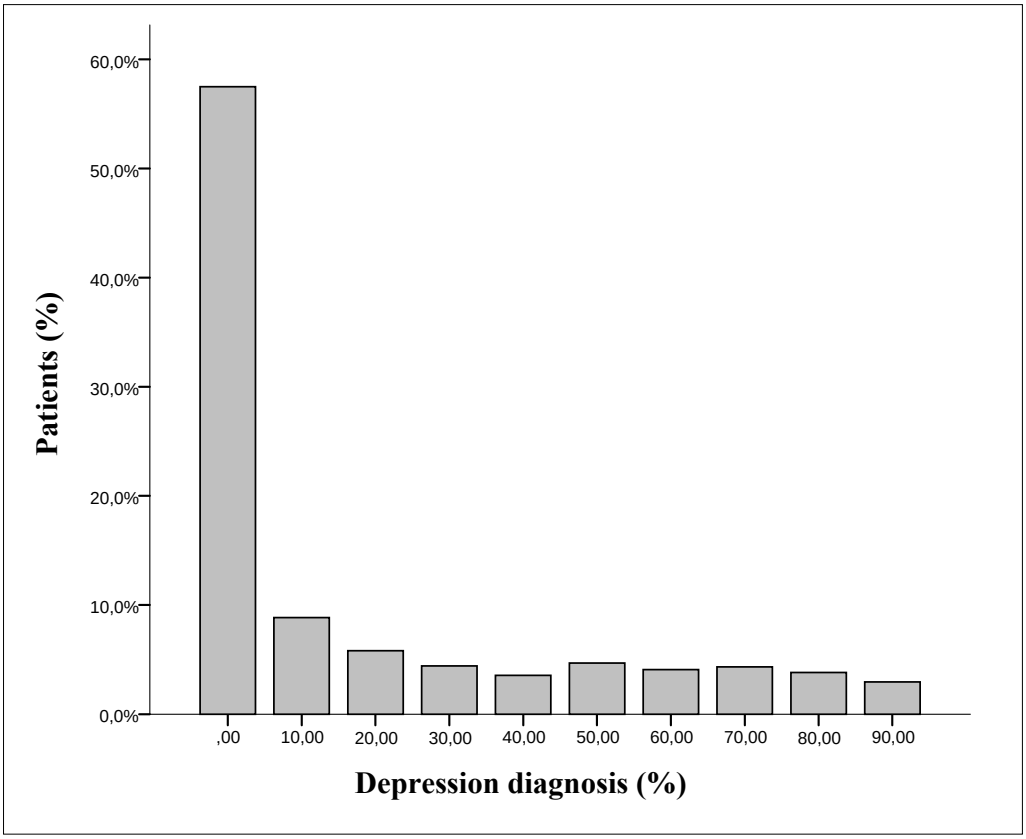



**Figure 10. Percentage of depression diagnoses in the sample**

## Misdiagnosis with other psychiatric diagnoses (comorbidity)

There are several other diagnoses in different categories that interfere with bipolar disorder stability. Some of them could be more properly considered comorbid entities, though as explained elsewhere (see Misdiagnosis and comorbidity, page 67) the discrimination has not been undertaken in the present study.

| | N | % on the total | Not stable | % not stable | Stable | % stable | FET[1] |
|---|---|---|---|---|---|---|---|
| F0 Organic, including symptomatic, mental disorders | 63 | 5,5% | 49 | 5,5% | 14 | 5,3% | 1,000 |
| F1 Mental and behavioural disorders due to psychoactive substance use | 140 | 12,1% | 120 | 13,5% | 20 | 7,5% | 0,007 |
| F2 Schizophrenia, schizotypal and delusional disorders | 381 | 33,0% | 330 | 37,2% | 51 | 19,2% | 0,000 |
| F4 Neurotic, stress-related and somatoform disorders | 572 | 49,6% | 510 | 57,5% | 62 | 23,3% | 0,000 |
| F5 Behavioural syndromes associated with physiological disturbances and physical factors | 61 | 5,3% | 56 | 6,3% | 5 | 1,9% | 0,003 |
| F6 Disorders of adult personality and behaviour | 218 | 18,9% | 200 | 22,5% | 18 | 6,8% | 0,000 |
| F7 Mental retardation | 10 | 0,9% | 8 | 0,9% | 2 | 0,8% | 1,000 |
| F8 Disorders of psychological development | 1 | 0,1% | 0 | 0,0% | 1 | 0,1% | 1,000 |
| F9 Behavioural and emotional disorders with onset usually occurring in childhood and adolescence | 23 | 2,0% | 20 | 2,3% | 3 | 1,1% | 0,323 |

**Table 12. Frequencies of not-F3 diagnoses related to the stability of BD**

[1] Fisher Exact Test (2-tailed)

Table 12 provides the number of diagnoses given in not-F3 categories and their relation with the stability of bipolar disorder. The frequencies of diagnoses in the categories of substance abuse, personality disorders, schizophrenia and anxiety disorders exhibit a significant statistical association with the instability of bipolar disorder diagnosis.

| F2 Schizophrenia, schizotypal and delusional disorders | N | % on the total | Non stable | % Non stable | Stable | % Stable | FET[1] |
|---|---|---|---|---|---|---|---|
| Total | 381 | 33,0% | 330 | 37,2% | 51 | 19,2% | 0,000 |
| Men | 161 | 42,1% | 141 | 46,7% | 20 | 25,0% | 0,001 |
| Women | 220 | 28,5% | 189 | 32,3% | 31 | 16,7% | 0,000 |

**Table 13. Frequencies of F2 diagnoses and gender differences**

[1] Fisher Exact Test (2-tailed)

Table 13 details the frequencies of F2 diagnoses depending both on the gender and the stability of bipolar disorder. Together with the association between instability of bipolar disorder and the number of schizophrenia spectrum diagnoses, differences between genders and the quantity of schizophrenia diagnoses are observed.

## *Probability of diagnostic changes*

Four Markov's models were calculated to study the evolution of diagnosis in bipolar disorder (See Appendix 2: Markov Models, page 113). Each pixel in Figure 11-Figure 14 represents the probability of a transition between the 'prior' diagnostic stage (e.g. F10) and the 'next' diagnostic stage (e.g. F20). In the pictures, the y-axis represents 'prior' diagnostic states, and the x-axis represents 'next' diagnostic states. The inventory of diagnostic states included can be seen on Table 14. The first model,

which incorporated the whole sample, is represented in Figure 11. 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 Figure 11.

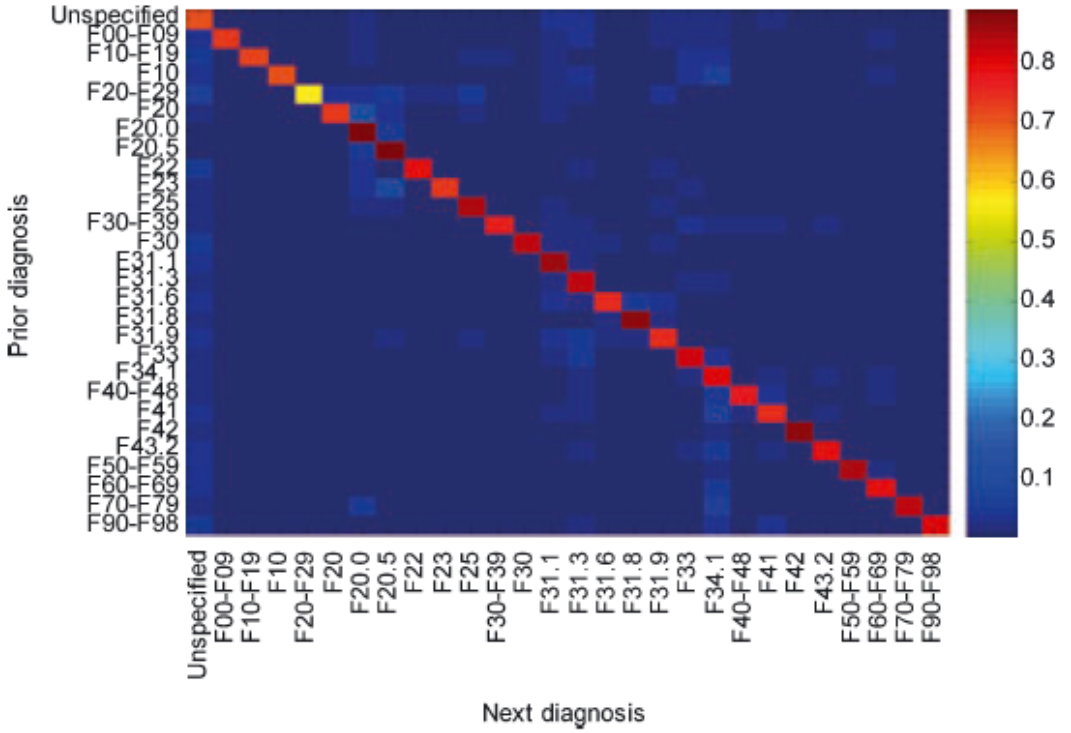



**Figure 11. Probability of transitions between 'prior' diagnoses and 'next' diagnoses in the whole sample (n = 1153). Highest probabilities are concentrated on the red-coloured diagonal**

| | |
|---|---|
| **F00–F09** | Organic, including symptomatic, mental disorders |
| **F10–F19** | Mental and behavioural disorders due to psychoactive substance use |
| **F10** | Mental and behavioural 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 mild or moderate depression |
| **F31.6** | Bipolar disorder, most recent episode mixed |
| **F31.8** | Other bipolar affective disorders |
| **F31.9** | Bipolar disorder, most recent episode unspecified |
| **F33** | Major depressive disorder, recurrent |
| **F34.1** | Dysthymia |
| **F40–F48** | Neurotic, stress-related and somatoform disorders |
| **F41** | Other anxiety disorders |
| **F42** | Obsessive–compulsive disorder |
| **F43.2** | Adjustment disorders |
| **F50–F59** | Behavioural syndromes associated with physiologic disturbances and physical factors |
| **F60–F69** | Disorders of adult personality and behaviour |
| **F70–F79** | Mental retardation |
| **F90–F98** | Behavioural and emotional disorders with onset usually occurring in childhood and adolescence |

**Table 14. Diagnostic states assigned in Markov's models**

The second model includes the patients who have received the diagnosis of bipolar disorder in at least 75% of the evaluations ('stable BD' group). These patients have consistently been assigned the diagnosis of bipolar disorder by the clinicians who have assessed them so that they can be regarded as 'true' bipolar patients. That makes this model the most appealing. We can see on Figure 12 that the most probable transitions across diagnostic blocks in this group were from other diagnoses to bipolar

disorder, but not from bipolar disorder to other diagnoses. Once the patients receive a stable diagnosis of bipolar disorder they do not switch to any other diagnostic block, in other words there is an absence of diagnostic change from 'stable' bipolar disorder. These results indicate that discordant diagnoses are given before those patients have reached a stable diagnosis of bipolar disorder. The most probable transitions across diagnostic blocks were:

1) from F10–F19, mental and behavioural disorders due to use of psychoactive substances, to F31.1 bipolar disorder, current episode manic without psychotic symptoms;
2) from F20, schizophrenia, to F31.1 bipolar disorder, current episode manic without psychotic symptoms;
3) from F23, acute and transient psychotic disorders, to F31.9 bipolar disorder, most recent episode unspecified;
4) from F42, obsessive–compulsive disorder, to F31.1 BD, current episode manic without psychotic symptoms;
5) from F60–F69, disorders of adult personality and behaviour, to F31.1 bipolar disorder, current episode manic without psychotic symptoms;
6) from F90–F98, behavioural and emotional disorders with onset usually occurring in childhood and adolescence, to F31.8, other BD;
7) from F90–F98, behavioural and emotional disorders with onset usually occurring in childhood and adolescence, to F41, other anxiety disorders.

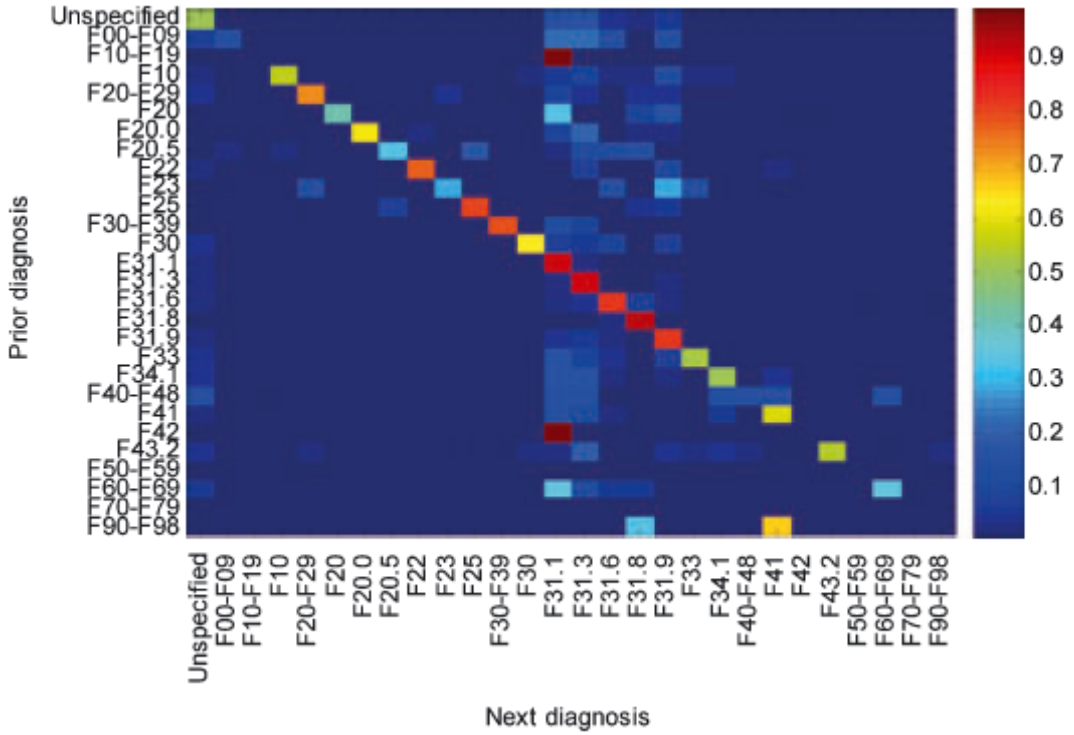



**Figure 12. Probability of transitions between prior diagnoses and next diagnoses in the 266 'stable bipolar affective disorder (BD)' patients (patients who have received the diagnosis of BD in ≥75% of the evaluations)**

The third model (Figure 13) was made including all the patients who had received the diagnosis of bipolar disorder at their last evaluation (last diagnosis bipolar disorder). The outcome is similar to the first model being the most probable transitions within the same diagnosis. Finally, the fourth model (Figure 14) included patients who had received the diagnosis of bipolar disorder at the first evaluation (first diagnosis bipolar disorder) and the results are similar to those of first and third models, showing the most probable transitions within the same diagnosis. In this model, time between stage changes (<1 month or ≥1 month) was incorporated to the analysis.

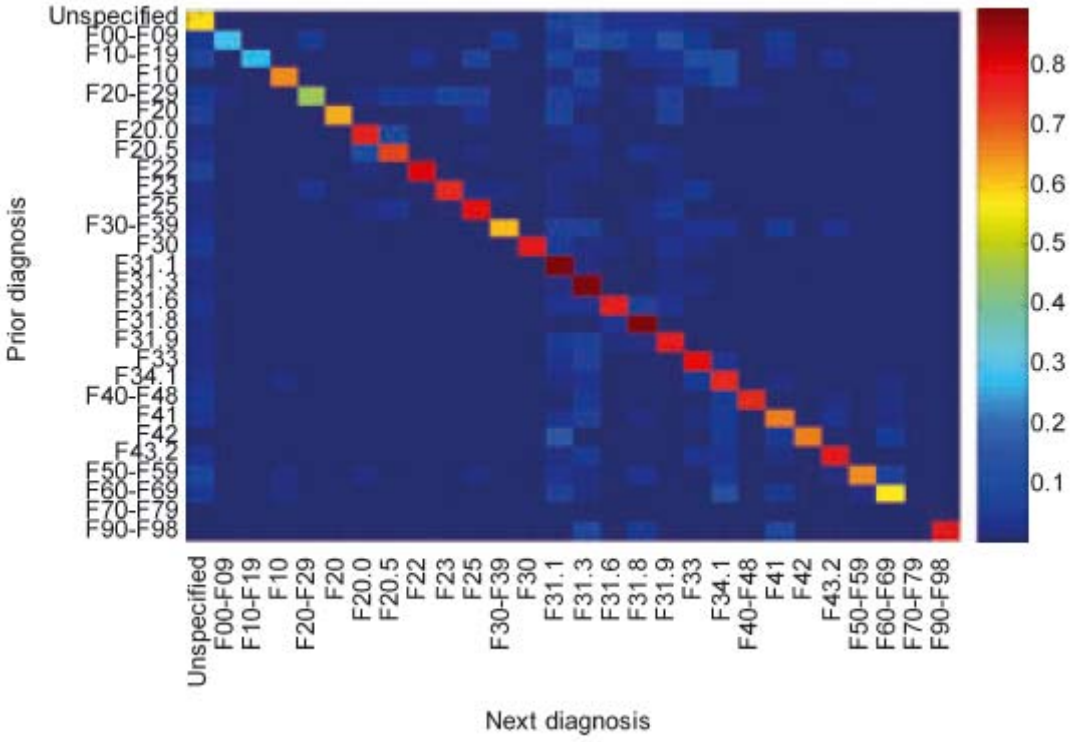



**Figure 13. Probability of transitions between prior diagnoses and next diagnoses of the 443 'last diagnosis bipolar affective disorder (BD)' patients (patients who received the diagnosis of BD at the last evaluation)**

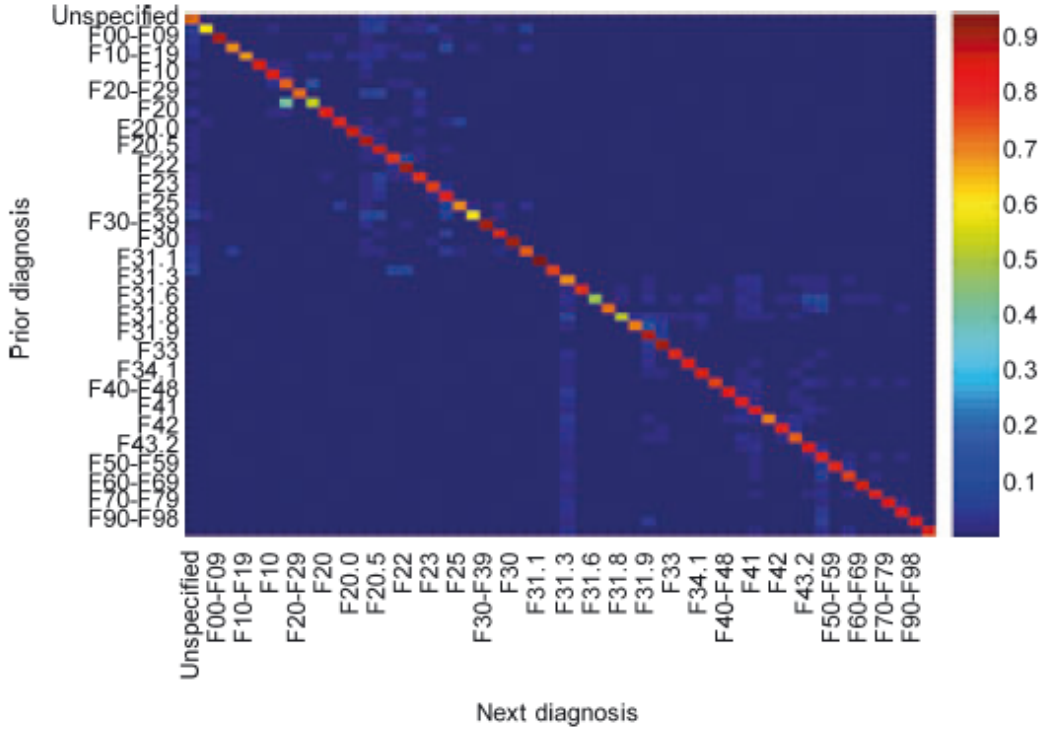



**Figure 14. Probability of transitions between prior diagnoses and next diagnoses of the 342 'first diagnosis bipolar affective disorder (BD)' patients (patients who received the diagnosis of BD at the first evaluation) including time between stages**

# DISCUSSION

The natural evolution of bipolar disorder is prone to a high variability of its clinical course, characterised by a pattern of manic and depressive episodes surrounded by euthymic periods. However, the easily recognizable core symptoms are not so often present and frequent comorbid conditions delude the clinicians in daily practice (10;119;201). In our study stability of bipolar disorder diagnosis was found to be low, and lower than in previous studies (7;14;17;22;23), with the three different indices used. Only 30% of the patients received the diagnosis of bipolar disorder in the first evaluation and only 23.1% of the total sample was considered 'stable' according to the criteria established in this study. Additionally we found diagnostic fluctuation involving the habitual diagnoses included in the differential diagnosis of bipolar disorder. All these findings are detailed and commented upon in the following sections.

## *Bipolar disorder diagnosis at the first evaluation*

Diagnostic errors in bipolar disorder are especially frequent in the first contact with the clinician, misleading initial symptoms of a 'masked' presentation due to substance abuse, depressive or transient psychotic symptoms may explain partly these difficulties. While only 30% of subjects received the diagnosis of bipolar disorder at the first evaluation, 70% got the diagnosis during later contacts. Similar findings were obtained by Kessing (14) who pointed out that these figures are consistent with the high prevalence of misdiagnosis of 48% and 69% found in naturalistic investigations using self-administered questionnaires on contact to doctors in general (79;80).

However in our sample only 46.1% of the 'stable' patients got the diagnosis of bipolar disorder in the first assessment which contradicts the figure reported by Chen (7) who signaled that about 70% of subjects with an initial bipolar disorder diagnosis

did not change to a different one. Our results confirm the problematic diagnosis of bipolar disorder in the first assessments (see Temporal consistency of bipolar disorder diagnoses, page 74).

## *Bipolar disorder diagnosis at the last evaluation*

Last assessment showed a major increase in the number of bipolar disorder diagnoses (38% of the sample) though 47.2% had not been 'stable' along the study. 88% of the 'stable' patients were accurately diagnosed in their last consultation.

This outcome may reflect a progressive increase in the stability of the diagnosis along the evaluations (in our case in a minimal number of ten) which is congruent with the idea that routine rediagnosis could improve chances of successful diagnostic process. However, Schwartz et al. (17) reported that the retrospective consistency of bipolar disorder was 85% when comparing 6- and 24-month diagnoses, but lowered to 73% when comparing baseline and 24-month diagnoses. This would mean that the rates of consistency for some diagnoses decreased as the follow-up period increased. At any rate and compared with Schwartz's data, the retrospective consistency of bipolar disorder across clinical settings in this study (38%) is strikingly low. A structured interview, the Structured Clinical Interview for DSMIII-R (SCID) provided DSM-III-R psychiatric diagnoses in the study by Schwartz et al. Perhaps the use of semistructured interviews would have enhanced reliability and therefore stability.

## *Diagnostic stability*

To our knowledge this has been the largest longitudinal study evaluating the diagnostic stability of bipolar disorder through the use of three complementary indices. Kessing (29) recently mentioned that no study had investigated the diagnostic stability for the most common ICD-10 psychiatric diagnoses given under ecological clinical

conditions. This is the case of our study, which has shown a low stability of ICD-10 bipolar disorder categories as measured by their temporal consistency, diagnostic constancy, and probability of diagnostic change; and considerably lower than in previous studies. The reasons for these differences in diagnostic temporal stability are unclear, but may be due to the large sample size, extensive duration of follow-up, high number of assessments, diagnostic criteria, or socio-demographic variables.

Temporal consistency showed low results with prospective consistency of 49% and retrospective consistency of 38.1%. It is worth pointing out that Kappa value was low (Kappa=0.40) between first and last diagnosis. Still, given that kappa values take into account stable positive cases and stable negative cases but also remitted cases and new cases, low kappa values may be observed if a high number of new or remitted cases occurred (46) and thus not necessarily reflect lack of diagnostic stability.

The findings in our study show that only 23.1% of the patients held the diagnosis of bipolar disorder in 75% of the assessments. The mean number of evaluations till the first diagnosis of bipolar disorder was 17.9, and this number was increased to 22.6 within the 'non stable' group (see Diagnostic constancy of bipolar disorder diagnoses, page 76). These results might be in consonance with previous reports by Hirschfeld in 2003 and Baldesarini in 1999 informing of a delay to correct diagnosis that could be in many cases around 8-10 years from the onset of the illness (79;84).

In a study on the reliability of Best Estimate Diagnosis, Roy et al (99) suggested that it is possible to identify cases that are more likely to lead to diagnostic disagreements, and signalled mixed psychotic and affective symptoms, shorter duration of illness, less certainty of diagnosis, and poorer quality of information as factors associated with poor reliability. In the case of bipolar disorder Chen et al (7) named gender, ethnicity and substance abuse/dependence as having a prominent role in the

diagnostic changes. The variables associated with the stability of bipolar disorder diagnosis were not a primary aim of our study, but the analysis was performed nevertheless as a mean of ascertaining future directions. We found four variables related to the diagnostic stability of the illness: gender, age ≥40 years, number of psychiatric consultations, and the out-patient mental health centres.

The association between gender and stability (women showing more frequent stability) could be explained by the most usual confounding factor in each gender as found in our study, schizophrenia in the case of men and depression in the case of women, accepting that depressive diagnoses would facilitate diagnostic shift to bipolar disorder. The widely reported (202-204) higher prevalence of drug abuse in men may as well influence this finding. Age would contribute to the constancy of bipolar disorder diagnosis through the stabilisation of symptomatology and the disposal of better knowledge on the clinical history of the patient, assuming that cases with longer durations of illness yield more clinical information. Criterion and observation variance as described by Spitzer may explain differences found between psychiatric care centres (5). Finally, the contradictory association with the number of assessments does not confirm our previous hypothesis on the importance of longitudinal diagnosis. Patients assessed less than 26 times in the period of study were more likely to attain a 'stable' diagnosis of bipolar disorder than patients with more than 78 visits. This detection may reflect the existence of cases that show an especially difficult diagnosis so explaining the high number of assessments and the persistence of instability. In any case further studies focused on these variables are needed to confirm or reject these findings (see Factors related with stability, page 78).

The fact that treating psychiatrists/psychologists often had access to past records and diagnoses, could turn out on an inclination to keep the previous diagnosis rather

than assign a different one. However, this possible bias is not supported by the strikingly low values of diagnostic temporal stability of bipolar disorder in our study, parallel to the values found for other chronic mental disorder diagnoses using similar methodology in an adult sample treated by the same team of psychiatrists and psychologists (12).

The higher rates of consistency found by other authors (7;14;17;22;23) may have been influenced by a number of drawbacks that decrease the generalizability of these studies: (i) most studies that have evaluated the stability of bipolar disorder have shorter follow-up periods than the present study; (ii) data used in most of the studies is obtained on a single clinical setting (mainly the in-patient setting); (iii) there is a scanty number of assessments in most of the studies.

## *Diagnostic change*

Patients with a stable diagnosis of bipolar disorder ('stable' group) presented some diagnostic fluctuation involving the typical diagnoses included in the differential diagnosis of bipolar disorder. The disorders that presented the highest probability of transition to bipolar disorder according to the second Markov's model were: mental and behavioral disorders due to use of psychoactive substances, schizophrenia, acute and transient psychotic disorders, obsessive–compulsive disorder (a result that was unexpected), personality disorders, and behavioural and emotional disorders with onset usually occurring in childhood and adolescence.

The probabilities of diagnostic shift found in the 'stable' group could be explained as a result of the comorbidity with drug abuse which may overlap bipolar disorder confounding diagnosis in earlier assessments and subsequently corrected. On the contrary schizophrenia spectrum disorders do more likely correspond with the initial

misdiagnosis of bipolar disorder. The clinical picture in cases of mixed psychotic and affective symptoms becomes clearer with time. Indeed, in many cases the onset of bipolar disorder is predominantly psychotic (205;206), with the affective phenomena often becoming clearer over time. Still the second Markov's model demonstrates that the number of diagnostic changes from 'stable' bipolar disorder is scarce, so that discordant diagnoses are given before those patients have reached a stable diagnosis of bipolar disorder (see Probability of diagnostic changes, page 83).

On the contrary a high proportion in the 'non stable' group received the diagnoses of paranoid schizophrenia (14%), residual schizophrenia (13%), dysthymia (11%), and major depressive disorder (8%). Similar findings were observed in the 'last diagnosis not bipolar disorder' and 'first diagnosis not bipolar disorder' groups. A significantly high number of patients (68%) with diagnosis of affective disorders (bipolar disorder excluded) is located among the 'non stable' group, a finding that could indicate that other affective disorders are the main confounding factor for diagnostic stability of bipolar disorder in a greater degree than other axis I categories. This is congruent with the results seen on Figure 10 (page 81) where more than 20% of the patients had a diagnosis of depression in more than 50% of the assessments. Previous reports coincide explaining the high rates of misdiagnosis out of the confusion with unipolar depression (79;80).

Gender is one potential source of bias. A remarkable gender difference is found in the 'non stable' group where women show a higher percentage of affective (not bipolar) diagnoses compared to men, and men show a higher percentage of schizophrenia (see Table 11, page 80, and Table 13, page 83). This difference may point out that between genders there are different diagnoses involved in the misdiagnosis of bipolar disorder.

## *Prevalence of bipolar disorder*

A significant difference between our study and previous DSM-based studies is that bipolar II subtypes may be included in the sample given that ICD-10 does not discriminate explicitly between bipolar disorder I and II. Following the ICD-10 definition for bipolar disorder, there must be at least two mood episodes among which at least one is a hypomanic or a manic episode (117). However this difference should have increased the prevalence of bipolar disorder compared to other studies, which is not the case.

Diagnostic stability plays a major role in the accurate estimation both of prevalence and incidence. Whereas incidence is a measure of risk, prevalence is influenced by episode duration (prognosis) and by mortality (105). Ideally, it should be possible to classify associations that are observed in epidemiologic prevalence data according to their main determinants: incidence and episode duration (207); an intermediate measure to attain is the precise description of the diagnostic stability of the disorder.

Unexpectedly we found that the prevalence of bipolar disorder in this psychiatric sample was low. In view of the total number of patients assessed about 3.35% had a diagnosis of bipolar disorder (n = 1153/34 368) and taking into consideration the total population of the area, only 0.38% ever had a diagnosis of bipolar disorder (n = 1153/300 000) in the study period. These figures are smaller than those found in most other studies performed on general and psychiatric population (72;73;127;131;148;208;209) though not all (102;126). Our results on prevalence are comparable to those of Perala et al (2007) who recently reported that the National Hospital Discharge Register was the most reliable means of screening for psychotic and bipolar disorder and found a lifetime prevalence of 0.2% for bipolar disorder. In this

paper difference with population surveys was explained on the basis of possible false positive results in the structured interviews like Composite International Diagnostic Interview (CIDI) compared to the use of multiple sources of information (102). Studies such like ours where patients are followed over long periods and across several settings, are closer to this approach than clinical trials based on diagnostic schedules and interviews performed in a research unit over a short period or large cross-sectional epidemiological studies based on a single assessment.

The use of a minimal number of ten assessments as a criterion to be included in the sample may have diminished the number of bipolar patients though it seems highly improbable that a patient actually suffering from a bipolar disorder would not have consulted at least ten times in the study period. Even taking account of the individuals diagnosed and treated in private consultations which would not be included in the sample, these are considerably low figures given that the Spanish psychiatric services are easily accessible by individuals in the community.

## *Strengths and weaknesses*

The main strengths of this study are the large representative sample, the length of follow-up (up to 12 years, mean = 6.2 years), and the high number of assessments (median = 34). It should be noted as well that the evaluation of bipolar disorder stability was made in three different clinical settings so that the diagnostic procedure was in agreement with the regular clinical practice. 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 to enhance reliability. The results of our study were based on the use of ICD-10 diagnoses established clinically and though reliability may have been affected, they possibly

reflect more accurately the real use of diagnostic classifications in psychiatric practice and may be more helpful in estimating the clinical utility of current psychiatric classification systems. The doctor sets a diagnosis based on the criteria of the ICD-10 system, but the validity of these criteria is in daily practice probed by the phenomenological information gathered about the patient in addition to the diagnostic criteria (210).

The limitations of our study are common to most large-scale surveys and inherent to a retrospective naturalistic study, which was performed in real world conditions (211). Structured or semi-structured clinical interviews were not used in this study for the assessment of ICD-10 bipolar disorder. The clinicians who assigned the diagnoses were not specifically trained to increase inter-rater reliability. Improved inter-rater reliability would have been likely to increase diagnostic stability by reducing random error.

Alternate pathways of treatment-seeking are another possible drawback of our clinical based study. Most Spaniards receive medical and mental heath care in public services (212), but we cannot discard the existence of a number of patients assessed in private consultations and private in-patient units. In the particular case of bipolar disorder it is highly probable that they had consulted in the psychiatric public service during the long study period, especially due to symptomatology present in manic episodes, but they might not reach the minimum number of consultations needed to be included in our sample. Similarly, patients might have moved or sought treatment elsewhere along the follow-up period, in particular those with the most unstable diagnoses so leading to a bias in the diagnostic stability. Nevertheless there are some reasons against this possibility. First of all, though residential changes within the same

province are not included, the rates of residential changes to other provinces in Spain or other countries among young people are estimated annually at less than 2% (193).

On the other hand the intentional selection of subjects with 10 or more visits to psychiatric clinics, imply that the results of this investigation may not be generalized to those subjects with more transient and less impairing disorders. We based our estimations on the notion that the follow-up of the patients was the result of a single episode of a disorder. Given the characteristics of our dataset, we could not take into consideration in our analysis the possibility that some patients may have been followed for independent episodes that not only could be distant in time but also of different nature. This limitation however would have resulted in decreasing rather than increasing the diagnostic stability of the bipolar disorder diagnoses studied.

The study of factors involved in the evolution of the bipolar disorder was a secondary aim of the present study and was finally regarded to be beyond the scope of it. Our results show that there is a significant influence of some of these factors, however the transversal collection of data made solely in the initial assessment of the patients and not continued along the study diminish the fiability of the results. Our intention is to generate new studies on this matter that combine longitudinal assessments and longitudinal recollection of relevant factors to explore their relationship regarding the evolution of bipolar disorder.

## *Meaning of the study and future directions*

Many factors can be involved in the instable evolution of a psychiatric diagnosis, Schwartz mentioned that diagnostic changes over time could reflect the evolution of an illness, the emergence of new information, or unreliability of measurement (17). 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 assessments. In our study depression was included as a diagnosis independent of bipolar disorder. In case of having decided otherwise results might have changed in a certain degree the final outcome on the stability of bipolar disorder. This could especially be true for women whose principal confounding factor has been found to be the affective disorders (not bipolar) in our study.

Notwithstanding this point, findings in 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. Actual methods of clinical evaluation may require further revisions to ensure its reliability.

Classic dichotomy as described by Kraepelin, disregarding the limitations of our study, may be at the origin of the low probability of transitions between bipolar disorder and psychotic disorder found in our sample using Markov's model. An independent or categorical view of these disorders is currently more popular in clinical fields according to the actual classification systems, which could lead to the delimitation of diagnoses on both disorders in clinical practice; however many researches have traditionally and more recently defended the opinion that bipolar and psychotic disorders are different dimensions in a continuum (92;109;213;214) and more specifically that a multidimensional approach instead of the categorical one may increase the predictive validity of operational and empirical models in research (215;216). New studies with bigger samples including the whole spectrum of psychotic and affective disorders would be needed to further explore these views.

The study of diagnostic stability on the basis of the three methods used in our study shapes the first step on a broader work and outlines the possibilities to come. These parallel markers allow us to quantify the probability of a given transversal diagnosis to attain the correct one. We expect that the present study may contribute as well to emphasize the important repercussions of diagnostic instability as reflected on the existing differences between comorbidity prevalence in 'stable' and 'not stable' patients (see Misdiagnosis with other psychiatric diagnoses (comorbidity), page 82).

From this point the quest for a better understanding on the temporal course of bipolar disorder can be started, and by extension be applicable to consolidate other psychiatric diagnoses. The route that follows a psychiatric diagnosis along the evolution of the illness frequently crosses different categories (10); its accurate description including the most frequent diagnostic pathways and the factors related to diagnostic change may help to program clinical decisions. The health care system also might benefit from a diagnostic system that included functioning and prognosis of the illness, not only the descriptive, criteriological diagnoses (210).

## *Summary*

Our study addresses the issue of diagnostic stability in the bipolar disorder using three different methods. The results demonstrate a very poor stability according to these three methods and remarkably lower than that found in previous studies, even though depressive diagnoses were considered independently. In the whole sample only 23.1% of the patients kept the diagnoses of bipolar disorder in 75% of the evaluations and both prospective and retrospective consistency were found to be low. Some methodological reasons could explain the differences with previous studies; especially the scanty number of assessments and the shorter follow-up period used which may not allow

enough time to reach the right diagnosis. Our study has the limitations derived from the methodological design of any retrospective naturalistic study and it is also limited by the possible existence of non-controlled pathways to psychiatric care, but may reflect more accurately the 'real' clinical process so raising concern on the precision of actual clinical evaluation systems. The prevalence of bipolar disorder in this psychiatric sample is lower than that found in other psychiatric populations but close to more accurate studies. The closer study of factors influencing the stability of bipolar disorder and a better knowledge on the course of diagnoses along its evolution are proposed as two future lines of investigation.

# CONCLUSIONES

1. La consistencia temporal del trastorno afectivo bipolar fue menor que la encontrada en otros estudios, resultado que puede estar en relación con el escaso número de evaluaciones y los menores períodos de seguimiento utilizados en estudios previos.
2. La estabilidad del diagnóstico de trastorno bipolar medida a partir del porcentaje de diagnósticos concordantes a lo largo de la evaluación fue baja. Sólo uno de cada cuatro pacientes (23.1%) recibió consistentemente el diagnóstico de trastorno bipolar a lo largo del seguimiento.
3. Los resultados del estudio muestran cambios diagnósticos desde otros trastornos psiquiátricos hacia el trastorno bipolar y errores en el diagnóstico del propio trastorno bipolar. El principal factor de confusión en nuestra muestra, coincidiendo con lo descrito en trabajos previos, han sido los diagnósticos del espectro de la esquizofrenia.
4. Los pacientes con un diagnóstico ‘estable’ de trastorno bipolar (‘stable BD’) mostraron una cierta fluctuación en el diagnóstico. La probabilidad de cambio en el diagnóstico estudiada a través de los modelos de Markov indica que los pacientes estables suelen oscilar antes del diagnóstico de trastorno bipolar entre los diagnósticos de abuso de sustancias, esquizofrenia, psicosis agudas y transitorias, trastorno obsesivo compulsivo y trastornos de personalidad (por orden de frecuencias). Una vez asignado el diagnóstico de trastorno bipolar en pacientes estables la probabilidad de cambio se sitúa dentro de esta categoría.

5. La prevalencia del trastorno bipolar en una población psiquiátrica de acuerdo con los resultados de nuestro estudio y considerando los criterios de inclusión utilizados fue más baja que la encontrada en estudios previos. La facilidad de acceso a los servicios psiquiátricos en España puede relacionarse con la menor prevalencia comparativa del trastorno bipolar.
6. Se ha hallado una asociación entre el género y la frecuencia de errores diagnósticos en pacientes inestables ('non stable' group). Las mujeres suelen recibir más diagnósticos de trastorno afectivo no bipolar mientras que los hombres suelen recibir un mayor número de diagnósticos dentro del espectro de la esquizofrenia.
7. Nuestros resultados enfatizan la necesidad en el ámbito de la clínica de adoptar una perspectiva longitudinal, opuesta a las evaluaciones realizadas de forma transversal, y la importancia de utilizar múltiples fuentes de información para prevenir los errores en el diagnóstico del trastorno bipolar.
8. Del mismo modo los resultados de nuestro estudio cuestionan la validez de los trabajos de investigación previos basados en estudios de seguimiento a corto plazo al no existir tiempo suficiente para lograr la estabilización de los diagnósticos. Algo que puede ser determinante en enfermedades crónicas como el trastorno bipolar. La medida en que esto ocurre está aún por determinar.
9. Son necesarios nuevos trabajos que analicen los itinerarios seguidos por los diagnósticos hasta la estabilidad del trastorno bipolar y de los factores relacionados con la inestabilidad diagnóstica estudiando simultáneamente la evolución del diagnóstico y la evolución de los factores relacionados. Estos factores podrían actuar como predictores del trastorno bipolar, por ejemplo si

se examina la relación entre las distintas presentaciones semiológicas de la enfermedad en su comienzo y la evolución posterior.

10. En la medida en que el presente estudio refleja el funcionamiento rutinario de las clasificaciones psiquiátricas y se ajusta a la realidad clínica y en vista de los resultados plantea dudas sobre la validez de los actuales métodos de evaluación clínica. Esto cobra especial importancia en vista de las graves repercusiones a las que dan lugar las dificultades en el diagnóstico correcto del trastorno bipolar, que aumentan significativamente los costes personales y económicos de la enfermedad.

# CONCLUSIONS

1. The temporal consistency of bipolar affective disorder (BD) was lower than that found in other studies, an outcome that may be related to the scarce number of assessments and inferior time periods of follow-up used in previous studies.
2. Diagnostic stability of bipolar disorder measured from the percentage of concordant diagnosis along the follow-up was low. Only one out of every four patients (23.1%) got consistently the bipolar disorder diagnosis along the follow-up.
3. Results of the study show diagnostic shifts from other psychiatric disorders towards bipolar affective disorder and misdiagnosis of this disorder. The main confounding factor in our sample, as reported in previous studies, were the diagnoses situated within the schizophrenia spectrum.
4. Patients with a 'stable' diagnosis of bipolar disorder ('stable BD') displayed some diagnostic fluctuation. The probability of diagnostic shift as studied through Markov's models indicates that stable patients are inclined to swing before the diagnosis of bipolar disorder is settled between the diagnoses of substance abuse, schizophrenia, acute and transient psychosis, obsessive-compulsive disorder and personality disorder (ordered by frequencies). Once the bipolar disorder diagnosis is set the probability of a diagnostic shift diminish deeply, so that patients tend to keep that diagnosis.
5. The prevalence of bipolar disorder in a psychiatric population according to the results obtained in our study and taking account of the inclusion criteria was lower than that found in previous studies. The easily access to

psychiatric services in Spain may be related with a smaller comparative prevalence of bipolar disorder.

6. A significant association was found between gender and the frequency of misdiagnoses in instable patients ('non stable' group). Women are more commonly diagnosed in the category of affective disorder (not bipolar) while men use to receive a bigger number of diagnoses inside the schizophrenia spectrum.
7. Our results emphasize the need in the clinical field for the adoption of a longitudinal perspective, opposite to the transversal evaluations, and the importance of using multiple sources of information to prevent errors in the diagnosis of bipolar disorder.
8. In a similar way, the outcome of our investigation questions the validity of previous research studies based on short follow-up periods that may not provide enough time to attain the stabilisation of a diagnosis. This could be especially important in chronic illness as is the case of bipolar disorder. The degree to which this happens is still to be determined.
9. New research studies are needed to analyse both the pathways followed by diagnoses assigned to bipolar patients through their evolution till the definitive stability of this diagnosis, and the factors related with diagnostic instability studying simultaneously the evolution of diagnoses and the evolution of related factors. These factors could act as predictors of bipolar disorder, for instance through the inspection of the relationship between different semiologic presentations of the illness in its beginning and the subsequent evolution.

10. In the measure that the present study may reflect the daily performance of psychiatric classifications and is adjusted to the clinical reality and in the view of the results obtained, some doubts arouse on the validity of actual methods of clinical evaluation. This acquires a special importance taking account of the severe repercussions that may derive from difficulties in the righteous diagnosis of bipolar disorder that increase notably the personal and economic costs of the disease.

# TABLES

# FIGURES

# Appendix 1: Recogida de datos

Servicios de Salud Mental

Madrid

FECHA:____/____/____ Nº HISTORIA:

Nº SEGURIDAD SOCIAL: ___ / _____

APELLIDOS: 1º: _____ 2º:_____

NOMBRE: _____ Nº D.N.I.: _____

DOMICILIO: _____ C. POSTAL: _____

TELÉFONO: _____/_____

FECHA DE NACIMIENTO: _____/_____/_____ HOMBRE: _____ MUJER: _____

**ESTADO CIVIL:**

| | |
|---|---|
| SOLTERO/A: | _____S |
| CASADO/A: | _____C |
| DIVORCIADO/A: | _____D |
| SEPARADO/A: | _____X |
| VIUDO/A: | _____V |

**TIPO DE CONVIVENCIA:**

| | |
|---|---|
| SOLO/A: | _____01 |
| CON CÓNYUGE: | _____02 |
| CON PAREJA: | _____03 |
| CON PADRES: | _____04 |
| SOLO CON PADRE: | _____05 |
| SOLO CON MADRE: | _____06 |
| CON HIJOS: | _____07 |
| CON OTROS FAMIL: | _____08 |
| EN INSTITUCIÓN: | _____09 |
| OTROS: | _____00 |

**TIPO DE ESTUDIOS:**

| | |
|---|---|
| ANALFABETO/A: | _____01 |
| SIN ESTUDIOS: | _____02 |
| ESTUDIOS PRIMARIOS: | _____03 |
| GRADUADO ESCOLAR: | _____04 |
| BACHILLER: | _____05 |
| COU: | _____06 |
| TITUL. UNIVERSITARIO: | _____07 |
| LICEN. UNIVERSITARIO: | _____08 |
| OTROS: | _____09 |

**OCUPACIÓN O PROFESIÓN:**

| | |
|---|---|
| SIN TRABAJO | _____00 |
| PROFESIONALES Y TÉCNICOS: | _____01 |
| DIRECTIVOS: | _____02 |
| PERSONAL ADMINISTRATIVO: | _____03 |
| VENDEDORES Y COMERCIANTES: | _____04 |
| HOSTELERÍA Y SERV. DE SEGURIDAD: | _____05 |
| AGRICULTURA Y GANADERÍA: | _____06 |
| PERSONAL DE INDUSTRIA, CONSTRUCCIÓN Y TRANSPORTE: | _____07 |
| OTROS: | _____08 |
| PERSONAL FUERZAS ARMADAS: | _____09 |

**SITUACIÓN LABORAL:**

| | |
|---|---|
| TRABAJANDO : | _____02 |
| BUSCANDO PRIMER EMPLEO: | _____03 |
| PARADO CON SUBSIDIO: | _____04 |
| PARADO SIN SUBSIDIO: | _____05 |
| RETIRADO, PENSIONISTA, JUBILADO: | _____06 |
| ESTUDIANDO: | _____08 |
| DEDICADO LABORES DEL HOGAR: | _____09 |
| INCAPACIDAD LABORAL TRANS.: | _____10 |
| INCAPACIDAD PERMANENTE: | _____11 |

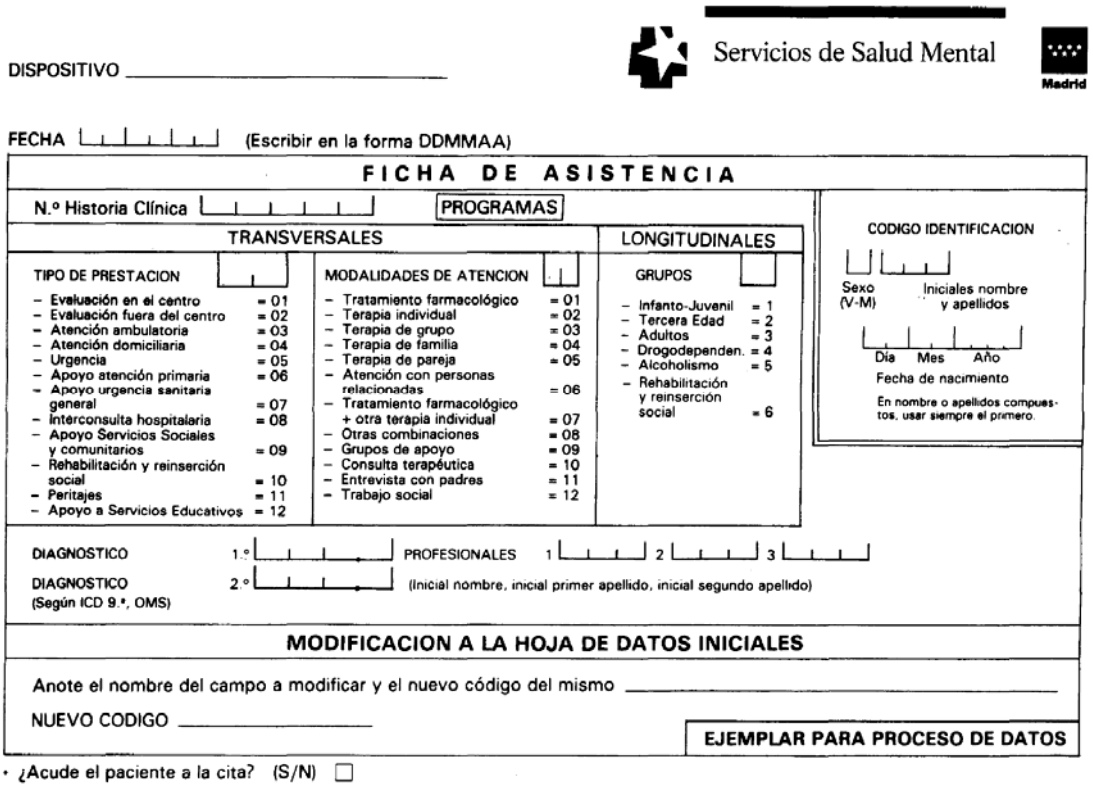
NOMBRE DEL CONSULTORIO: _____

NOMBRE DEL MÉDICO DE CABECERA: _____

¿HA TENIDO CONTACTO CON PSIQUIÁTRA O PSICÓLOGO ANTERIORMENTE?:

SI: _____ ¿DE QUÉ TIPO?: PARTICULAR: _____ A; AMBULATORIO: _____ A; HOSPITALARIO: _____ H NO: _____

Ficha para la recogida de datos socio-laborales de los pacientes.



DISPOSITIVO _____

Servicios de Salud Mental — Madrid

FECHA ⌊⌊⌊⌊⌊⌊⌋ (Escribir en la forma DDMMAA)

**FICHA DE ASISTENCIA**

N.º Historia Clínica ⌊⌊⌊⌊⌊⌋ PROGRAMAS

TRANSVERSALES | LONGITUDINALES

TIPO DE PRESTACION ⌊⌋
- Evaluación en el centro = 01
- Evaluación fuera del centro = 02
- Atención ambulatoria = 03
- Atención domiciliaria = 04
- Urgencia = 05
- Apoyo atención primaria = 06
- Apoyo urgencia sanitaria general = 07
- Interconsulta hospitalaria = 08
- Apoyo Servicios Sociales y comunitarios = 09
- Rehabilitación y reinserción social = 10
- Peritajes = 11
- Apoyo a Servicios Educativos = 12

MODALIDADES DE ATENCION ⌊⌋
- Tratamiento farmacológico = 01
- Terapia individual = 02
- Terapia de grupo = 03
- Terapia de familia = 04
- Terapia de pareja = 05
- Atención con personas relacionadas = 06
- Tratamiento farmacológico + otra terapia individual = 07
- Otras combinaciones = 08
- Grupos de apoyo = 09
- Consulta terapéutica = 10
- Entrevista con padres = 11
- Trabajo social = 12

GRUPOS ⌊⌋
- Infanto-Juvenil = 1
- Tercera Edad = 2
- Adultos = 3
- Drogodependen. = 4
- Alcoholismo = 5
- Rehabilitación y reinserción social = 6

CODIGO IDENTIFICACION
⌊⌋⌊⌊⌊⌋
Sexo (V-M) Iniciales nombre y apellidos
⌊⌊⌋⌊⌊⌋⌊⌊⌋
Día Mes Año
Fecha de nacimiento
En nombre o apellidos compuestos, usar siempre el primero.

DIAGNOSTICO 1.º ⌊⌊⌊.⌊⌋ PROFESIONALES 1 ⌊⌊⌊⌋ 2 ⌊⌊⌊⌋ 3 ⌊⌊⌊⌋
DIAGNOSTICO 2.º ⌊⌊⌊.⌊⌋ (Inicial nombre, inicial primer apellido, inicial segundo apellido)
(Según ICD 9.ª, OMS)

**MODIFICACION A LA HOJA DE DATOS INICIALES**

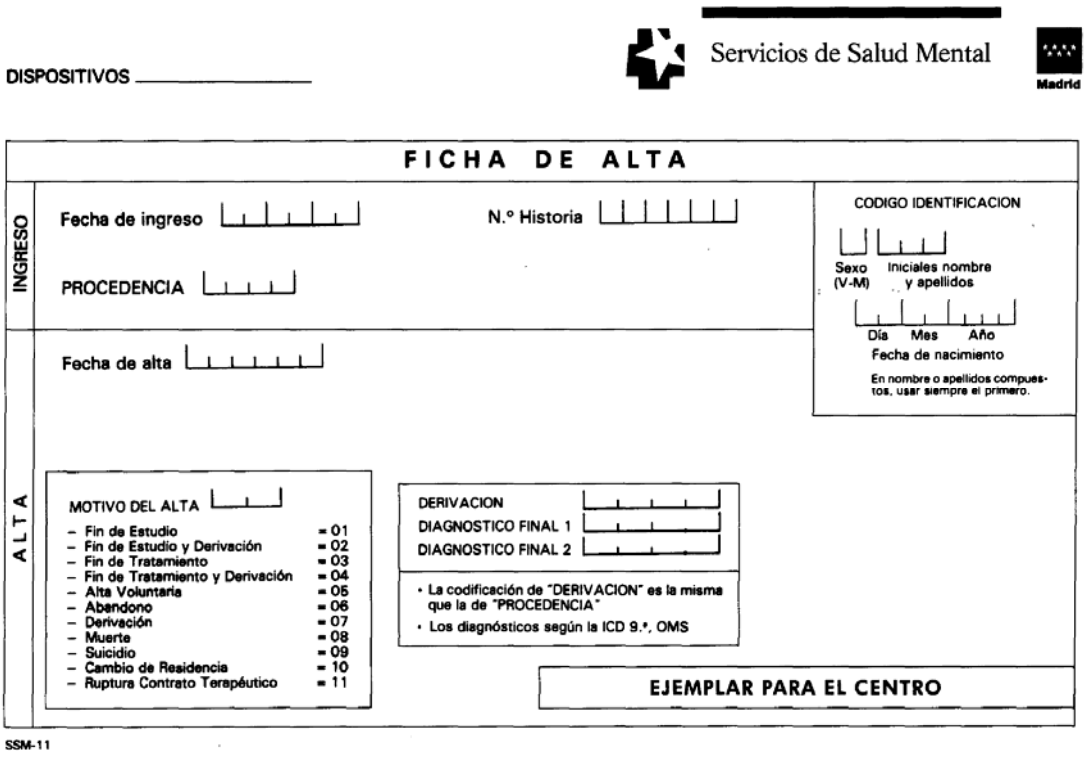
Anote el nombre del campo a modificar y el nuevo código del mismo _____

NUEVO CODIGO _____

**EJEMPLAR PARA PROCESO DE DATOS**

• ¿Acude el paciente a la cita? (S/N) ☐

Ficha de asistencia en consulta, donde se especifica el tipo de prestación, la modalidad de atención y el diagnóstico.



DISPOSITIVOS _____

Servicios de Salud Mental — Madrid

**FICHA DE ALTA**

INGRESO

Fecha de ingreso ⌊⌊⌊⌊⌊⌋ N.º Historia ⌊⌊⌊⌊⌊⌋

PROCEDENCIA ⌊⌊⌊⌋

CODIGO IDENTIFICACION
⌊⌋⌊⌊⌊⌋
Sexo (V-M) Iniciales nombre y apellidos
⌊⌊⌋⌊⌊⌋⌊⌊⌋
Día Mes Año
Fecha de nacimiento
En nombre o apellidos compuestos, usar siempre el primero.

ALTA

Fecha de alta ⌊⌊⌊⌊⌊⌋

MOTIVO DEL ALTA ⌊⌊⌋
- Fin de Estudio = 01
- Fin de Estudio y Derivación = 02
- Fin de Tratamiento = 03
- Fin de Tratamiento y Derivación = 04
- Alta Voluntaria = 05
- Abandono = 06
- Derivación = 07
- Muerte = 08
- Suicidio = 09
- Cambio de Residencia = 10
- Ruptura Contrato Terapéutico = 11

DERIVACION ⌊⌊⌊⌋
DIAGNOSTICO FINAL 1 ⌊⌊⌊⌋
DIAGNOSTICO FINAL 2 ⌊⌊⌊.⌊⌋

• La codificación de "DERIVACION" es la misma que la de "PROCEDENCIA"
• Los diagnósticos según la ICD 9.ª, OMS

**EJEMPLAR PARA EL CENTRO**

SSM-11

Ficha para el alta de un paciente, se especifica el motivo y el/los diagnóstico/s final/es

# Appendix 2: Markov Models

Some problems can be modeled as a state machine, i.e., a sequence of states and the set of probabilities of jumping from one to the others and to itself; the changes in the state take place at fixed periods of time. For these problems a useful mathematical tool are Markov Models (MM). The basic assumption of the (first order) Markov Model is that all the information needed to estimate the next state at t+1 is contained in the present state at time t. The resulting model is very tractable at the expenses of a lack of modeling capability. Nevertheless, for the problem at hand MM has revealed very useful.

More deeply, the MM consists of a sequence of T states and the probabilities $a_{ij}$:

$$\vec{\omega} = \{\omega_1, \omega_2, \cdots, \omega_T\}; \qquad a_{ij} = \Pr\{\omega(t+1) = \omega_j \mid \omega(t) = \omega_i\}; (i, j) \in [0, T]$$

Once the states have been chosen, the training of the model consists of simply computing the a priori probabilities of the T $\omega_i$ states and the estimation of the transitions $a_{ij}$. This is achieved by a frequentist approach: the count of the number of times a transition occurs divided by the total number of transitions.

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 (1). The probability of each transition is represented by a gradient, from dark blue (the lowest probability=0) to dark red (the highest probability=1).

In some of the experiments, the Markov Model presents a little modification, due to the fact that the changes of state of the patients occur in times not uniformly spaced. Our solution has been to include the time information while encoding the states of the model.

Consequently, the transitions of the model occur in a virtual time index that has not a literal sense.

Bibliography:

Richard Duda, Peter Hart, David Stork. “Pattern Classification”. Second Edition. John Wiley & Sons, Inc. 2001

Rabiner, L.; Juang, B. An introduction to hidden Markov models, ASSP Magazine, IEEE Volume 3, Issue 1, Part 1, Jan 1986 Page(s):4 – 16

# Appendix 3: ICD-10 manic episode and bipolar disorder diagnostic criteria

## F30 Manic episode

All the subdivisions of this category should be used only for a single episode. Hypomanic or manic episodes in individuals who have had one or more previous affective episodes (depressive, hypomanic, manic, or mixed) should be coded as bipolar affective disorder (F31.-).

Includes: bipolar disorder, single manic episode

**F30.0 Hypomania**

A disorder characterized by a persistent mild elevation of mood, increased energy and activity, and usually marked feelings of well-being and both physical and mental efficiency. Increased sociability, talkativeness, over-familiarity, increased sexual energy, and a decreased need for sleep are often present but not to the extent that they lead to severe disruption of work or result in social rejection. Irritability, conceit, and boorish behaviour may take the place of the more usual euphoric sociability. The disturbances of mood and behaviour are not accompanied by hallucinations or delusions.

**F30.1 Mania without psychotic symptoms**

Mood is elevated out of keeping with the patient's circumstances and may vary from carefree joviality to almost uncontrollable excitement. Elation is accompanied by increased energy, resulting in overactivity, pressure of speech, and a decreased need for sleep. Attention cannot be sustained, and there is often marked distractibility. Self-esteem is often inflated with grandiose ideas and overconfidence. Loss of normal social inhibitions may result in behaviour that is reckless, foolhardy, or inappropriate to the circumstances, and out of character.

**F30.2 Mania with psychotic symptoms**

In addition to the clinical picture described in F30.1, delusions (usually grandiose) or hallucinations (usually of voices speaking directly to the patient) are present, or the excitement, excessive motor activity, and flight of ideas are so extreme that the subject is incomprehensible or inaccessible to ordinary communication.

Mania with:

- mood-congruent psychotic symptoms
- mood-incongruent psychotic symptoms

Manic stupor.

**F30.8 Other manic episodes**

**F30.9 Manic episode, unspecified**

Mania NOS.

## F31 Bipolar affective disorder

A disorder characterized by two or more episodes in which the patient's mood and activity levels are significantly disturbed, this disturbance consisting on some occasions of an elevation of mood and increased energy and activity (hypomania or mania) and on others of a lowering of mood and decreased energy and activity (depression). Repeated episodes of hypomania or mania only are classified as bipolar.

Includes: manic-depressive:

- illness
- psychosis
- reaction

Excludes:

- bipolar disorder, single manic episode ( F30.- )

- cyclothymia ( F34.0 )

**F31.0 Bipolar affective disorder, current episode hypomanic**

The patient is currently hypomanic, and has had at least one other affective episode (hypomanic, manic, depressive, or mixed) in the past.

**F31.1 Bipolar affective disorder, current episode manic without psychotic symptoms**

The patient is currently manic, without psychotic symptoms (as in F30.1), and has had at least one other affective episode (hypomanic, manic, depressive, or mixed) in the past.

**F31.2 Bipolar affective disorder, current episode manic with psychotic symptoms**

The patient is currently manic, with psychotic symptoms (as in F30.2), and has had at least one other affective episode (hypomanic, manic, depressive, or mixed) in the past.

**F31.3 Bipolar affective disorder, current episode mild or moderate depression**

The patient is currently depressed, as in a depressive episode of either mild or moderate severity (F32.0 or F32.1), and has had at least one authenticated hypomanic, manic, or mixed affective episode in the past.

**F31.4 Bipolar affective disorder, current episode severe depression without psychotic symptoms**

The patient is currently depressed, as in severe depressive episode without psychotic symptoms (F32.2), and has had at least one authenticated hypomanic, manic, or mixed affective episode in the past.

**F31.5 Bipolar affective disorder, current episode severe depression with psychotic symptoms**

The patient is currently depressed, as in severe depressive episode with psychotic symptoms (F32.3), and has had at least one authenticated hypomanic, manic, or mixed affective episode in the past.

**F31.6 Bipolar affective disorder, current episode mixed**

The patient has had at least one authenticated hypomanic, manic, depressive, or mixed affective episode in the past, and currently exhibits either a mixture or a rapid alteration of manic and depressive symptoms.

Excludes: single mixed affective episode (F38.0)

**F31.7 Bipolar affective disorder, currently in remission**

The patient has had at least one authenticated hypomanic, manic, or mixed affective episode in the past, and at least one other affective episode (hypomanic, manic, depressive, or mixed) in addition, but is not currently suffering from any significant mood disturbance, and has not done so for several months. Periods of remission during prophylactic treatment should be coded here.

**F31.8 Other bipolar affective disorders**

Bipolar II disorder

Recurrent manic episodes NOS

**F31.9 Bipolar affective disorder, unspecified**

# Appendix 4: Papers

The results of this thesis have produced some scientific papers. One of them has been already published in an international journal. Another has been submitted for publication and is actually in press. Nonetheless, the paper drafts are included.

*Acta Psychiatr Scand 2007: 115: 473–480*
*All rights reserved*
DOI: 10.1111/j.1600-0447.2006.00984.x



*ACTA PSYCHIATRICA SCANDINAVICA*

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

Baca-Garcia E, Perez-Rodriguez MM, Basurte-Villamor I, López-Castromán J, Fernandez del Moral AL, Jimenez-Arriero MA, Gonzalez de Rivera JL, Saiz-Ruiz J, Leiva-Murillo JM, de Prado-Cumplido M, Santiago-Mozos R, Artés-Rodríguez A, Oquendo MA, de Leon J. 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.

**E. Baca-Garcia[1], M. M. Perez-Rodriguez[2], I. Basurte-Villamor[1], J. López-Castromán[1], A. L. Fernandez del Moral[3], M. A. Jimenez-Arriero[4], J. L. Gronzalez de Rivera[1], J. Saiz-Ruiz[2,5], J. M. Leiva-Murillo[6], M. de Prado-Cumplido[6], R. Santiago-Mozos[6], A. Artés-Rodríguez[6], M. A. Oquendo[7], J. de Leon[8]**

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

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

Enrique Baca-Garcia, Department of Psychiatry, Fundacion Jimenez Diaz University Hospital, Autonoma University of Madrid, Avda, Reyes Católicos 2, Madrid 28040, Spain.
E-mail: ebacgar2@yahoo.es

Accepted for publication November 28, 2006

**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.

**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 life-long 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-Garcia 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 is 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

Table 1. Diagnostic frequencies of the most common diagnoses

| Diagnosis | Whole sample [% (*n*)] | Stable BD* [% (*n*)] | Non-stable BD† [% (*n*)] | Last diagnosis BD‡ [% (*n*)] | Last diagnosis not BD§ [% (*n*)] | First diagnosis BD¶ [% (*n*)] | First diagnosis not BD** [% (*n*)] |
|---|---|---|---|---|---|---|---|
| F20.0 paranoid schizophrenia | 11.8 (8447) | 0.3 (38) | 14.4 (8409) | 2.4 (525) | 16.0 (7922) | 3.3 (561) | 14.5 (7886) |
| F20.5 residual schizophrenia | 10.6 (7576) | 0.3 (33) | 12.9 (7543) | 2.0 (434) | 14.5 (7142) | 3.3 (558) | 12.9 (7018) |
| F31.1 BD, current episode manic without psychotic symptoms | 9.9 (7113) | 28.7 (3767) | 5.7 (3346) | 20.8 (4589) | 5.1 (2524) | 20.0 (3430) | 6.8 (3683) |
| F31.3 BD, current episode mild or moderate depression | 11.4 (8131) | 29.6 (3896) | 7.3 (4235) | 21.8 (4817) | 6.7 (3314) | 20.0 (3425) | 8.7 (4706) |
| F31.6 BD, most recent episode mixed | 2.2 (1551) | 7.5 (981) | 1.0 (570) | 4.6 (1014) | 1.1 (537) | 5.3 (909) | 1.2 (642) |
| F31.8 other BD | 4.6 (3288) | 15.3 (2017) | 2.2 (1271) | 10.0 (2206) | 2.2 (1082) | 12.3 (2100) | 2.2 (1188) |
| F31.9 BD, most recent episode unspecified | 3.5 (2518) | 8.9 (1174) | 2.3 (1344) | 6.8 (1504) | 2.1 (1014) | 5.8 (984) | 2.8 (1534) |
| F33. major depressive disorder, recurrent | 6.7 (4855) | 1.0 (136) | 8.0 (4691) | 7.1 (1577) | 6.6 (3250) | 3.3 (557) | 7.9 (4270) |
| F34.1 dysthymia | 8.8 (6314) | 0.8 (109) | 10.6 (6205) | 5.8 (1278) | 10.2 (5036) | 6.3 (1069) | 9.6 (5245) |
| Total assessments | 100.0 (71543) | 100.0 (13148) | 100.0 (58395) | 100.0 (22117) | 100.0 (49426) | 100.0 (17122) | 100.0 (54421) |

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 the 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

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

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 22 117 assessments of the 443 'last diagnosis BD' 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%, 4589/22 117); other BD (F31.8; 10%, 2206/22 117); major depressive disorder, recurrent (F33; 7%, 1577/22 117); BD, most recent episode unspecified (F31.9; 7%, 1504/22 117); and dysthymia (F34.1; 6%, 1278/22 117).

The most frequent diagnoses during the study period among all 49 426 assessments of the 710 'last diagnosis not BD' patients were: paranoid schizophrenia (F20.0; 16%, 7922/49 426); residual schizophrenia (F20.5; 14%, 7142/49 426); dysthymia (F34.1; 10%, 5036/49 426); BD, current episode mild or moderate depression (F31.3; 7%, 3314/49 426); major depressive disorder, recurrent (F33; 7%, 3250/49 426); and BD, 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.

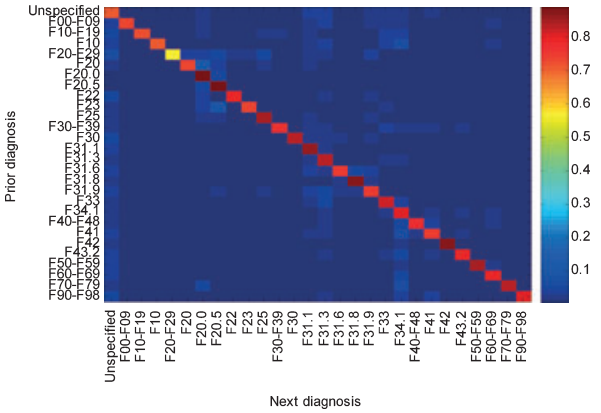



*Fig. 1.* Probability of transitions between prior diagnoses [a] and next diagnoses in the whole sample ($n = 1153$).
[a]Diagnoses: F00–F09, organic, including symptomatic, mental disorders; F10–F19, mental and behavioral disorders due to psychoactive substance use; F10, mental and behavioral disorders 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 mild or moderate depression; F31.6, bipolar disorder, most recent episode mixed; F31.8, other bipolar affective disorders; F31.9, bipolar disorder, most recent episode unspecified; F33, major depressive disorder, recurrent; F34.1, dysthymia; F40–F48, neurotic, stress-related and somatoform disorders; F41, other anxiety disorders; F42, obsessive–compulsive disorder; F43.2, adjustment disorders; F50–F59, behavioral syndromes associated with physiologic disturbances and physical factors; F60–F69, disorders of adult personality and behavior; F70–F79, mental retardation; F90–F98, behavioral and emotional disorders with onset usually occurring in childhood and adolescence.

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:

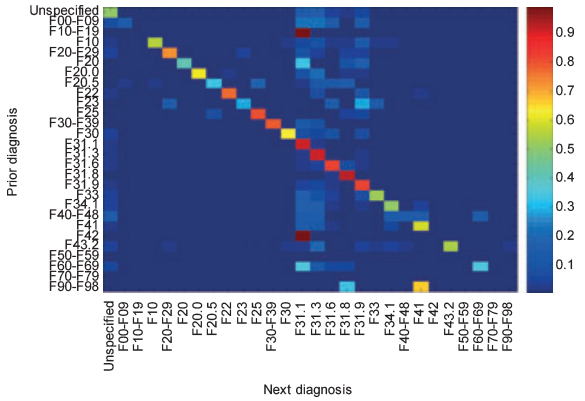



*Fig. 2.* Probability of transitions between prior diagnoses and next diagnoses in the 266 'stable bipolar affective disorder (BD)' patients (patients who have received the diagnosis of BD in ≥75% of the evaluations).

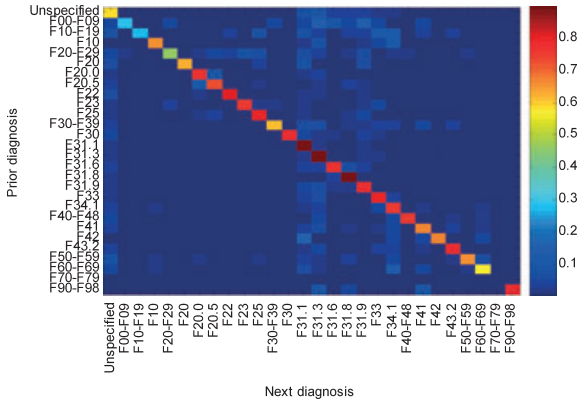



*Fig. 3.* Probability of transitions between prior diagnoses and next diagnoses of the 443 'last diagnosis bipolar affective disorder (BD)' patients (patients who received the diagnosis of BD at the last evaluation).

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.

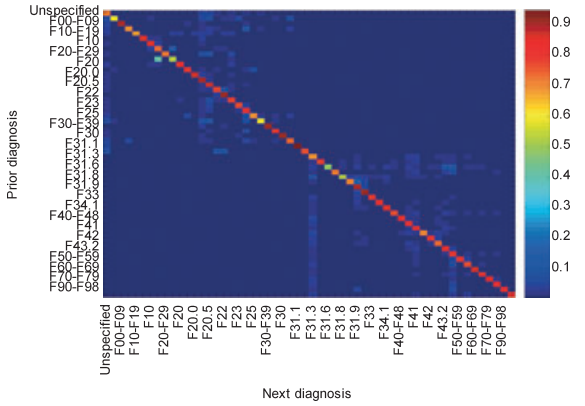



*Fig. 4.* Probability of transitions between prior diagnoses and next diagnoses of the 342 'first diagnosis bipolar affective disorder (BD)' patients (patients who received the diagnosis of BD at the first evaluation) including time between stages.

## 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:

mental and behavioral disorders due to use of psychoactive substances, schizophrenia, acute and transient psychotic disorders, obsessive–compulsive disorder, personality disorders, and behavioral and emotional disorders with onset usually occurring in childhood and adolescence.

A high proportion in the 'non-stable BD' group had received the diagnoses of paranoid schizophrenia (14%), residual schizophrenia (13%), dysthymia (11%), and major depressive disorder (8%). Similar findings were observed in the 'last diagnosis not BD' and 'first diagnosis not BD' groups.

The temporal consistency of BD was lower than that found in other studies (1, 3, 12, 13, 15). Four variables were related to the diagnostic stability of BD: gender, age ≥40 years, number of psychiatric consultations, and treatment at the out-patient mental health centers.

### Strengths and weaknesses of the study

The main strengths of this study are the large representative sample, the length of follow-up (up to 12 years, mean = 6.2 years), and the high number of assessments (median = 34). The stability of BD was evaluated in three different clinical settings, using the same diagnostic procedure that is used during regular clinical practice. 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

(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

1. Schwartz JE, Fennig S, Tanenberg-Karant M et al. Congruence of diagnosis 2 years after a first-admission diagnosis of psychosis. Arch Gen Psychiatry 2000;**57**:593–600.
2. Weeke A. Admission pattern and diagnostic stability among unipolar and bipolar manic-depressive patients. Acta Psychiatr Scand 1984;**70**:603–613.
3. Chen YR, Swann AC, Johnson BA. Stability of diagnosis in bipolar disorder. J Nerv Ment Dis 1998;**186**:17–23.
4. Beiser M, Iacono WG, Erickson D. Temporal stability in the major mental disorders. In: Robins LN, Barrett JE, eds. The validity of psychiatric diagnosis. New York: Raven, 1989:77–98.
5. Blacker D, Tsuang MT. Contested boundaries of bipolar disorder and the limits of categorical diagnosis in psychiatry. Am J Psychiatry 1992;**149**:1473–1483.
6. Kendler KS. Toward a scientific nosology: strengths and limitations. Arch Gen Psychiatry 1990;**47**:969–973.
7. Cooper JE. Diagnostic change in a longitudinal study of psychiatric patients. Br J Psychiatry 1967;**113**:129–142.
8. Amin S, Singh SP, Brewin J, Jones PB, Medley I, Harrison G. Diagnostic stability of first-episode psychosis. Comparison of ICD-10 and DSM-III-R systems. Br J Psychiatry 1999;**175**:537–543.
9. Fennig S, Kovasznay B, Rich C et al. Six-month stability of psychiatric diagnoses in first-admission patients with psychosis. Am J Psychiatry 1994;**151**:1200–1208.
10. Lenz G, Simhandl C, Thau K, Berner P, Gabriel E. Temporal stability of diagnostic criteria for functional psychoses. Results from the Vienna follow-up study. Psychopathology 1991;**24**:328–335.
11. Marneros A, Deister A, Rohde A. Stability of diagnoses in affective, schizoaffective and schizophrenic disorders. Cross-sectional versus longitudinal diagnosis. Eur Arch Psychiatry Clin Neurosci 1991;**241**:187–192.
12. Schimmelmann BG, Conus P, Edwards J, Mcgorry PD, Lambert M. Diagnostic stability 18 months after treatment initiation for first-episode psychosis. J Clin Psychiatry 2005;**66**:1239–1246.
13. Kessing LV. Diagnostic stability in bipolar disorder in clinical practise as according to ICD-10. J Affect Disord 2005;**85**:293–299.
14. Rufino ACTBF, Uchida RR, Vilela JAA, Marques JMA, Zuardi AW, Del-Ben CM. Stability of the diagnosis of first–episode psychosis made in an emergency setting. Gen Hosp Psychiatry 2005;**27**:189–193.
15. Veen ND, Selten JP, Schols D et al. Diagnostic stability in a Dutch psychosis incidence cohort. Br J Psychiatry 2004;**185**:460–464.
16. World Health Organization. Tenth revision of the International Classification of Diseases and related health problems (ICD-10). Geneva: World Health Organization, 1992.
17. Baca-Garcia E, Perez-Rodriguez MM, Basurte-Villamor I et al. Diagnostic stability of psychiatric disorders in clinical practice. Br J Psychiatry (in press).
18. Duda R, Hart P, Stork D. Pattern classification, 2nd edn. New York: John Wiley & Sons Inc., 2001.
19. Rabiner L, Juang B. An introduction to hidden Markov models. ASSP Magazine 1986;**1**(part I):4–16.
20. Lish JD, Dime-Meenan S, Whybrow PC, Price RA, Hirschfeld RM. The National Depressive and Manic Depressive Association (DMDA) survey of bipolar members. J Affect Disord 1994;**31**:281–294.
21. Hirschfeld RM, Lewis L, Vornik LA. Perceptions and impact of bipolar disorder: how far have we really come? Results of the national depressive and manic-depressive association 2000 survey of individuals with bipolar disorder. J Clin Psychiatry 2003;**64**:161–174.
22. Robins E, Guze SB. Establishment of diagnostic validity in psychiatric illness: its application to schizophrenia. Am J Psychiatry 1970;**126**:983–987.
23. Kendler KS. The nosologic validity of paranoia (simple delusional disorder): a review. Arch Gen Psychiatry 1980;**37**:699–706.
24. Andreasen NC. The validation of psychiatric diagnosis: new models and approaches (editorial). Am J Psychiatry 1995;**152**:161–162.
25. Jorgensen P, Bennedsen B, Christensen J, Hyllested A. Acute and transient psychotic disorder: a 1-year follow-up study. Acta Psychiatr Scand 1997;**96**:150–154.
26. Kendell R, Jablensky A. Distinguishing between the validity and utility of psychiatric diagnoses. Am J Psychiatry 2003;**160**:4–12.
27. Patten SB. Markov models of major depression for linking psychiatric epidemiology to clinical practice. Clin Pract Epidemol Ment Health 2005;**1**:2.

# ERRORES DE DIAGNÓSTICO Y ESTABILIDAD TEMPORAL EN EL TRASTORNO BIPOLAR

**Diagnostic errors and temporal stability of bipolar disorder**

Autores: López Castromán J [a], Baca García E [a], Botillo Martín C [a], Quintero Gutiérrez del Álamo J [a], Navarro Jiménez R [a], Negueruela López M [a], Pérez Rodríguez MM [b], Basurte Villamor I [a], Fernández del Moral A [c], Jiménez Arriero MA [d], González de Rivera JL [a].

[a] Servicio de Psiquiatría, Hospital Universitario Fundación Jiménez Díaz, Madrid.

[b] Servicio de Psiquiatría, Hospital Universitario Ramón y Cajal, Madrid.

[c] Centro de Salud Mental del distrito Centro, Madrid.

[d] Centro de Salud Mental del distrito de Arganzuela, Madrid.

Contacto: Jorge López Castromán (Residente de Psiquiatría de la Fundación Jiménez Díaz).

Dirección: Fundación Jiménez Díaz. Avenida de los Reyes Católicos nº2. C.P: 28040

Teléfono: 915504987/616943402

Correo electrónico: jorgecastroman@gmail.com

**Resumen:** El diagnóstico de trastorno bipolar se modifica con frecuencia a lo largo de la evolución de la enfermedad. Se describen los cambios de diagnóstico y error asociado de 1153 pacientes mayores de 18 años, diagnosticados de trastorno bipolar y con un seguimiento mínimo de 10 visitas, en base a un registro clínico de atención ambulatoria especializada en Psiquiatría y hospitalizaciones psiquiátricas de 25152 pacientes representativos de un área urbana de 240.000 habitantes. Se

usó como criterio de estabilidad diagnóstica mantener el diagnóstico de trastorno bipolar en al menos el 75% de las visitas. De los 342 pacientes diagnosticados de trastorno bipolar en la primera consulta el 46,1% mantuvieron el diagnóstico estable. Se cometió un error inicial de infradiagnóstico con 108 pacientes estables no diagnosticados en la primera visita. 184 de los 342 pacientes diagnosticados en la primera visita obtuvieron posteriormente al menos un 25% de diagnósticos diferentes de bipolar y podrían ser considerados como sobrediagnóstico inicial. 209 de 443 pacientes diagnosticados como bipolares en la última visita no mantuvieron criterios de estabilidad en su evolución y podrían por tanto considerarse como sobrediagnóstico final. 32 pacientes estables no diagnosticados en la última visita constituirían el error final de infradiagnóstico. Diagnósticos del espectro de esquizofrenia (F2), aparecen casi en una de cada cuatro visitas al psiquiatra de los pacientes del estudio. Otras tres categorías presentan solapamiento: los trastornos de ansiedad (F4), los trastornos de personalidad (F6) y los trastornos por consumo de sustancias. Conclusión: El trastorno bipolar es un trastorno de difícil diagnóstico en su evolución inicial.

**Abstract:** Bipolar disorder diagnosis is frequently modified along the evolution of the illness. Diagnostic changes and associated errors were described for 1153 patients diagnosed as bipolar disorder, aged over 18 years and with at least ten consultations in their follow-up. Data was extracted from a clinical registry of ambulatory care specialised on Psychiatry and psychiatric hospitalizations of 25152 patients representing an urban area of 240000 inhabitants. Limit for diagnostic stability was established as keeping bipolar disorder diagnosis in at least 75% of the consultations. Out of 342 patients diagnosed as bipolar disorders in the first

consultation, 46,1% kept this diagnosis stable. Infradiagnostic initial error was committed on 108 stable patients that were not diagnosed in the first visit. 184 out of 342 patients diagnosed in the first consultation got in their follow-up at least 25% of diagnosis differing from bipolar disorder and could be seen as initial overdiagnosis. 209 out of 443 patients that were diagnosed as bipolar disorder in their last visit did not keep stability criteria in their follow-up and could be considered therefore as final overdiagnosis. 32 stable patients not diagnosed in their last visit could be considered as infradiagnosis final error. Diagnosis from schizophrenia spectrum (F2), appear in one of every four psychiatric consultations of the patients included in this study. Overlap was seen in three other categories: anxiety disorders (F4), personality disorders (F6) and substance abuse disorders. Conclusion: Initial evolution of bipolar disorder causes difficulties in the diagnosis.

Palabras clave: Trastorno bipolar, estabilidad, diagnóstico, evolución.

Key words: Bipolar disorder, stability, diagnosis, evolution.

**Introducción:**

El concepto de estabilidad diagnóstica comienza a cobrar importancia con el artículo publicado por Robins y Guze en 1970, donde lo relacionan por primera vez con la validez predictiva de los diagnósticos en psiquiatría[1]. Este concepto ha sido definido como la medida en que un diagnóstico es confirmado en evaluaciones consecutivas[2]. También se han propuesto distintos métodos para potenciar la estabilidad de un diagnóstico aunque ninguno de ellos asegura la fiabilidad del resultado. Estos métodos incluirían: la evaluación u observación longitudinal[3,4],

pero también estudios diagnósticos más avanzados como los de tipo genético, la monitorización de la respuesta al tratamiento[5] o la evaluación de los efectos de la enfermedad sobre la función psicosocial.

Existen múltiples factores que pueden provocar inestabilidad en los diagnósticos psiquiátricos. La propia evolución de las enfermedades mentales, y en concreto del trastorno bipolar, hace difícil a menudo diferenciar la entidad clínica de los cuadros. Asimismo, la falta de una información completa sobre el curso de la enfermedad y posibles errores diagnósticos previos pueden también dar lugar a un sesgo en la valoración.

En el caso concreto del trastorno bipolar las dificultades diagnósticas parecen especialmente relevantes al inicio de la enfermedad y a menudo el diagnóstico se asegura únicamente con el seguimiento de los pacientes a largo plazo. De hecho, el de trastorno bipolar es longitudinalmente uno de los diagnósticos que con más frecuencia es modificado antes de estabilizarse de forma definitiva. En muchas ocasiones se producen cambios en su clasificación durante la evolución del trastorno, sobre todo hacia diagnósticos del espectro de la esquizofrenia[6,7].

El concepto de trastorno bipolar en la actualidad incluye una gran variedad de cuadros clínicos con una prevalencia similar en los distintos grupos étnicos, variando desde el 0.5 al 7.5 % de la población según los distintos estudios publicados. Sin embargo la mayoría de los expertos en este trastorno coinciden en afirmar que los datos actuales probablemente subestimen la prevalencia real de la enfermedad. Según la clasificación internacional de enfermedades en su última versión[8], el trastorno bipolar se define por la aparición de episodios reiterados, al menos dos, en los que el estado de ánimo y los niveles de actividad del enfermo

están profundamente alterados. Esta alteración puede ser de dos tipos: exaltación del estado de ánimo y aumento de la vitalidad y del nivel de actividad que correspondería con las fases maníacas o hipomaníacas del trastorno o bien una disminución del estado de ánimo y un descenso de la vitalidad y de la actividad que correspondería a las fases depresivas. Entre los episodios aislados se produciría una recuperación completa. La incidencia es similar en ambos sexos.

Pocos estudios han investigado hasta el momento el cambio diagnóstico en el trastorno bipolar o las relaciones entre los distintos diagnósticos en su evolución[9]. En la búsqueda bibliográfica encontramos sólo un autor[7] que analiza la estabilidad del diagnóstico de trastorno bipolar utilizando criterios CIE-10 para la clasificación de los pacientes. Otros autores han realizado estudios en base a la clasificación DSM-IV pero utilizando menor número de pacientes o reduciendo significativamente el período de seguimiento[3]. Varios estudios diferentes han abordado la problemática de la estabilidad de los diagnósticos tras primeros episodios psicóticos[2,6,10-12]. En nuestro estudio intentamos abordar estos aspectos utilizando una población importante de pacientes diagnosticados de trastorno bipolar que ha mantenido seguimiento durante once años lo que nos permite evaluar adecuadamente la evolución de sus diagnósticos en ese tiempo[13].

**Metodología:**

Se utiliza un registro clínico que incluye todas las actuaciones médicas llevadas a cabo en las consultas ambulatorias de psiquiatría para una población de aproximadamente 240.000 personas. Esta base de datos corresponde a dos Centros de Salud Mental (CSM) y también incluye las actuaciones médicas

realizadas tanto en el servicio de urgencias como en la unidad de hospitalización breve del hospital de tercer nivel que cubre esta área. Estos recursos cubren una zona heterogénea en el centro de la ciudad de Madrid en la que destaca el alto porcentaje de inmigración, que se corresponde con cerca del 21,7 % de la población actualmente [13,14]. 25152 pacientes recibieron asistencia desde el 1 de enero de 1992 al 31 de diciembre de 2004. El total de visitas realizadas por pacientes con algún diagnóstico de trastorno bipolar y al menos 10 consultas registradas fue de 71543. Los diagnósticos asignados en cada consulta se detallan en la tabla 1.

La amplitud del período de inclusión ha hecho que los códigos nosológicos empleados hayan variado en cierta medida a lo largo del estudio. Los clínicos asignaron los diagnósticos usando criterios CIE-10 y DSM-IV [8,15] aunque por razones administrativas fueron codificados a CIE-9[16]. En nuestro análisis se ha realizado la conversión previa y automática de todos los diagnósticos a CIE-10.

De ellos se consideraron los pacientes que en algún momento habían recibido el diagnóstico F31 (trastorno bipolar) según la CIE-10 y habían sido atendidos por lo menos 10 veces en los dispositivos asistenciales del área durante el mencionado periodo. 10 valoraciones diferentes garantizan un período mínimo de seguimiento de cinco meses contabilizando un intervalo entre consultas de 15 días, de esta forma se reduce el riesgo de que las valoraciones hayan sido realizadas por diferentes especialistas y de que la brevedad del seguimiento condicione el diagnóstico. 10025 pacientes fueron atendidos al menos en 10 ocasiones y de ellos 1153 recibieron al menos en una ocasión el diagnóstico F31.

Para el análisis de los datos se utilizó el índice kappa como medida de acuerdo entre el primer y último diagnóstico [17,18]. Se consideró como pacientes

correctamente diagnosticados desde el punto de la estabilidad diagnóstica a aquellos que habían recibido el diagnóstico de trastorno bipolar en un 75% de las consultas por lo menos [17,18]. Se calculó el intervalo de confianza de los distintos porcentajes utilizando el método de Wald [19].

**Resultados:**

Encontramos que del total de 1153 sujetos que al menos en una ocasión recibieron el diagnóstico de TB, un 23,1% (266/1153) lo mantuvieron en el 75% de las visitas. Los 1153 pacientes analizados recibieron una media de 62 asistencias. 342 pacientes fueron diagnosticados de trastorno bipolar en la primera consulta con una consistencia prospectiva (grado de coincidencia en la última consulta registrada) del 49,4%. 443 pacientes fueron diagnosticados de trastorno bipolar en la última consulta con una consistencia retrospectiva (coincidencia con la primera consulta) del 38,1%. Destaca que el acuerdo entre el primer y el último diagnóstico siendo ambos trastorno bipolar es bajo, Kappa=0,40. No hallamos diferencias entre los datos procedentes de las asistencias en consulta y aquellos procedentes de la unidad de hospitalización o de las atenciones en urgencias, ni tampoco en relación con la duración del seguimiento.

En el gráfico 1 se puede observar la diferencia en la evolución del diagnóstico de trastorno bipolar entre la primera y la última visita registrada para los pacientes del estudio. De un total de 342 pacientes diagnosticados de trastorno bipolar en la primera consulta, únicamente 158 mantuvieron el diagnóstico F31 en al menos un 75% de las consultas a lo largo del seguimiento, de modo que se diagnosticó a un 46,1% de los 266 pacientes considerados estables a lo largo del

seguimiento del estudio (158/342, IC95%: 40,9-50,1). Los 108 pacientes estables pero no diagnosticados en la primera consulta constituirían el error inicial (31%, 108/266; IC 95%: 26,6-36,5). En esta primera cita 184 de los 342 pacientes diagnosticados obtuvieron posteriormente al menos un 25% de diagnósticos diferentes del F31 y podrían ser considerados como sobrediagnóstico inicial o falsos positivos (FP), se corresponderían con un 53% de los diagnósticos iniciales (184/342; IC95%: 48,5-59,0).

En la última consulta aumentó considerablemente el número de diagnósticos de trastorno bipolar, 234 de los 443 pacientes que obtuvieron F31 en esta visita se corresponden con los 266 pacientes estables durante el seguimiento (88%). Sin embargo 209 pacientes (47,2%, 209/443; IC95% 42,5-51,8) diagnosticados como bipolares en esta última consulta no habían mantenido criterios de estabilidad en su evolución y podrían por lo tanto ser considerados como sobrediagnóstico final (FP). Consideramos que los 32 pacientes estables no diagnosticados en esta consulta constituirían el error final en el diagnóstico (12%, 32/266; IC95%: 8,1-15,9%).

Podemos observar que existe una gran variabilidad a lo largo del tiempo para estos pacientes. Hay un amplio margen de categorías que podrían actuar como factores de confusión en el diagnóstico de trastorno bipolar. Pero esto es especialmente cierto en el caso del espectro de esquizofrenia (F2), diagnóstico que llega a aparecer casi en una de cada cuatro visitas al psiquiatra de los pacientes incluidos en el estudio. A pesar de ello en muy pocos casos estos pacientes llegarían a cumplir el criterio de estabilidad diagnóstica establecido en nuestro estudio para la esquizofrenia u otros trastornos afines. La confusión parece menor pero también consistente con otras tres categorías: los trastornos de ansiedad (F4), los trastornos de personalidad (F6) y los trastornos por consumo de sustancias.

Como era de esperar la inmensa mayoría de los diagnósticos mantenidos de forma estable en la evolución del seguimiento se corresponden con el de trastorno bipolar.

En el gráfico 2 desglosamos la categoría F3 en sus distintas categorías, de este modo se puede observar que aunque la mayoría de los diagnósticos corresponden al F31 de trastorno bipolar, muchos de estos pacientes fueron incluidos en algún momento dentro de la categoría F33 (trastorno depresivo recurrente) e incluso con cierta frecuencia en la última visita registrada. No se llegan a cumplir en todo caso criterios de estabilidad para la mayoría de estos pacientes.

**Conclusiones:**

Las dificultades diagnósticas del trastorno bipolar destacan en el primer contacto con el especialista, bien por la dificultad en reconocer los síntomas iniciales o por la presentación “enmascarada” debido al abuso de tóxicos, a la clínica inicial depresiva o con síntomas psicóticos transitorios[7]. En nuestro estudio encontramos que la estabilidad del diagnóstico es baja, ya que menos del 25% de los pacientes mantienen el diagnóstico en su evolución y de ellos, menos del 60% son diagnosticados en su primer contacto con el especialista. Encontramos también frecuentes oscilaciones diagnósticas dentro del grupo de nuestro estudio. Estas oscilaciones se producen más frecuentemente hacia o desde diagnósticos como la esquizofrenia o los episodios psicóticos agudos pero también con trastornos relacionados con el uso de sustancias psicoactivas, trastornos de ansiedad y trastornos de personalidad. El principal factor de confusión por lo tanto y de

acuerdo con los trabajos previos, parece ser el relacionado con los diagnósticos del espectro de la esquizofrenia.

Varios aspectos diferencian nuestro trabajo de otros realizados con anterioridad. En primer lugar el registro incluía evaluaciones en tres diferentes escenarios: consultas ambulatorias, urgencias hospitalarias y unidad de hospitalización de agudos. Los diagnósticos fueron asignados por psiquiatras que desconocían los procedimientos del estudio y sus diagnósticos fueron otorgados dentro de la práctica clínica habitual, mientras en otros trabajos realizados a menudo la evaluación se realizaba mediante entrevistas estructuradas u otros métodos diagnósticos. De esta forma el estudio puede tener importancia en cuanto refleja el funcionamiento rutinario de las clasificaciones psiquiátricas al uso y se ajusta mejor a la realidad clínica. La alta representatividad de la muestra y el largo periodo de seguimiento favorecen asimismo los resultados del estudio.

La consistencia temporal del trastorno bipolar es menor en nuestro estudio que en otros realizados previamente[3,7], esto podría ser explicable porque en algunos de estos estudios el periodo de seguimiento fue de menor duración[6,11]. En todo caso la inestabilidad en los diagnósticos de enfermedad bipolar es llamativa considerando el probable sesgo presente en las evaluaciones al contar los clínicos con referencias previas de los pacientes y en muchos casos con el historial clínico. La propia evolución de la enfermedad podría explicar en parte las dificultades diagnósticas debido a su variabilidad, pero es posible que los actuales métodos de evaluación clínica requieran nuevas revisiones para asegurar su fiabilidad.

Del mismo modo los trabajos de investigación actuales (sobre todo en enfermedades crónicas como el trastorno bipolar) se basan en gran medida en estudios de seguimiento a corto plazo que pueden afectar a la validez de los

mismos al no existir tiempo suficiente para lograr la estabilización de los diagnósticos. También los estudios de prevalencia e incidencia se pueden ven influidos por la estabilidad de los diagnósticos que investigan. La medida en que esto ocurre está aún por determinar.

Los resultados resaltan la necesidad de un proceso diagnóstico longitudinal y de nuevas herramientas diagnósticas de mayor precisión[13]. Nuevos estudios que analicen los factores relacionados con la inestabilidad diagnóstica podrían establecer factores predictores de la misma, por ejemplo examinando la relación entre las distintas presentaciones semiológicas de la enfermedad en su comienzo y la evolución posterior.

En cuanto a las limitaciones de nuestro estudio, en primer lugar es destacable la probable infraestimación de la inestabilidad del diagnóstico por no incluir las evaluaciones realizadas antes del comienzo del mismo, en caso de que el paciente hubiera sido valorado antes por un psiquiatra. La ausencia de datos sobre la actividad de especialistas privados y el sesgo poblacional o cultural debido a la importancia de la población inmigrante en nuestro área son dificultades añadidas al estudio. Para asegurar un estudio adecuado de la estabilidad diagnóstica en el trastorno bipolar se debería probablemente realizar un registro más prolongado en el tiempo y con un mayor número de pacientes de modo que se pudiera investigar la evolución histórica de los diagnósticos de cada paciente incluyendo el inicio del trastorno.

## 6. Referencias:


1. Robins E, Guze SB. Establishment of diagnostic validity in psychiatric illness: its application to schizophrenia. Am J Psychiatry. 1970; 126: 983–7.
2. Fennig S, Kovasznay B, Rich C, Ram R, Pato C, Miller A, et al. Six-month stability of psychiatric diagnoses in first-admission patients with psychosis. Am J Psychiatry. 1994; 151: 1200-1208.
3. Chen YR, Swann AC, Johnson BA. Stability of diagnosis in bipolar disorder. J Nerv Ment Dis. 1998; 186: 17-23.
4. Marneros A, Deister A, Rohde A. Stability of diagnoses in affective, schizoaffective and schizophrenic disorders. Cross-sectional versus longitudinal diagnosis. Eur Arch Psychiatry Clin Neurosci. 1991; 241: 187-192.
5. Blacker D, Tsuang MT. Contested boundaries of bipolar disorder and the limits of categorical diagnosis in psychiatry. Am J Psychiatry. 1992; 149: 1473-1483.
6. Schwartz JE, Fennig S, Tanenberg-Karant M, Carlson G, Craig T, Galambos N, et al. Congruence of diagnosis 2 years after a first-admission diagnosis of psychosis. Arch Gen Psychiatry. 2000; 57: 593-600.
7. Kessing LV. Diagnostic stability in bipolar disorder in clinical practise as according to ICD-10. J Affect Disord. 2005; 85: 293-299.
8. World Health Organization. Tenth revision of the International Classification of Diseases and related health problems (ICD-10). Geneva (Switzerland): World Health Organization, 1992.

9. Sánchez González A, Baca-García E, Diaz-Sastre C, Baca Baldomero E. Los límites del diagnóstico de trastrono bipolar. Actas Luso-Españolas de Neurología, Psiquiatría y Ciencias Afines 1998; 26 (1): 12-19.
10. Amin S, Singh SP, Brewin J, Jones PB, Medley I, Harrison G. Diagnostic stability of first-episode psychosis. Comparison of ICD-10 and DSM-III-R systems. Br J Psychiatry. 1999; 175: 537–543.
11. Schimmelmann BG, Conus P, Edwards J, McGorry PD, Lambert M. Diagnostic stability 18 months after treatment initiation for first-episode psychosis. J Clin Psychiatry 2005; 66: 1239-1246.
12. Rufino ACTBF, Uchida RR, Vilela JAA, Marques JMA, Zuardi AW, Del-Ben CM. Stability of the diagnosis of first –episode psychosis made in an emergency setting. Gen Hosp Psychiatry. 2005; 27: 189-193.
13. Perez-Rodriguez MM, Baca-Garcia E, Quintero J, Burguillo F, Saiz D, Lozano MC, Botillo C, Diaz-Marsa M, De Rivera JL. Psychiatric care and emigration. A survey in a Spanish hospital's catchments area. European Journal of Psychiatry 2005; 19 (2): 107-111.
14. Perez-Rodriguez MM, Baca-Garcia E, Quintero-Gutierrez J, Gonzalez G, Saiz-Gonzalez D, Bustillo C, Basurte Vilamor I, Sevilla J, De Rivera JL. Demand for psychiatric emergency services and emigration. Findings in a Spanish hospital's catchments area during the year 2003. Eur J Public Health 2006 Aug;16(4):383-7.
15. American Psychiatric Association. Diagnostic and Statistical Manual of Mental Disorders. 4th ed. Washington, DC: American Psychiatric Association, 1994.

16. Organización Mundial de la Salud. Novena revisión de la Clasificación Internacional de Enfermedades, modificación clínica, 2ª edición (CIE-9 MC). Madrid, Ministerio de Sanidad y Consumo, 1994.
17. Baca-Garcia E, Perez-Rodriguez MM, Basurte-Villamor I, Lopez-Castroman J, Fernandez del Moral AL, Jimenez-Arriero MA et al. Diagnostic stability and evolution of bipolar disorder in clinical practice: a prospective cohort study. Acta Psychiatr Scand. 2007 Jun; 115(6):473-80.
18. Baca-Garcia E, Perez-Rodriguez MM, Basurte-Villamor I, Fernandez del Moral AL, Jimenez-Arriero MA, Gonzalez de Rivera JL et al. Diagnostic stability of psychiatric disorders in clinical practice. Br J Psychiatry. 2007 Mar; 190: 210-6.
19. Baca García E, Pérez Rodríguez MM, Basurte Villamor I, López Castromán J, Fernández del Moral A, Jiménez Arriero MA, González de Rivera JL, Saiz J, Oquendo MA, De Leon J. Diagnostic stability and evolution of bipolar disorder in clinical practice: A prospective cohort study. En prensa.
20. Arce Cordon R, Perez-Rodriguez MM, Baca-Baldomero E, Oquendo MA, Baca-Garcia E. Routine laboratory screening in newly admitted psychiatric patients: Is it worthwhile? Psychiatry Services. En prensa.

| | |
|---|---|
| Trastorno Bipolar | 31,6% |
| Esquizofrenia | 23,9% |
| Trastornos de ansiedad | 9,1% |
| Depresión crónica | 6,8% |
| Trastornos de personalidad | 3,7% |
| Consumo de sustancias | 2,2% |
| Trastornos psicosomáticos | 0,8% |
| Depresión | 0,5% |

Tabla 1: Clasificación de los diagnósticos en función de la frecuencia con que fueron emitidos durante las consultas de los pacientes registrados en el estudio.

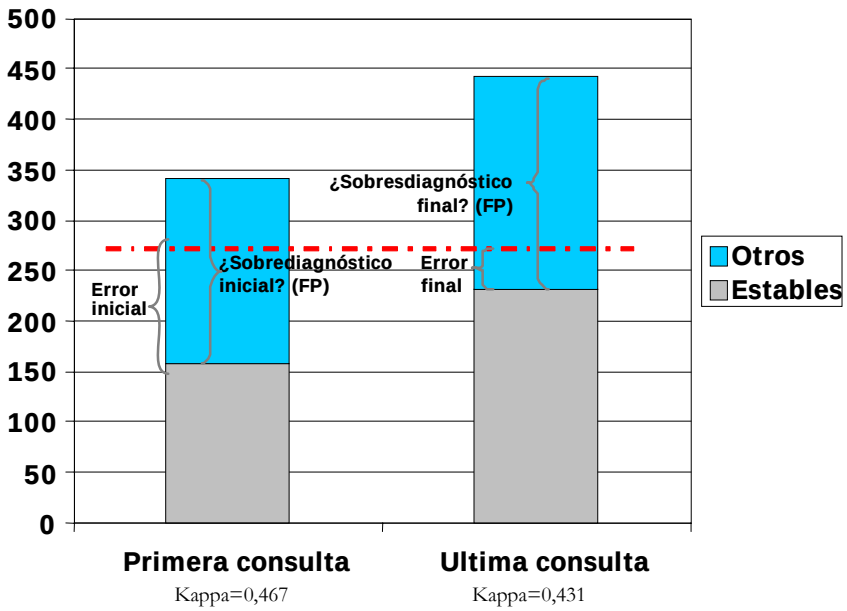



Gráfico 1: Comparación de diagnósticos de trastorno bipolar (F31) entre la primera y la última consulta. La línea roja representa el número de pacientes que han mantenido este diagnóstico en más del 75% de las consultas.

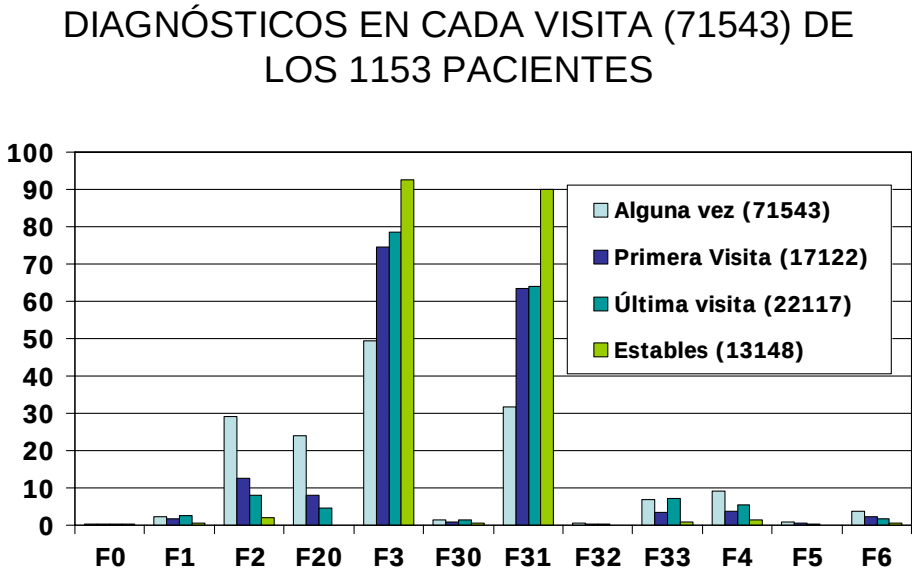



Grafico 2: Diagnósticos otorgados a los 1153 pacientes que en alguna ocasión fueron evaluados como trastorno bipolar.

# BIBLIOGRAFY

(1) Robins E, Guze SB. Establishment of diagnostic validity in psychiatric illness: its application to schizophrenia. Am J Psychiatry 1970;126:983-7.

(2) Fennig S, Kovasznay B, Rich C, et al. Six-month stability of psychiatric diagnoses in first-admission patients with psychosis. Am J Psychiatry 1994;151:1200-8.

(3) Kim-Cohen J, Caspi A, Moffitt TE, Harrington H, Milne BJ, Poulton R. Prior juvenile diagnoses in adults with mental disorder: developmental follow-back of a prospective-longitudinal cohort. Arch Gen Psychiatry 2003 Jul;60(7):709-17.

(4) Whitty P, Clarke M, McTigue O, Browne S, Kamali M, Larkin C, et al. Diagnostic stability four years after a first episode of psychosis. Psychiatric Services 2005;56(9):1084-8.

(5) Spitzer RL, Endicott J, Robins E. Clinical criteria for psychiatric diagnosis and DSM-III. Am J Psychiatry 1975;132(11):1187-92.

(6) Spitzer RL, Endicott J, Robins E. Research diagnostic criteria: rationale and reliability. Arch Gen Psychiatry 1978;35(773):782.

(7) Chen Y, Richard D, Swann AC, Johnson BA. Stability of Diagnosis in Bipolar Disorder. The Journal of Nervous and Mental Disease 1998;186(1):17-23.

(8) Marneros A, Deister A, Rohde A. Stability of diagnoses in affective, schizoaffective and schizophrenic disorders. Cross-sectional versus longitudinal diagnosis. European Archives Psychiatry Clinical Neurosciences 1991;241:187-92.

(9) Taylor MA. Are schizophrenia and affective disorders related? A selective literature review. Am J Psychiatry 1992;149(1):22-32.

(10) Blacker D, Tsuang MT. Contested boundaries of bipolar disorder and the limits of categorical diagnosis in psychiatry. Am J Psychiatry 1992;149(11):1473-83.

(11) Goodwin FK, Jamison KR. Manic-Depressive Illness. New York: Oxford University Press; 1990.

(12) Baca-Garcia E, Perez-Rodriguez MM, Basurte-Villamor I, Fernandez del Moral AL, Jimenez-Arriero MA, Gonzalez de Rivera JL, et al. Diagnostic

stability of psychiatric disorders in clinical practice. Br J Psychiatry 2007;190:210-6.

(13) Baca-Garcia E, Perez-Rodriguez MM, Basurte-Villamor I, López-Castromán J, Fernandez del Moral AL, Jimenez-Arriero MA, et al. Diagnostic stability and evolution of bipolar disorder in clinical practice: a prospective cohort study. Acta Psychiatr Scand 2007;115(6):473-80.

(14) Kessing LV. Diagnostic stability in bipolar disorder in clinical practise as according to ICD-10. Journal of Affective Disorders 2005;85:293-9.

(15) Rice JP, McDonald-Scott P, Endicott J, Coryell W, Grove WM, Keller MB, et al. The stability of diagnosis with an application to bipolar II disorder. Psychiatry Research 1986;19(4):285-96.

(16) Weeke A. Admission pattern and diagnostic stability among unipolar and bipolar manic-depressive patients. Acta Psychiatr Scand 1984;70:603-13.

(17) Schwartz JE, Fennig S, Tanenberg-Karant M, Carlson G, Craig T, Galambos N, et al. Congruence of diagnoses 2 years after a first-admission diagnosis of psychosis. Arch Gen Psychiatry 2000;57:593-600.

(18) Rufino A, Uchida R, Vilela J, Marques J, Zuardi A, Del-Ben C. Stability of the diagnosis of first-episode psychosis made in an emergency setting. Gen Hosp Psychiatry 2005;27:189-93.

(19) Amin S, Singh SP, Brewin J, Jones PB, Medley I, Harrison G. Diagnostic stability of first-episode psychosis. Comparison of ICD-10 and DSM-III-R systems. Br J Psychiatry 1999;(175):537-43.

(20) Mojtabai R, Susser ES, Bromet EJ. Clinical characteristics, 4-year course, and DSM-IV classification of patients with nonaffective acute remitting psychosis. Am J Psychiatry 2003 Dec;160(12):2108-15.

(21) Jørgensen P, Bennedsen B, Christensen J, Hyllested A. Acute and transient psychotic disorder: a 1-year follow-up study. Acta Psychiatr Scand 1997;96(2):150-4.

(22) Schimmelmann BG, Conus P, Edwards J, McGorry PD, Lambert M. Diagnostic stability 18 months after treatment initiation for first-episode psychosis. Journal Clinical Psychiatry 2005;66(10):1239-46.

(23) Veen N, Selten JP, Schols D, Laan W, Hoek HW, Van der Tweel I, et al. Diagnostic stability in a Dutch psychosis incidence cohort. Br J Psychiatry 2004;185:460-4.

(24) Addington J, Chaves A, Addington D. Diagnostic stability over one year in first-episode psychosis. Schizophrenia Research 2006;86:71-5.

(25) Culverhouse R, Bucholz KK, Crowe RR, Hesselbrock V, Nurnberger JI, Jr., Porjesz B, et al. Long-term stability of alcohol and other substance dependence diagnoses and habitual smoking: an evaluation after 5 years. Arch Gen Psychiatry 2005 Jul;62(7):753-60.

(26) Grilo CM, Sanislow CA, Gunderson JG, Pagano ME, Yen S, Zanarini MC, et al. Two-year stability and change of schizotypal, borderline, avoidant, and obsessive-compulsive personality disorders. J Consult Clin Psychol 2004 Oct;72(5):767-75.

(27) McGlashan TH, Grilo CM, Sanislow CA, Ralevski E, Morey LC, Gunderson JG, et al. Two-year prevalence and stability of individual DSM-IV criteria for schizotypal, borderline, avoidant, and obsessive-compulsive personality disorders: toward a hybrid model of axis II disorders. Am J Psychiatry 2005 May;162(5):883-9.

(28) Shea MT, Stout R, Gunderson J, Morey LC, Grilo CM, McGlashan T, et al. Short-term diagnostic stability of schizotypal, borderline, avoidant, and obsessive-compulsive personality disorders. Am J Psychiatry 2002 Dec;159(12):2036-41.

(29) Kessing LV. Diagnostic stability in depressive disorder as according to ICD-10 in clinical practice. Psychopathology 2005 Jan;38(1):32-7.

(30) Lieb R, Zimmermann P, Friis RH, Hofler M, Tholen S, Wittchen HU. The natural course of DSM-IV somatoform disorders and syndromes among adolescents and young adults: a prospective-longitudinal community study. Eur Psychiatry 2002 Oct;17(6):321-31.

(31) Tsuang MT, Woolson RF, Winokur G, Crowe RR. Stability of psychiatric diagnosis. Schizophrenia and affective disorders followed up over a 30- to 40-year period. Arch Gen Psychiatry 1981;38(5):535-9.

(32) Barkow K, Heun R, Wittchen HU, Bedirhan UT, Gansicke M, Maier W. Mixed anxiety-depression in a 1 year follow-up study: shift to other diagnoses or remission? J Affect Disord 2004 Apr;79(1-3):235-9.

(33) Chen Y, Richard D, Swann AC, Burt DB. Stability of diagnosis in schizophrenia. Am J Psychiatry 1996;153(5):682-7.

(34) Cooper JE. Diagnostic change in a longitudinal study of psychiatric patients. Br J Psychiatry 1967;113:129-42.

(35) Lenz G, Simhandi C, Thau K, Berner P, Gabriel E. Temporal stability of diagnostic criteria for funtional psychoses. Results from the Vienna follow-up study. Psychopathology 1991;24:328-35.

(36) Amara IB. Consistency in the diagnosis of the functional psychoses. Can Psychiatr Assoc J 1978;23(5):329-36.

(37) Babigian HM, Gardner EA, Miles HC, Romano J. Diagnostic consistency and change in a follow-up study of 1215 patients. Am J Psychiatry 1965;121:895-901.

(38) Kendell RE. The stability of psychiatric diagnosis. Br J Psychiatry 1974;124:352-6.

(39) Rabinowitz J, Slyuzberg M, Ritsner M, Mark M, Popper M, Ginath Y. Changes in diagnosis in a 9-year national longitudinal sample. Compr Psychiatry 1994;35(5):361-5.

(40) Stanton MW, Joyce PR. Stability of psychiatric diagnoses in New Zealand psychiatric hospitals. Austr New Zeal J Psychiatry 1993;27(1):2-8.

(41) Lee AS, Murray RM. The long-term outcome of Maudsley depressives. Br J Psychiatry 1988;153:741-51.

(42) Mason P, Harrison G, Croudace T, Glazebrook C, Medley I. The predictive validity of a diagnosis of schizophrenia. A report from the International Study of Schizophrenia (ISoS) coordinated by the World Health Organization and the Department of Psychiatry, University of Nottingham. Br J Psychiatry 1997;170:321-7.

(43) Daradkeh TK, El-Rufaie OEF, Younis YO, Ghubash R. The diagnostic stability of ICD-10 psychiatric diagnoses in clinical practice. Eur Psychiatry 1997;12:136-9.

(44) Zhang-Wong J, Beiser M, Bean G, Iacono WG. Five-year course of schizophreniform disorder. Psychiatry Research 1995;59(1-2):109-17.

(45) Weissman MM, Wolk S, Goldstein RB, Moreau D, Adams P, Greenwald S, et al. Depressed adolescents grown up. JAMA 1999 May 12;281(18):1707-13.

(46) Mattanah JJ, Becker DF, Levy KN, Edell WS, McGlashan TH. Diagnostic stability in adolescents followed up 2 years after hospitalization. Am J Psychiatry 1995 Jun;152(6):889-94.

(47) Helgeland MI, Torgersen S. Stability and prediction of schizophrenia from adolescence to adulthood. Eur Child Adolesc Psychiatry 2005 Mar;14(2):83-94.

(48) Hollis C. Adult outcomes of child- and adolescent-onset schizophrenia: diagnostic stability and predictive validity. Am J Psychiatry 2000 Oct;157(10):1652-9.

(49) Jarbin H, von Knorring AL. Diagnostic stability in adolescent onset psychotic disorders. Eur Child Adolesc Psychiatry 2003 Jan;12(1):15-22.

(50) Maziade M, Gingras N, Rodrigue C, Bouchard S, Cardinal A, Gauthier B, et al. Long-term stability of diagnosis and symptom dimensions in a systematic sample of patients with onset of schizophrenia in childhood and early adolescence. I: nosology, sex and age of onset. Br J Psychiatry 1996 Sep;169(3):361-70.

(51) McClellan J, McCurry C. Early onset psychotic disorders: diagnostic stability and clinical characteristics. Eur Child Adolesc Psychiatry 1999;8 Suppl 1:I13-I19.

(52) Chanen AM, Jackson HJ, McGorry PD, Allot KA, Clarkson V, Yuen HP. Two-year stability of personality disorder in older adolescent outpatients. J Personal Disord 2004 Dec;18(6):526-41.

(53) Grilo CM, Becker DF, Edell WS, McGlashan TH. Stability and change of DSM-III-R personality disorder dimensions in adolescents followed up 2 years after psychiatric hospitalization. Compr Psychiatry 2001 Sep;42(5):364-8.

(54) Harrington R, Fudge H, Rutter M, Pickles A, Hill J. Adult outcomes of childhood and adolescent depression. I. Psychiatric status. Arch Gen Psychiatry 1990 May;47(5):465-73.

(55) Kovacs M, Feinberg TL, Crouse-Novak M, Paulauskas SL, Pollock M, Finkelstein R. Depressive disorders in childhood. II. A longitudinal study of the risk for a subsequent major depression. Arch Gen Psychiatry 1984;41(7):643-9.

(56) Valevski A, Ratzoni G, Sever J, Apter A, Zalsman G, Shiloh R, et al. Stability of diagnosis: a 20-year retrospective cohort study of Israeli psychiatric adolescent inpatients. J Adolesc 2001 Oct;24(5):625-33.

(57) Colman I, Wadsworth ME, Croudace TJ, Jones PB. Forty-year psychiatric outcomes following assessment for internalizing disorder in adolescence. Am J Psychiatry 2007 Jan;164(1):126-33.

(58) Pettit JW, Morgan S, Paukert AL. The stability of axis I diagnoses in youth across multiple psychiatric hospitalizations. Child Psychiatry Hum Dev 2005;36(1):53-71.

(59) Bernstein DP, Cohen P, Velez CN, Schwab-Stone M, Siever LJ, Shinsato L. Prevalence and stability of the DSM-III-R personality disorders in a community-based survey of adolescents. Am J Psychiatry 1993;150(8):1237-43.

(60) Geller B, Zimerman B, Williams M, Bolhofner K, Craney JL. Bipolar disorder at prospective follow-up of adults who had prepubertal major depressive disorder. Am J Psychiatry 2001;158:125-7.

(61) Fraguas D, de Castro MJ, Medina O, Parellada M, Moreno D, Graell M, et al. Does Diagnostic Classification of Early-Onset Psychosis Change Over Follow-Up? Child Psychiatry Hum Dev 2007.

(62) Baca-Garcia E, Perez-Rodriguez MM, Basurte-Villamor I, Quintero Gutiérrez del Álamo J, Sevilla-Vicente J, Matinez-Vigo M, et al. Patterns of mental health service utilization in a general hospital and outpatient mental health facilities: analysis of 365,262 psychiatric consultations. Eur Arch Psychiatry Clin Neurosci 2007;xx:1-8.

(63) Basurte Villamor I. Análisis de itinerarios y coste en salud mental en un area metropolitana de madrid durante 5 años. Departamento de Psiquiatría. Universidad Complutense; 2006.

(64) Baca-Garcia E, Carballo JJ, Blanco C, Perez-Rodriguez MM, Jimenez-Arriero MA, Ferre-Navarrete F, et al. Diagnostic stability of anxiety disorders in infants, children and adolescents: a 13-year-follow-up study of 1,869 patients (1992-2005). 2007.
Ref Type: Unpublished Work

(65) López Castromán J, Baca García E, Botillo Martín C, Quintero Gutiérrez del Álamo J, Navarro Jiménez R, Negueruela López M, et al. Errores de diagnóstico y estabilidad en el trastorno bipolar. Act Esp Psiquiatr . 2007.
Ref Type: Unpublished Work

(66) Beiser M, Iacono WG, Erickson D. Temporal stability in the major mental disorders. In: L.N.Robins & J.E.Barrett, editor. The validity of psychiatric diagnosis.New York: Raven Press; 2007. p. 77-98.

(67) Kendler KS. Toward a scientific psychiatric nosology. Strengths and limitations. Arch Gen Psychiatry 1990;47:969-73.

(68) Sánchez González A, Baca-García E, Diaz-Sastre C, Baca Baldomero E. Los límites del diagnóstico de trastorno bipolar. Actas Luso-Españolas de Neurología, Psiquiatría y Ciencias Afines 1998;26(1):12-9.

(69) Sajatovic M. Bipolar disorder: Disease Burden. The American Journal of Managed Care 2005;11(3, sup.):80-3.

(70) Weissman MM, Bland RC, Camino GJ, et al. Cross-national epidemiology of major depression and bipolar disorder. JAMA 1996;276:293-9.

(71) Kessler RC, Chiu WT, Demler O, Merikangas KR, Walters EE. Prevalence, severity, and comorbidity of 12-month *DSM-IV* disorders in the National Comorbidity Survey Replication. Arch Gen Psychiatry 2005;62(617):627.

(72) Bijl RV, Ravelli A, van Zessen G. Prevalence of psychiatric disorder in the general population: results of the Netherlands Mental Health Survey and Incidence Study (NEMESIS). Soc Psychiatry Psychiatr Epidemiol 1998;33:581-6.

(73) ten Have M, Vollenberg W, Bija R, Nolen WA. Bipolar disorder in the general population in the Netherlands (prevalence, consequences and care utilisation): results from the Netherlands Mental Health Survey and Incidence Study (NEMESIS). Journal of Affective Disorders 2002;68:203-13.

(74) Grant BF, Stinson FS, Hasin DS, Dawson DA, Chou SP, Ruan WJ, et al. Prevalence, correlates, and comorbidity of bipolar I disorder and axis I and II disorders: results from the National Epidemiologic Survey on Alcohol and Related Conditions. Journal Clinical Psychiatry 2005;66:1205-15.

(75) Angst J. The emerging epidemiology of hypomania and bipolar II disorder. Journal of Affective Disorders 1998;50:143-51.

(76) Dunner DL. Clinical consequences of under-recognized bipolar spectrum disorder. Bipolar Disorders 2003;5:456-63.

(77) Hirschfeld RM. Bipolar spectrum disorder: improving its recognition and diagnosis. The Journal of Clinical Psychiatry 2001;62(14):5-9.

(78) Carta MG, Angst J. Epidemiological and clinical aspects of bipolar disorders: controversies or a common need to redefine the aims and methodological aspects of surveys. Clin Pract Epidemol Ment Health 2007;1(4):1-4.

(79) Hirschfeld RM, Lewis L, Vornik LA. Perceptions and impact of bipolar disorder: how far have we really come? Results of the national depressive and manic-depressive association 2000 survey of individuals with bipolar disorder. The Journal of Clinical Psychiatry 2003;64(2):161-74.

(80) Lish JD, Dime-Meenan S, Whybrow PC, Price RA, Hirschfeld RM. The National Depressive and Manic-depressive Association (DMDA) survey of bipolar members. Journal of Affective Disorders 1994;31(4):281-94.

(81) Ghaemi NS, Sachs GS, Chiou AM. Is bipolar disorder still under diagnosed? Are the antidepressants over utilized? Journal of Affective Disorders 1999;52:135-44.

(82) Angst J, Gamma A, Lewinsohn PM. The evolving epidemiology of bipolar disorder. World Psychiatry 2002;1:146-8.

(83) Berk M, Berk L, Moss K, Dodd S, Malhi GS. Diagnosing bipolar disorder: how can we do it better? The Medical Journal of Australia 2006 May 1;184(9):459-62.

(84) Baldessarini RJ, Tondo L, Hennen J. Treatment delays in bipolar disorders. Am J Psychiatry 1999;156(5):811-2.

(85) Chen YW, Dilsaver SC. Lifetime rates of suicide attempts among subjects with bipolar and unipolar disorders relative to subjects with other Axis I disorders. Biol Psychiatry 1996;39:896-9.

(86) Tondo L, Isacsson G, Baldessarini R. Suicidal behaviour in bipolar disorder: risk and prevention. CNS Drugs 2003;17:491-511.

(87) World Health Organisation. The World Health Report 2001; Mental Health: New Understanding, New Hope. World Health Organization,Geneva; 2001.

(88) Hirschfeld RM, Vornik LA. Bipolar Disorder - Costs and Comorbidity. The American Journal of Managed Care 2005;11(3, sup.):85-90.

(89) McElroy SL. Diagnosing and treating comorbid (complicated) bipolar disorder. The Journal of Clinical Psychiatry 2004;65(15):35-44.

(90) Vieta E, Colom F, Corbella B, et al. Clinical correlates of psychiatric comorbidity in bipolar I patients. Bipolar Disorders 2001;3:253-8.

(91) Sanjuán J. ¿Qué Psiquiatría biológica deben aprender los psiquiatras del futuro? Interpsiquis . 2001.
Ref Type: Internet Communication

(92) Zachar P, Kendler KS. Psychiatric Disorders: A Conceptual Taxonomy. Am J Psychiatry 2007;164:557-65.

(93) Kendler KS. Reflections on the Relationship Between Psychiatric Genetics and Psychiatric Nosology. Am J Psychiatry 2006;163:1138-46.

(94) Segurado R, Detera-Wadleigh SD, Levinson DF, Lewis CM, et al. Genome scan meta-analysis of schizophrenia and bipolar disorder, part III: bipolar disorder. Am J Hum Genet 2003;73:49-62.

(95) Krishnan K. Psychiatric disease in the genomic era: rational approach. Molecular Psychiatry 2005;10:978-84.

(96) Boutros N, et al. A Four-Step Approach for Developing Diagnostic Test in Psychiatry. The Journal of Neuropsychiatry and Clinical Neurosciences 2005;17:455-64.

(97) Leckman JF, Sholomskas D, Thompson D, Bélanger A, Weissman MM. Best estimate of lifetime psychiatric diagnosis. Arch Gen Psychiatry 1982;39:879-83.

(98) Kosten TA, Rounsaville BJ. Sensitivity of psychiatric diagnosis based on the estimate procedure. Am J Psychiatry 1992;149(9):1225-7.

(99) Roy MA, Lanctôt C, Mérette C, Cliche D, et al. Clinical and Methodological Factors Related to Reliability of the Best-Estimate Diagnostic Procedure. Am J Psychiatry 1997;154(12):1726-33.

(100) Torrens M, Serrano D, Astals M, Pérez-Domínguez G, Martín-Santos R. Diagnosing comorbid psychiatric disorders in substance abusers: validity of the Spanish versions of the Psychiatric Research Interview for Substance and Mental Disorders and the Structured Clinical Interview for DSM-IV. Am J Psychiatry 2004;161(7):1231-7.

(101) Spitzer RL. Psychiatric diagnosis: are clinicians still necessary? Compr Psychiatry 1983;24:399-411.

(102) Perala J, Suvisaari J, Saarni SI, Kuoppasalmi K, Isometsa E, Pirkola S, et al. Lifetime Prevalence of Psychotic and Bipolar I Disorders in a General Population. Arch Gen Psychiatry 2007;64:19-28.

(103) Gelder MG, Lopez Ibor Jr JJ, Andreasen N. Tratado de psiquiatría. Barcelona: 2003.

(104) Vallejo Ruiloba J, Leal Cercós C. Tratado de psiquiatría. Barcelona: 2005.

(105) Burger H, Neeleman J. A glossary on psychiatric epidemiology. J Epidemiol Community Health 2007;61:185-9.

(106) Kendler KS. The nosologic validity of paranoia (simple delusional disorder): a review. Arch Gen Psychiatry 1980;37:699-706.

(107) Andreasen NC. The validation of psychiatric diagnosis: new models and approaches (editorial). Am J Psychiatry 1995;152:161-2.

(108) First MB, Pincus HA, Levine JB, Williams JBW, Ustun B, Peele R. Clinical utility as a criterion for revising psychiatric diagnoses. Am J Psychiatry 2004;161(6):946-54.

(109) Kendell R, Jablensky A. Distinguishing between the validity and utility of psychiatric diagnoses. Am J Psychiatry 2003;160:4-12.

(110) Mezzich JE, Berganza CE. Purposes and models of diagnostic systems. Psychopathology 2005;38:162-5.

(111) Helmuth L. In Sickness and in Health? Science 302, 808-810. 2003.
Ref Type: Magazine Article

(112) IGDA workgroup W. IGDA. I: Conceptual bases - historical, cultural and clinical perspectives. Br J Psychiatry 2003;182(45):s40-s41.

(113) Kendell RE. The diagnosis of mania. Journal of Affective Disorders 1985;8:207-13.

(114) Bartko JJ. Measurement and reliability: statistical thinking considerations. Schizophr Bull 1991;17(3):483-9.

(115) Langenbucher J, Labouvie E, Morgenstern J. Measuring diagnostic agreement. J Consult Clin Psychol 1996;64(6):1285-9.

(116) Altman DG, Machin D, ryant TN, Gardner MJ. Statistics with Confidence: Confidence Intervals and Statistical Guidelines. 2nd ed. London: BMJ Books; 2000.

(117) World Health Organisation. Tenth revision of the International Classification of Diseases and related health problems (ICD-10). Geneva: World Health Organisation; 1992.

(118) Thomas P. The many forms of bipolar disorder: a modern look at an old illness. Journal of Affective Disorders 2004;79:s3-s8.

(119) Managing bipolar disorder: misdiagnosis and quality of life. Am J Manag Care 2005;11(9 Suppl):S267-S270.

(120) ESEMeD /MHEDEA 2000 investigators. Prevalence of mental disorders in Europe: results from the ESEMeD European Study of Epidemiology of Mental Disorders (ESEMeD) project. Acta Psychiatr Scand 2004;109(Suppl 420):21-7.

(121) Alonso J, Angermeyer MC, Lépine JP. The European Study of the Epidemiology of Mental Disorders (ESEMeD) project: an epidemiological basis for informing mental health policies in Europe. Acta Psychiatr Scand 2004;109(Suppl 420):5-7.

(122) Bebbington P, Ramana R. The epidemiology of bipolar affective disorder. Soc Psychiatry Psychiatr Epidemiol 1995;30:279-92.

(123) Vazquez-Barquero JL, Diez-Manrique JF, Peña C, Quintanal RG, Labrador Lopez M. Two stage design in a community survey. Br J Psychiatry 1986;149:88-97.

(124) Vazquez-Barquero JL, Diez-Manrique JF, Peña C, Aldama J, Rodriguez CS, Arango JM, et al. A community mental health survey in Cantabria: a general discription of morbidity. Psychol Med 1987;17:227-41.

(125) Canals J, Domenech E, Carbajo C, Blade J. Prevalence of DSMIII-R and ICD-10 psychiatric disorders in a Spanish population of 18-year-olds. Acta Psychiatr Scand 1997;96:287-94.

(126) Pini S, de Queiroz V, Pagnin D, Pezawas L, Angst J, Cassano GB, et al. Prevalence and burden of bipolar disorders in European countries. Eur Neuropsychopharmacol 2005;15(4):425-34.

(127) Waraich P, Goldner EM, Somers JM, Hsu L. Prevalence and incidence studies of mood disorders: a systematic review of the literature. Can J Psychiatry 2004;49:124-38.

(128) Jefferson JW. Trastornos bipolares: guía breve de diagnóstico y tratamiento. In: Medical Trends SL, editor. Clinical psychiatry self assessment program: Trastornos del estado de ánimo.Barcelona: 2006.

(129) Akiskal HS, Bourgeois ML, Angst J, Post R, Möller HJ, Hirschfeld RM. Re-evaluating the prevalence of and diagnostic composition within the broad clinical spectrum of bipolar disorders. Journal of Affective Disorders 2000;59(1):S5-S30.

(130) Berk M, Dodd S. El trastorno bipolar de tipo II: una revisión. Bipolar Disorders (Ed Esp) 2006;4:1-11.

(131) Angst J. Epidemiologie du spectre bipolaire. Encephale 1995;Supl. VI:37-42.

(132) Benazzi F. Borderline personality-bipolar spectrum relationship. Progress in Neuro-psychopharmacology and Biological Psychiatry 2006;30(1):68-74.

(133) Tsuang MT, Winokur C, Crowe RR. Morbidity risks of schizophrenia and affective disorders among first degree relatives of patients with schizophrenia, mania, depression and surgical conditions. Br J Psychiatry 1980;137:497-504.

(134) Angst J, Gamma A, Sellaro R, Lavori PW, Zhang H. Recurrence of bipolar disorders and major depression. A life-long perspective. Eur Arch Psychiatry Clin Neurosci 2003;253:236-40.

(135) Lewinsohn PM, Rohde P, Klein DN, Seeley JR. Natural course of adolescent major depressive disorder: I. Continuity into young adulthood. J Am Acad Child Adolesc Psychiatry 1999 Jan;38(1):56-63.

(136) Lewinsohn PM, Seeley JR. Bipolar disorder in adolescents. Epidemiology and suicidal behavior in Child and early adolescent bipolar disorder. New York: Guilford Press; 2003.

(137) Thase ME. Bipolar depression: diagnostic and treatment considerations. Dev Psychopathol 2006;18(4):1213-30.

(138) Leboyer M, Henry C, Paillere-Martinot ML, Bellivier F. Age at onset in bipolar affective disorders: a review. Bipolar Disorders 2005;7:111-8.

(139) Emilien G, Septien L, Brisard C, Corruble E, Bourin M. Bipolar disorder: how far are we from a rigorous definition and effective management? Prog Neuropsychopharmacol Biol Psychiatry 2007;31(5):975-96.

(140) Hantouche EG., Azorin JM, Chatenet-Duchene L, Lancrenon S, Allilaire J.F, Akiskal HS. Caracterisation de la manie dans une cohorte nationale de 1090 patients dans «EPIMAN-II Mille»: frequence des sous-types cliniques, debut et errances diagnostiques. Ann Med Psychol 2003;161:359-66.

(141) Kohn R, Saxena Sh, Levav I, Sarraceno B. The treatment gap in mental health care. Bulletin of the World Health Organization 2004;82(11):858-66.

(142) Li J, McCombs JS, Stimmel GL. Cost of treating bipolar disorder in the California Medicaid Program. Journal of Affective Disorders 2002;71(131-139).

(143) Fagiolini A, Kupfer DJ, Masalehdan A, et al. Functional impairment in the remission phase of bipolar disorder. Bipolar Disorders 2005;7:281-5.

(144) Chisholm D, Van Ommeren M, Ayuso-Mateos JL, Saxena Sh. Cost-effectiveness of clinical interventions for reducing the global burden of bipolar disorder. Br J Psychiatry 2005;187:559-67.

(145) Seva A. El coste de las patologías psiquiátricas en España. Un seguimiento de 26 años y 10.974 ingresos en una Unidad Psiquiátrica de Corta Estancia de un hospital general. European Journal of Psychiatry 2002;16(1):57-67.

(146) Fernández A, Haro JM, Codony M, Vilagut G, Martínez-Alonso M, Autonell J, et al. Treatment adequacy of anxiety and depressive disorders: primary versus specialised care in Spain. Journal of Affective Disorders 2006;96(1-2):9-20.

(147) Strakowski SM, Tohen M, Stoll AL, Faedda GL, GoodwinDC. Comorbidity in mania at first hospitalization. Am J Psychiatry 1992;149:554-6.

(148) Kessler RC, McGonagle KA, Zao S. Lifetime and 12-month prevalence of DSM-III-R psychiatric disorders in the United States. Results from the National Comorbidity Survey. Arch Gen Psychiatry 1994;51:8-19.

(149) Cassano GB, Pini S, Saettoni M, Rucci P, Dell'Osso L. Occurrence and clinical correlates of psychiatric comorbidity in patients with psychotic disorders. J Clin Psychiatry 1998;59:60-8.

(150) Black DW, Winokur G, Bell S, Nasrallah A, Hulbert J. Complicated mania: comorbidity and immediate outcome in the treatment of mania. Arch Gen Psychiatry 1988;a45:232-6.

(151) McElroy SL, Altshuler LL, Suppes T, Keck PE Jr, Frye MA, Denicoff KD, et al. Axis I psychiatric comorbidity and its relationship to historical illness variables in 288 patients with bipolar disorder. Am J Psychiatry 2001;158(3):420-6.

(152) Cassidy F, Ahearn EP, Carroll BJ. Substance abuse in bipolar disorder. Bipolar Disorders 2001;3:181-8.

(153) Regier DA, Farmer ME, Rae DE, et al. Comorbidity of mental disorders with alcohol and other drug abuse. Results from the Epidemiological Catchment Area (ECA) Study. JAMA 1990;264:2511-8.

(154) Chen YW, Dilsaver SC. Comorbidity of panic disorder in bipolar illnes: evidence from the Epidemiological Catchment Area Survey. Am J Psychiatry 1995;152:280-2.

(155) Chen YW, Dilsaver SC. Comorbidity of obsessive-compulsive disorder in bipolar and unipolar disorders. Am J Psychiatry 1995;59:57-64.

(156) George EL, Miklowitz DJ, Richards JA, Simoneau TL, Taylor DO. The comorbidity of bipolar disorder and axis II personality disorders: prevalence and clinical correlates. Bipolar Disorders 2003;5:115-22.

(157) Vieta E, Colom F. Personality disorder in bipolar II patients. J Nerv Ment Dis 1999;187:245-8.

(158) Barbato N, Hafner RJ. Comorbidity of bipolar and personality disorder. Austr New Zeal J Psychiatry 1998;32:276-80.

(159) Kay JH, Altshuler LL, Ventura J, Mintz J. Prevalence of axis II comorbidity in bipolar patients with and without alcohol use disorders. Ann Clin Psychiatry 1999;11:187-95.

(160) Pica S, Edwards J, Jackson HJ. Personality disorders in recent-onset bipolar disorder. Compr Psychiatry 1990;31:499-510.

(161) O'Connell RA, Mayo JA, Eng LK, Jones JS, Gabel RH. Social support and long-term lithium outcome. Br J Psychiatry 1985;147:272-5.

(162) Colom F, Vieta E, Martinez-Aran A. Clinical factors associated with treatment noncompliance in euthymic bipolar patients. J Clin Psychiatry 2000;61:549-55.

(163) Schou M. No help from lithium? About patients who might have been but were not helped by prophylactic lithium treatment. Compr Psychiatry 1988;29:83-90.

(164) Dilsaver SC, Henderson-Fuller S, Akiskal HS. Occult mood disorders in 104 consecutively presenting children referred for the treatment of attention-deficit/hyperactivity disorder in a community mental health clinic. J Clin Psychiatry 2003;64(10):1170-6.

(165) Biederman J, Mick E, Faraone SV, Spencer T, Wilens TE, Wozniak J. Pediatric mania: a developmental subtype of bipolar disorder? Biol Psychiatry 2000;48:458-66.

(166) Courtet P, Jollant F, Castelnau D, Astruc B, Buresi C, Malafosse A. Implication of genes of the serotonergic system on vulnerability to suicidal behavior. J Psychiatry Neurosci 2004;29(5):350-9.

(167) Berrios GE. The history of mental symptoms: Descriptive psychopathology since the nineteenth century. 1996.

(168) Álvarez JM. La invención de las enfermedades mentales. Madrid: 1999.

(169) Benavent Rodríguez P. Clasificación de los trastornos bipolares. Interpsiquis . 2001.
Ref Type: Internet Communication

(170) Kendler KS, Gardner CO. Boundaries of major depression: an evaluation of DSM-IV criteria. Am J Psychiatry 1998;155:172-7.

(171) Krishnan KR. Psychiatric and medical comorbidities of bipolar disorder. Psychosom Med 2005;67(1):1-8.

(172) Akiskal HS. Demystifying borderline personality: critique of the concept and unorthodox reflections on its natural kinship with the bipolar spectrum. Acta Psychiatr Scand 2004;110(6):401-7.

(173) Magill CA. The boundary between borderline personality disorder and bipolar disorder: current concepts and challenges. Can J Psychiatry 2004;49(8):551-6.

(174) Casas M. Trastornos duales. In: Vallejo J GC, editor. Trastornos afectivos: Ansiedad y depresión. 2 ed. Barcelona: Masson; 2000.

(175) Carlson GA, Bromet EJ, Driessens C, Mojtabai R, Schwartz JE. Age at onset, childhood psychopathology, and 2-year outcome in psychotic bipolar disorder. Am J Psychiatry 2002 Feb;159(2):307-9.

(176) Costello EJ, Mustillo S, Erkanli A, Keeler G, Angold A. Prevalence and development of psychiatric disorders in childhood and adolescence. Arch Gen Psychiatry 2003 Aug;60(8):837-44.

(177) Hofstra MB, van der EJ, Verhulst FC. Continuity and change of psychopathology from childhood into adulthood: a 14-year follow-up study. J Am Acad Child Adolesc Psychiatry 2000 Jul;39(7):850-8.

(178) Fergusson DM, Horwood LJ, Lynskey MT. Prevalence and comorbidity of DSM-III-R diagnoses in a birth cohort of 15 year olds. J Am Acad Child Adolesc Psychiatry 1993 Nov;32(6):1127-34.

(179) Cohen P, Cohen J, Brook J. An epidemiological study of disorders in late childhood and adolescence--II. Persistence of disorders. J Child Psychol Psychiatry 1993 Sep;34(6):869-77.

(180) Ramos-Quiroga JA, Bosch-Munsó R, Castells-Cervelló X, Nogueira-Morais M, García-Giménez E, Casas-Brugué M. Trastorno por déficit de atención con hiperactividad en adultos: caracterización clínica y terapéutica. Rev Neurol 2006;42(10):600-6.

(181) Perugi G, Akiskal HS. The soft bipolar spectrum redefined: focus on the cyclothymic, anxious-sensitive, impulse-dyscontrol, and binge-eating connection in bipolar II and related conditions. Psychiatr Clin North Am 2002;25(4):713-37.

(182) Savitz J, Solms M, Ramesar R. Neuropsychological dysfunction in bipolar affective disorder: a critical opinion. Bipolar Disorders 2005;7(3):216-35.

(183) Malhi GS, Lagopoulos J, Sachdev PS, Ivanovski B, Shnier R. An emotional Stroop functional MRI study of euthymic bipolar disorder. Bipolar Disorders 2005;7(Suppl. 5):58-69.

(184) Malhi GS, Ivanovski B, Hadzi-Pavlovic D, Mitchell PB, Vieta E, Sachdev P. Neuropsychological deficits and functional impairment in bipolar depression, hypomania and euthymia. Bipolar Disorders 2007;9:114-25.

(185) Cahill CM, Green MJ, Jairam R, Malhi GS. Bipolar disorder in children and adolescents: obstacles to early diagnosis and future directions. Early Intervention in Psychiatry 2007;1:138-49.

(186) Doyle AE, Wilens T, Kwon A, et al. Neuropsychological functioning in youth with bipolar disorder. Biol Psychiatry 2005;58:540-8.

(187) DeKosky ST, Marek K. Looking backward to move forward: early detection of neurodegenerative disorders. Science 302, 830-834. 2003.
Ref Type: Magazine Article

(188) Hasler G, Drevets WC, Gould TD, Gottesman II, Manji HK. Toward constructing an endophenotype strategy for bipolar disorders. Biol Psychiatry 2006;60(2):93-105.

(189) Blumberg HP, Kaufman J, Martin A, Whiteman R, Zhang JH, Gore JC, et al. Amygdala and hippocampal volumes in adolescents and adults with bipolar disorder. Arch Gen Psychiatry 2003;60(12):1201-8.

(190) Young RC, Murphy CF, Heo M, Schulberg HC, Alexopoulos GS. Cognitive impairment in bipolar disorder in old age: Literature review and findings in manic patients. Journal of Affective Disorders 2006;92:125-31.

(191) Comunidad de Madrid. Información y estadísticas. www.madrid.org 2007

(192) Ayuntamiento de Madrid. Anuario estadístico. http://www.munimadrid.es 2007

(193) Instituto Nacional de Estadistica "National Statistics Institute". http://www.ine.es 2007

(194) Organización Mundial de la Salud DdSM. CIE 10, Trastornos mentales y del comportamiento: tablas de conversión entre la CIE 8, la CIE 9 y la CIE 10. Madrid: Meditor; 1993.

(195) Cohen J. A coefficient of agreement for nominal scales. Educational and Psychological Measurement 1960;20:37-46.

(196) Duda R, Hart P, Stork D. Pattern Classification. 2 ed. John Wiley & Sons, Inc; 2001.

(197) Rabiner L, Juang B. An introduction to hidden Markov models. ASSP Magazine 1986;3(Issue 1, Part 1):4-16.

(198) Hougaard P. Analysis of multivariate survival data, statistics for biology and health. New York: Springer-Verlag; 2000.

(199) Fogarty F, Russell JM, Newman SC, Bland RC. Epidemiology of psychiatric disorder in Edmonton: Mania. Acta Psychiatr Scand 1994;376 (Suppl.):16-23.

(200) Feinstein AR. The pretherapeutic classification of comorbidity in chronic disease. J Chronic Dis 1970;23:455-68.

(201) Ghaemi N, Sachs GS, Goodwin FK. What is to be done? Controversies in the diagnosis and treatment of manic-depressive illness. World J Biol Psychiatry 2000;1(2):65-74.

(202) Becker JB, Hu M. Sex differences in drug abuse. Front Neuroendocrinol 2007.

(203) Brady KT, Randall CL. Gender differences in substance use disorders. Psychiatr Clin North Am 1999;22(2):241-52.

(204) Rubin A, Stout RL, Longabaugh R. Gender differences in relapse situations. Addiction 1996;91(Suppl):S111-S120.

(205) Werry JS, McClellan JM, Chard L. Childhood and adolescent schizophrenic, bipolar, and schizoaffective disorders: a clinical and outcome study. J Am Acad Child Adolesc Psychiatry 1991.

(206) Joyce PR. Age of onset in bipolar affective disorder and misdiagnosis as schizophrenia. Psychol Med 1984;14:145-9.

(207) Patten SB. Markov models of major depression for linking psychiatric epidemiology to clinical practice. Clin Pract Epidemol Ment Health 2005;1-2.

(208) Haro JM, Palacín C, Vilagut G, Bernal M, Luque I, Codony M, et al. Prevalence of mental disorders and associated factors: results from the ESEMeD-Spain study. Med Clin (Barc) 2006;126(12):445-51.

(209) Bland RC, Newman SC, Om H. Period prevalence of psychiatric disorders in Edmonton. Acta Psychiatr Scand 1988;77:33-42.

(210) Sandanger I, Nygard JF, Sorensen T. The concept of psychiatric illness - a core problem in psychiatric epidemiology. Norwegian Journal of Epidemiology 2002;12(3):181-7.

(211) Das PP, Grover S. Diagnostic stability: clinical v. research. Br J Psychiatry 2007;191(Correspondence):265-6.

(212) Ministerio de Salud y Consumo. http://www.msc.es/estadEstudios/estadisticas/docs/FOLLETO-BAJA PDF 2007

(213) Möller HJ. Bipolar disorder and schizophrenia: distinct illnesses or a continuum? J Clin Psychiatry 2003;64(Suppl 6):23-7.

(214) Llopis B. La psicosis única: escritos escogidos. 1 ed. Madrid: 2003.

(215) Peralta V, Cuesta MJ. The nosology of psychotic disorders: a comparison among competing classification systems. Schizophr Bull 2003;29(3):413-25.

(216) van Os J, Verdoux H, Maurice-Tison S, Gay B, Liraud F, Salamon R, et al. Self-reported psychosis-like symptoms and the continuum of psychosis. Soc Psychiatry Psychiatr Epidemiol 1999;34(9):459-63.