UNIVERSIDAD AUTONOMA DE MADRID

FACULTAD DE MEDICINA

Departamento de Psiquiatría

TESIS DOCTORAL

TRAYECTORIA ASISTENCIAL EN EL TRASTORNO BIPOLAR

FANNY CEGLA SCHVARTZMAN

Bajo la dirección de:
Enrique Baca-García

Madrid, 2022

AUTONOMA UNIVERSITY OF MADRID

FACULTY OF MEDICINE

Department of Psychiatry

DOCTORAL THESIS

PATHWAYS OF CARE IN BIPOLAR DISORDER

FANNY CEGLA SCHVARTZMAN

Directed by:
Enrique Baca-García

Madrid, 2022

DR. ENRIQUE BACA GARCÍA, PROFESOR TITULAR DEL DEPARTAMENTO DE PSIQUIATRÍA DE LA UNIVERSIDAD AUTÓNOMA DE MADRID Y JEFE DEL DEPARTAMENTO DE PSIQUIATRÍA DE LA FUNDACIÓN JIMÉNEZ DIAZ

CERTIFICA QUE:

El trabajo de investigación presentado por **Dña. Fanny Cegla-Schvartzman**, titulado "***Trayectoria asistencial en el Trastorno Bipolar***" realizado bajo mi dirección, reúne los requisitos científicos, metodológicos y de originalidad suficientes para ser defendido como Tesis Doctoral ante el Tribunal que legalmente proceda.

Y para que surta los efectos oportunos, se firma la presente en Madrid a primero de marzo de 2022.

Dr. Enrique Baca García

A todos aquellos que padecen algún trastorno mental.

AGRADECIMIENTOS

A mi director de tesis, al Prof. Dr. Enrique Baca, por el tiempo dedicado, por sus consejos, sus conocimientos, por todo su apoyo, su infinita paciencia y por todas las oportunidades brindadas. Gracias por crear un ambiente que permite el crecimiento profesional y personal.

Al Dr. Ovejero, tutor y referente. Por todo lo enseñado y por estimular la excelencia en la práctica clínica.

Al Dr. Martinez-Ales por su inestimable ayuda y al Dr. Lopez Castromán por sus valiosas aportaciones.

A mis padres, Sofia y Enrique, mis hermanos Daniel y Nelson, por la ayuda y el apoyo incondicional.

A mis amigos, mi familia elegida, por estar siempre presentes.

A mis compañeros del Departamento de Psiquiatría de la Fundación Jiménez Díaz y en especial a Marta y Nora, porque en equipo todo es más llevadero.

Gracias a todos los que de una u otra manera han ayudado a la realización de esta tesis.

INDEX

ABBREVIATIONS AND ACRONYMS

1. ABBREVIATIONS AND ACRONYMS

- **BD:** Bipolar disorder
- **CT-HMM**: Continuous-Time hidden Markov model
- **DSM**: Diagnostic and Statistical Manual of Mental Disorders
- **HMM:** Hidden Markov model
- **IDC:** International Classification of Diseases
- **PC**: Prospective consistency
- **RC:** Restroscpective consistency

RESUMEN

2. RESUMEN

2.1 Introducción

Los trastornos mentales graves, como la esquizofrenia y el trastorno bipolar, son afecciones crónicas, caracterizadas por síntomas recurrentes y consideradas enfermedades de por vida, que pueden tener un gran impacto en la capacidad de funcionamiento de un individuo y que, una vez establecido el diagnóstico, deberían ser estables a lo largo del tiempo; sin embargo, esto no siempre es así en la práctica clínica.

En otras ramas de la medicina, los diagnósticos suelen apoyarse en la identificación de los procesos biológicos subyacentes, mientras que, en psiquiatría, los diagnósticos, se basan principalmente en una evaluación transversal y longitudinal de las presentaciones clínicas.

2.2 Metodología

La estabilidad diagnóstica es el grado en que un diagnóstico se mantiene sin cambios durante el seguimiento. Se revisó la literatura en relación a la estabilidad diagnóstica en el trastorno bipolar, a través de una búsqueda bibliográfica exhaustiva, incluyendo todos los estudios publicados desde 1980 hasta 2016, para evaluar la estabilidad diagnóstica en el trastorno bipolar.

Posteriormente, analizamos el Registro Acumulativo de Casos de la Comunidad de Madrid, que incluye datos sociodemográficos y códigos diagnósticos según la Clasificación Internacional de Enfermedades (CIE) de todas las consultas psiquiátricas ambulatorias

realizadas entre enero de 1980 y diciembre de 2009 en los Centros de Salud Mental de la Comunidad de Madrid. Con dos objetivos, comprobar hasta que punto el proceso diagnostico de la esquizofrenia esta precedido de otros diagnóstios y establecer los cambios en los diagnoticos que recibe un paciente con trastorno bipolar durante su evolución.

Se investigó una gran muestra clínica de 26.163 pacientes con diagnóstico de esquizofrenia en al menos una visita ambulatoria. Se aplicó el Modelo oculto de Markov para describir la probabilidad de transición de otros diagnósticos a la esquizofrenia teniendo en cuenta la proximidad temporal.

Los cambios de diagnótico en el trastorno bipolar se investigaron en una muestra de 14.557 que fueron diagnosticados de trastorno bipolar en al menos una evaluación, tuvieron al menos 10 visitas y un año de seguimiento. Se utilizaron dos índices: consistencia temporal (mantenimiento del diagnóstico en el tiempo) y constancia diagnóstica (presencia del diagnóstico de BD en al menos el 75% de las visitas). El coeficiente Kappa midió la concordancia entre los diagnósticos de la primera y la última evaluación (consistencia prospectiva y retrospectiva).

2.3 Resultados

Los diagnósticos más frecuentes antes del diagnóstico de la esquizofrenia eran los trastornos de ansiedad y del estado de ánimo. Sin embargo, las transiciones directas a la esquizofrenia solían provenir de trastornos del espectro psicótico.

El diagnóstico inicial de esquizofrenia no solía cambiar en dos de cada tres pacientes si se confirmaba unos meses después de su aparición. Cuando no se confirmaba, los

diagnósticos alternativos más frecuentes eran los trastornos de personalidad, afectivos o psicóticos no esquizofrénicos. Los diagnósticos erróneos o la comorbilidad con trastornos afectivos, de ansiedad y de personalidad son frecuentes antes y después del diagnóstico de esquizofrenia. Nuestros hallazgos apoyan parcialmente una visión dimensional de la esquizofrenia y enfatizan la necesidad de una evaluación longitudinal.

Respecto al diagnóstio de trastorno bipolar, se encontró una mayor estabilidad si el diagnóstico se hizo después de la hospitalización, si la edad de inicio era >25 años, si el diagnóstico de trastorno bipolar se hizo a la edad < 24 años, si >65% de las visitas fueron realizadas por el mismo psiquiatra y si el paciente había sido evaluado por el mismo psiquiatra en la primera y última evaluación.

2.4 Conclusiones

Este trabajo refleja las prácticas clínicas comunes en pacientes con esquizofrenia y trastorno bipolar en un entorno ambulatorio durante un período de observación de 30 años. Se encontraron frecuentes cambios de diagnóstico en relación con el trastorno bipolar y varios factores parecen tener un impacto en la estabilidad del diagnóstico.

ABSTRACT

3. ABSTRACT

3.1 Background

Severe mental disorders, such as schizophrenia and bipolar disorder, are chronic affections, characterized by recurrent symptoms and considered life-long illness, that can have a major impact on an individual's ability to function and once the diagnosis is established, should be stable over time; however, this is not always the case in clinical practice.

In other branches of medicine, diagnoses are often supported by the identification of the underlying biological processes, whereas in psychiatry, the diagnoses, rely principally on a transversal evaluation of clinical presentations.

3.2 Methodology

Diagnostic stability is the degree to which a diagnosis remains unchanged during follow-up. In this study, we first review the literature concerning diagnostic stability in Bipolar Disorder, through a comprehensive literature search, including all studies published from 1980 to 2016, to evaluate the diagnostic stability of Bipolar Disorder. Secundarily, using the Cumulative Register of Cases of the Community of Madrid, an electronic healthcare record that includes socio-demographic data and International Classification of Diseases (ICD) diagnostic codes for all outpatient psychiatric visits held between January, 1980 - December, 2009 at the Madrid's Community Mental Healthcare Centers, we selected the patients. We established two aims to asses the (I) shift changes to Schizophrenia and (II)

shift changes to Bipolar Disorder.

A large clinical sample of 26,163 patients with a diagnosis of schizophrenia in at least one outpatient visit was investigated. We applied a Continuous Time Hidden Markov Model to describe the probability of transition from other diagnoses to schizophrenia considering time proximity.

A total of 14,557 patients were diagnosed with Bipolar Disorder for at least one evaluation and had at least 10 visits and one year of follow-up. Two indices were measured in order to assess Shift changes to Bipolar Disorder: temporal consistency (maintenance of the diagnosis over time) and diagnostic constancy (presence of the Bipolar Disorder diagnosis in at least 75% of visits). Kappa coefficient measured the agreement between diagnoses in the first and the last evaluation (prospective and retrospective consistency).

3.3 Results

Although the most frequent diagnoses before schizophrenia were anxiety and mood disorders, direct transitions to schizophrenia usually came from psychotic-spectrum disorders. The initial diagnosis of schizophrenia was not likely to change for two of every three patients if it was confirmed some months after its onset. When not confirmed, the most frequent alternative diagnoses were personality, afective or non-schizophrenia psychotic disorders. Misdiagnosis or comorbidity with afective, anxiety and personality disorders are frequent before and after the diagnosis of schizophrenia. Our findings give partial support to a dimensional view of schizophrenia and emphasize the need for longitudinal assessment.

Greater stability was found if the diagnosis of Bipolar Disorder when it was made after hospitalization, if the age of onset was >25 years, if the diagnosis of BD was made at age < 24 years, if >65% of the visits were held by the same psychiatrist and if the patient had been assessed by the same psychiatrist in the first and last assessments.

3.4 Conclusions

This work reflects common diagnostic clinical practices in patients with schizophrenia and Bipolar Disorder in outpatient settings over a 30-year period of observation. Frequent diagnostic shifts were found in relation to Bipolar Disorder and several factors appear to have an impact on the diagnostic stability.

INTRODUCCIÓN

4. INTRODUCCIÓN

En otras ramas de la medicina, los diagnósticos suelen apoyarse en la identificación de los procesos biológicos subyacentes, mientras que, en psiquiatría, los diagnósticos se basan principalmente en los síndromes clínicos (1), por lo que los diagnósticos de los trastornos psiquiátricos, se basan principalmente en una evaluación transversal de las presentaciones clínicas.

La esquizofrenia es un trastorno mental crónico, caracterizado por síntomas recurrentes de psicosis, que suele conllevar impedimentos sociales y laborales (2).

El trastorno bipolar es una enfermedad mental crónica y grave, caracterizada por episodios recurrentes de estado de ánimo deprimido, elevado y mixto. Y dado que se considera una enfermedad que dura toda la vida, el diagnóstico de trastorno bipolar, una vez establecido, debería ser estable en el tiempo (3). Sin embargo, esto no siempre es así en la práctica clínica (4).

4.1 Relevancia y epidemiología

La esquizofrenia y el trastorno esquizoafectivo tienen una prevalencia a lo largo de la vida de alrededor del 1% y se encuentran entre las principales causas de discapacidad (5–8). Debido a su temprana aparición y a deterioro progresivo, la esquizofrenia es causa de una inmensa carga económica. Una gran mayoría de los pacientes con esquizofrenia están desempleados, y las dificultades en el funcionamiento social, como profesional y a nivel residencial, siguen siendo graves, incluso durante los períodos de remisión de psicosis

activa (9–13).

Según el estudio Global Burden of Disease del 2013, el número de casos de trastorno bipolar aumentó en un 50% en los últimos 20 años, alcanzando los 50 millones de casos y situándose como la 16ª causa de años vividos con discapacidad en el 2013 (14). El trastorno bipolar puede tener un gran impacto en la capacidad de un individuo para funcionar (14). Además, el trastorno bipolar está asociado a un mayor riesgo de mortalidad por todas las causas y, en particular, de mortalidad por suicidio. (15,16). Por todo ello, investigaciones previas han puesto de manifiesto el elevado coste y la carga que supone el trastorno bipolar (17,18).

4.2 Estabilidad en psiquiatría

Una limitación común en los estudios de epidemiología psiquiátrica es que los diagnósticos psiquiátricos se basan en evaluaciones clínicas y no en mediciones biológicas (19). En ausencia de una sintomatología biológica objetiva del trastorno, la estabilidad en el tiempo representa la mejor prueba para validar el diagnóstico y, en gran medida puede utilizarse para predecir el curso de un trastorno (20). Se supone que la estabilidad diagnóstica a lo largo del tiempo es característica de los trastornos psiquiátricos con tendencia a la cronicidad y a las recaídas en el tiempo, como es el caso de la esquizofrenia y del trastorno bipolar. Sin embargo, la estabilidad varía notablemente entre los trastornos psiquiátricos crónicos. Por ejemplo, la esquizofrenia ha resultado ser uno de los diagnósticos más estables (21–24). Por el contrario, se estima que la tasa de pacientes bipolares que reciben diagnósticos inexactos en los centros de salud mental es de entre el 20% y el 60% (25–27).

La ausencia de estabilidad en un diagnóstico puede tener graves consecuencias (28). El infradiagnóstico conduce a un tratamiento tardío o ineficaz (29) y el sobrediagnóstico también puede tener consecuencias personales y sociales adversas, como la exposición innecesaria a los riesgos de la medicación, la pérdida de oportunidades para el tratamiento de otras afecciones, así coomo en la participación laboral (30).

4.2.1 Estabilidad en la esquizofrenia

La esquizofrenia es una enfermedad muy heterogénea, bautizada con este término por Bleuler en 1911 (31,32). Aunque han habido avances significativos en la comprensión de la enfermedad durante los últimos 130 años (desde la clasificación original de Kraepelin en 1887), los fundamentos de su etiología siguen siendo desconocidos y a día de hoy no existen biomarcadores que puedan utilizarse regularmente en la práctica clínica para el diagnóstico de la esquizofrenia (33,34). Por lo tanto, la validación longitudinal proporciona una de las pruebas más directas de la validez del diagnóstico (35).

Varios estudios han indagado la estabilidad diagnóstica a largo plazo de la esquizofrenia. Los metaanalísis muestra una alta estabilidad diagnóstica prospectiva en el espectro de la esquizofrenia, pero la mayor parte de la literatura está centrada en muestras pequeñas de primeros episodios psicóticos o en individuos con alto riesgo de esquizofrenia (36–39). Existe un conocimiento limitado de los diferentes diagnósticos que reciben los pacientes en torno al inicio de la esquizofrenia. En base a dos estudios epidemiológicos, An der Heiden et al. (2000) informaron de que con una media de 5 años, en el 75% de los pacientes el primer episodio psicótico en la esquizofrenia fue precedido de síntomas

prodrómicos, normalmente síntomas negativos y depresivos (40,41). Otros autores también han señalado la presencia de alteraciones a nivel conductual y afectivo, previo al diagnóstico de esquizofrenia (42,43). Asimismo, los estudos prospectivos de pacientes con síntomas prodomico muestran que fueron frecuentemente diagnosticadas de depresión, ansiedad o trastornos por consumo de sustancias, en particular de cannabis, antes de hacer la transición a la psicosis (44–46).

Por otro lado, el paradigma científico actual está desafiando la perspectiva categórica de hace tiempo (47), a favor de una conceptualización más dimensional de las psicosis (48). Según el modelo dimensional, un fenotipo extendido de esquizofrenia en la población general (vulnerabilidad) subyace al fenotipo clínico menos común de la esquizofrenia (49). Los elevados niveles de gravedad en las diferentes dimensiones de los síntomas conducirían a la evaluación clínica, identificando los síntomas correlacionados en otras dimensiones y, finalmente, al diagnóstico de esquizofrenia. Cabría esperar diagnósticos previos correspondientes a las diferentes dimensiones que se han propuesto (negativa, afectiva, psicótica y cognitiva) al estudiar una muestra amplia de pacientes con esquizofrenia.

4.2.2 Estabilidad en el trastorno bipolar

Hay una escasez de estudios centrados en la estabilidad diagnóstica, a pesar del interés que supone para la planificación en salud mental; aunque su número ha aumentado en los últimos años. La mayoría de ellos centrados en el primer episodio psicótico, y particularmente en la estabilidad diagnóstica de la esquizofrenia (50). Sin embargo, hay

menos estudios centrados en la estabilidad diagnóstica del trastorno bipolar a lo largo del tiempo. La validez del diagnóstico en los trastornos psiquiátricos requiere estabilidad a lo largo del tiempo, y debido a las potenciales implicaciones del tratamiento, los cambios en el diagnóstico son un tema fundamental a tener en cuenta (35,51). En el caso del trastorno bipolar, la literatura es inconsistente.

La mayoría de los estudios recientes sugieren niveles de estabilidad diagnóstica de moderados a altos para el trastorno bipolar (3,52–60). Sin embargo, estos estudios están limitados por dificultades técnicas. Por ejemplo, muchos estudios sólo han utilizado unos pocos puntos de evaluación, dos o tres a lo sumo, durante períodos de seguimiento limitados (61–64), lo que genera dudas sobre la generalización a periodos de tiempo más amplios y sugiere la necesidad de realizar estudios que incluyan más puntos de evaluación durante periodos de seguimiento más largos.

Un punto a tener en cuenta es que el diagnóstico inicial del trastorno bipolar suele ser problemático. Los retrasos de 8 a 10 años en el diagnóstico son comunes, y los estudios con evaluaciones longitudinales repetidas han puesto en duda la estabilidad de este diagnóstico en la práctica real (64–67).

4.3 Antecedentes

A través de una exhaustiva búsqueda bibliográfica (68), que incluyó estudios publicados entre 1980 y 2016, que evaluaban la estabilidad diagnóstica en el trastorno bipolar, se seleccionaron 37 estudios; seis de ellos centrados principalmente en el trstorno bipolar (4,51,54,61,65,69) (ver apendice), dieciocho centrado en trastornos psicóticos

(3,20,24,36,37,39,53,55–57,62–64,66,70–73) (see apenix), diez en depresión (74–83) (**¡Error! No se encuentra el origen de la referencia.**) y tres artículos centrados en la e stabilidad diagnóstica de los trastornos psiquiátricos en general (67,84,85) (ver apendice).

La mayoría de los estudios son de diseño prospectivo (n=25; 67,5%), y 12 (32,4%) son retrospectivos (**Figure** 1). Los principales criterios diagnósticos utilizados fueron el Manual Diagnóstico y Estadístico de los Trastornos Mentales (DSM) (n= 17; 46%), y la Clasificación Internacional de Enfermedades (CIE) (n= 15; 40,5%); 3 estudios utilizaron ambos (8,1%). Un estudio utilizó los Criterios Diagnósticos de Investigación (2,7%), y otro los criterios de Feighner (2,7%). De los 37 estudios incluidos, se utilizaron 40 instrumentos de evaluación diferentes para clasificar a los pacientes y los síntomas. Los criterios más utilizados para valorar la estabilidad diagnóstica fueron las consistencias prospectiva y retrospectiva, que fueron utilizadas por 27 estudios (73%). Las herramientas de medición menos comunes fueron la proporción de cambio diagnóstico, utilizada por 12 estudios (32%), y el coeficiente kappa de Cohen, utilizado en siete (19%).

Al examinar los 37 estudios, encontramos una consistencia media prospectiva del 77,4% y una consistencia retrospectiva del 67,6%. Sin embargo, debido a la variabilidad en el tamaño de las muestras y la importancia relativa de cada grupo, calculamos la media ponderada de las consistencias prospectiva y retrospectiva en cada grupo diagnóstico, controlada por el tamaño de la muestra . Por otra parte, los estudios centrados en la depresión estudiaron la tendencia de cambios diagnósticos hacia el trastorno bipoar, y por lo tanto, las consistencias prospectivas y retrospectivas no pueden ser evaluadas. Las consistencias prospectivas y retrospectivas se resumen en la *Figure 1*. Una mayoría

considerable de los estudios se realizaron en Europa o en Norteamérica (n= 25; 67,5%), mientras que el 21,6% (n=8) se realizaron en Asia y el 10,8% (n=4) en África, Oceanía y Sudamérica. El tamaño de las muestras varía entre 48 y 69792 pacientes. Cuatro estudios (10,8%) tienen menos de 100 pacientes, 25 estudios (67,5%) tienen entre 100 y 1000 pacientes, 7 estudios (20%) incluyeron entre 1000 y 10000 sujetos, y sólo 1 (2,7%) tenía más de 10000 pacientes.

INTRODUCTION

5. INTRODUCTION

In other branches of medicine, diagnoses are often supported by the identification of the underlying biological processes, whereas in psychiatry, diagnoses are based primarily on clinical syndromes (1), therefore, the diagnoses of psychiatric disorders, rely principally on a transversal evaluation of clinical presentations.

Schizophrenia is a chronic mental disorder, characterized by recurrent symptoms of psychosis which usually leads to social and occupational impairements (2).

BD is a chronic and severe mental disorder, characterized by recurrent episodes of depressed, elevated, and mixed mood. Moreover, since it is considered a life-long illness, the diagnosis of BD, once established, should be stable over time (3). However, this is not always the case in clinical practice (4).

5.1 Relevace and epidiemiology

Schizophrenia and schizoaffective disorder have a lifetime prevalence of about 1% and are among the leading causes of disability (5–8). Due to its early onset and its deteriorating course, schizophrenia causes an immense economic burden. A large majority of patients with schizophrenia are unemployed, and impairments in functioning across social, vocational and residential domains remain severe even during periods of remission from active psychosis (9–13).

According to the Global Burden of Disease 2013 study, the number of cases of BD increased by 50% in the last 20 years, attaining 50 million cases and being ranked as the

16th leading cause of years lived with disability in 2013 (14). BD can have a major impact on an individual's ability to function (14). In addition, BD is associated to an increased risk of all-cause mortality and, in particular, of suicide mortality (15,16,86). As a result, previous research has highlighted the high cost and burden driven by BD (17,18).

5.2 Stability in psychiatry

A common limitation of psychiatric epidemiology studies is that psychiatric diagnoses are based on clinical assessments rather than biological measurements (19). In the absence of objective biological symptomatology of the disorder, the stability over time represents the best proof to validate the diagnosis and to a great extent it can be used to predict the course of a disorder (20). Diagnostic stability over time is presumed to be characteristic of psychiatric conditions with a tendency to chronicity and relapses over time, such as schizophrenia and BD. However, stability varies markedly across chronic psychiatric disorders. For instance, schizophrenia has been found to be one of the most stable diagnosis (21–24). Conversely, the rate at which bipolar patients receive inaccurate diagnoses in mental health facilities is estimated to be as high as 20%–60% (25–27).

The absence of stability in a diagnosis may have serious implications (28). Underdiagnosis leads to delayed or ineffective treatment (29) and overdiagnosis may also have adverse personal and social consequences, including unnecessary exposure to the risks of medication, missed opportunities for treatment of other conditions, and effects on work participation (30).

5.2.1 Stability in Schizophrenia

Schizophrenia is a very heterogeneous illness, as originally described by Bleuler in 1911 and later confirmed in several studies (31,32). Although significant advances have been made in the understanding of the illness during the last 130 years (since Kraepelin's original classification in 1887) the underpinnings of its etiology remain unknown. There are no biomarkers that could regularly used in clinical practice for the diagnosis of schizophrenia (33,34). Thus, longitudinal validation provides one of the most direct evidences of diagnostic validity (35).

Several studies have investigated the long-term diagnostic stability of schizophrenia. Meta-analytical evidence shows high prospective diagnostic stability in schizophrenia spectrum, but most of the literature was focused on small samples of first-episode psychosis or individuals at high-risk of schizophrenia (36–39). There is also limited knowledge of the different diagnoses received by the patients around the onset of schizophrenia. Based on two epidemiological studies, An der Heiden et al. (2000) reported that the first psychotic episode in schizophrenia was preceded in 75% of patients by an average of 5 years of prodromic symptoms, usually negative and depressive symptoms (40,41). Other authors have also noted the presence of prior behavioral and affective abnormalities in patients with schizophrenia (42,43). Likewise, prodromal samples studied prospectively were frequently diagnosed with depression, anxiety or substance use disorders, particularly cannabis, before making a transition to psychosis (44–46).

On the other hand, the current scientific paradigm is challenging the long-standing categorical perspective (47), in favor of a more dimensional conceptualization of psychoses (48). According to the dimensional model, an extended phenotype of schizophrenia in the general population (vulnerability) would underlie the less common clinical phenotype of schizophrenia (49). High levels of severity in different symptom dimensions would lead to clinical assessment, identification of correlated symptoms in other dimensions and finally, the diagnosis of schizophrenia. We might expect prior diagnoses corresponding to the different dimensions that have been proposed (negative, affective, psychotic, and cognitive) when studying a large sample of patients with schizophrenia.

5.2.2 Stability in BD

There is a paucity of studies focused on diagnostic stability, despite the interest for mental health planning; though its number has increased in recent years. The majority of them are focused on the first episode of psychosis, and particularly the subsequent consistency of schizophrenia diagnoses (50). However, there are fewer studies focused on the BD diagnostic consistency over time. Diagnostic validity in psychiatric disorders requires stability over time, and because of the potential treatment implications, changes in diagnosis are a major issue to be considered (35,51). In the case of BD, the literature is inconsistent.

Most recent studies suggest moderate to high levels of diagnostic stability for BD (3,52–60). However, these studies are limited by technical difficulties. For instance, many

studies have only used a few evaluation points, two or three at most, over limited follow-up periods (61–64), raising concerns about the generalization to wider time periods and suggesting the need of studies including more evaluation points over longer follow-up periods.

A point also to be noted, is that the initial diagnosis of BD, is often problematic. Diagnostic delays of 8 to 10 years are common, and studies consisting of repeated longitudinal evaluations have called into question the stability of this diagnosis in actual practice (64–67).

5.3 Previous evidence

Through a comprehensive literature search (60) including all studies published from 1980 to 2016, evaluating the diagnostic stability of BD, we selected 37 studies; six of them are mainly focused on BD (4,51,54,61,65,69) (see apendix), eighteen on psychotic disorders (3,20,24,36,37,39,53,55–57,62–64,66,70–73) (see apendix), ten on depression (74–83) (see apendix), and three articles focused on diagnostic stability in psychiatric disorders in general (67,84,85) (see apendix).

Most of the studies are prospective in design (n=25; 67.5%), and twelve (32.4%) are retrospective (**Figure** 1). The main diagnostic criteria used were the Diagnostic and Statistical Manual of Mental Disorders (DSM) (n= 17; 46%), and the International Classification of Diseases (ICD) (n= 15; 40.5%); three studies used both of these (8.1%). One study used the Research Diagnostic Criteria (RDC) (2.7%), and one used the Feighner criteria (2.7%). Of the 37 included studies, 40 different assessment instruments were used

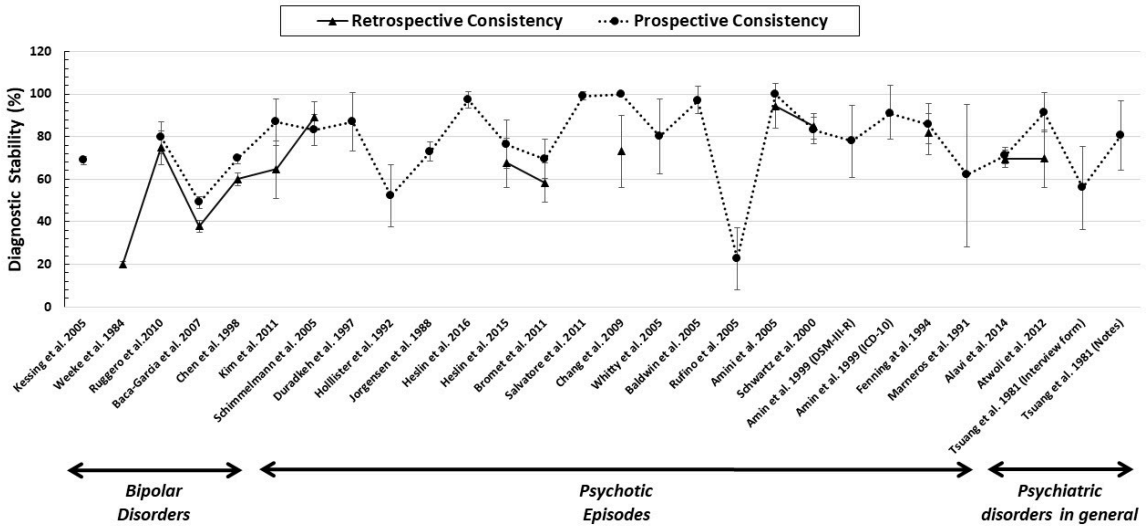
to classify patients and symptoms. The most widely used criteria for diagnostic stability were prospective and retrospective consistencies, which were used by 27 studies (73%). Less common measurement tools included the proportion of diagnostic change, used by 12 studies (32%), and Cohen's kappa inter-rater agreement, used in seven (19%).

Examining all 37 studies, we found a mean prospective consistency of 77.4% and a retrospective consistency of 67.6%. However, due to the variability in samples sizes and relative importance of each group, we calculated the weighted mean for prospective and retrospective consistencies in each diagnostic group, controlled by sample size (table 1). On the other hand, studies focusing on depression revealed a trend of diagnostic shifts to BD, and therefore, prospective and retrospective consistencies cannot be assessed. Prospective and retrospective consistencies are summarized in *Figure 1*. A sizeable majority of studies were performed in Europe or in North America (n= 25; 67.5%), whereas 21.6% (n=8) were performed in Asia and 10.8% (n=4) in Africa, Oceania and South America. The sample sizes vary from 48 to 69792 patients. Four studies (10.8%) have fewer than 100 patients, 25 studies (67.5%) have between 100 and 1000 patients, 7 studies (20%) included between 1000 and 10000 subjects, and only 1 (2.7%) had more than 10000 patients. We focused our description of the results of these articles on the most common criteria for diagnostic stability: prospective and retrospective consistency.

Table 1. Weighted means controlled by sample size for prospective and retrospective consistency in each diagnostic group

STUDIES FOCUSED ON	PROSPECTIVE CONSISTENCY X (95% C.I.)	RETROSPECTIVE CONSISTENCY X (95% C.I.)
Bipolar Disorders	65.7 (64.5 - 66.9)	32.9 (31.6 - 34.2)
Psychotic Episodes	80.4 (79.4 - 81.4)	75.7 (74.1 -77.3)
Psychiatric Disorders in General	70.9 (68.6 - 73.2)	69.4 (65.8 - 73.1)

Figure 1. Prospective and Retrospective Consistencies among the reviewed studies



JUSTIFICATION

6. JUSTIFICATION

Severe mental conditions can have a major impact on an individual's ability to function. Symptoms often overlap and a correct diagnosis followed by an appropriate intervention is of utmost importance to favor its evolution.

Schizophrenia is among the leading causes of disability and due to its early onset and its deteriorating course, can cause an immense economic burden.

Bipolar disorder is a chronic and severe mental disorder that is considered a life-long illness, and its diagnosis, once established, should be stable over time. However, this is not always the case in clinical practice.

The diagnoses of psychiatric disorders, including BD and schizophrenia, rely principally on a transversal evaluation of clinical presentations. Diagnostic stability in BD is an under-studied area, even though it represents the best means of validating the diagnosis.

The objective of this work is to establish the diagnostic evolution of patients with schizophrenia before and after this diagnosis is made for the first time and to establish the diagnostic stability of BD, in real conditions in a daily practice over a 30-year period of observation, including outpatient consultations.

HYPOTHESIS AND OBJECTIVES

7. HYPOTHESIS AND OBJETIVES

7.1 Hipothesis

It is postulated that there are difficulties in the correct diagnosis of certain psychiatric conditions, including non-affective (schizophrenia) and affective psychosis (bipolar disorder). Specifically, in BD, where both, over- and under-diagnosis errors occur.

7.1.1 Hipothesis I: Shift changes in schizophrenia

We hypothesized that the early onset of symptoms in different dimensions would be reflected in correspondingly different diagnostic pathways leading to the diagnosis of schizophrenia.

7.1.2 Hipothesis II: Shift changes in BD

We hypothesized that bipolar disorder is an unnestable diagnosis, and there is a clinical profile of patients with unstable/stable diagnosis.

7.2 Main objective

The main objective of this study is to carry out an ecological assessment, in psychiatric outpatient clinics in the Autonomous Community of Madrid, to examine the diagnostic evolution of patients with severe mental conditions in the early stages of the disease and long term stablity. Specifically we will focus on two major mental disoreders

1. To examine the diagnostic evolution of patients with schizophrenia before and after this diagnosis is made for the first time in public mental health facilities, we investigated a large clinical sample to identify which diagnoses preceded and followed that of schizophrenia.
2. The validity of the diagnosis of bipolar disorder is determined by its diagnostic stability during multiple assessments over a long period of time. The degree to which a patient is consistently classified as having bipolar disorder during follow-up is an important marker for the validity of the diagnosis itself.

7.3 Secondary objectives

A secondary objective of this work is to determine the clinical profile of patients with Severe mental conditions and unnestable/stable diagnosis.

METHODS

8. METHODS

8.1 Narrative review

In order to define more precisely the hipotheis and the methodology we conducted a narrative review (included in introduction) following the PRISMA checklist.

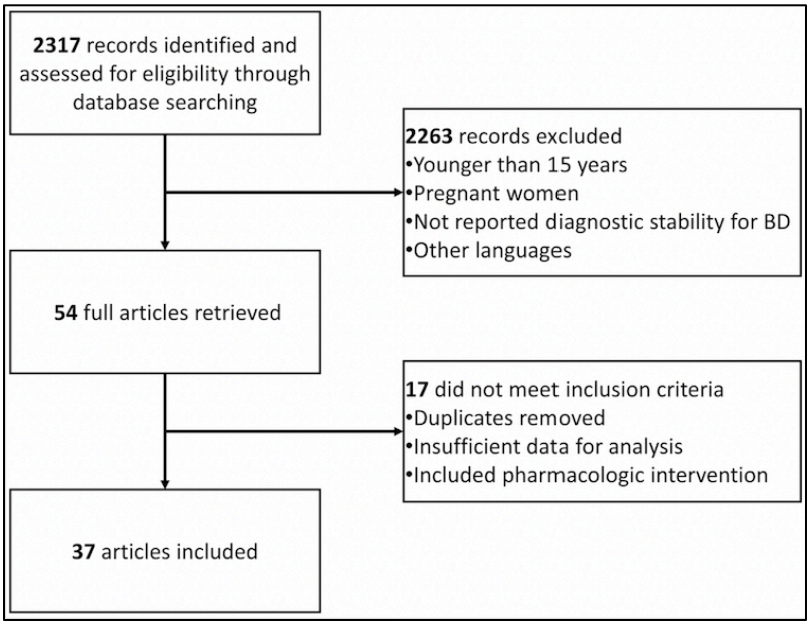
A comprehensive literature search was conducted in the following electronic databases: MEDLINE, EMBASE (the Excerpta Medica database) and PubMed, from 1980 to 2016. Key words used for the search included: BD or manic-depression, mania, bipolar spectrum AND diagnostic stability, diagnostic change, diagnostic consistency, diagnostic shift, diagnostic concordance, diagnostic conversion, diagnostic progression. Keywords were limited to the title and abstracts. The abstracts of the retrieved articles were then checked by applying the eligibility criteria. The PRISMA guidelines were followed to report findings (87).

Articles were included if they were published in English or Spanish language, including adolescents (over 15 years) and adult cases, and reported the assessment of diagnostic stability in BD. Our definition of BD was not overly-inclusive; we required that studies evaluated either BDs, psychotic episodes and major depression (which included emergency visits, inpatients, and outpatients). Samples with pediatric patients and / or pregnant women were not included. One study with patients younger than 15 years was included due to their small number (n=20) in the total sample (n= 69.792). No studies have been included in which the diagnoses were not performed by specialized mental

health personnel. Cross sectional, retospective and prospective studies were included. Citations within identified papers were included as additional sources. We omitted studies that included fewer than 10 patients, uncompleted studies, as well as conference abstracts, dissertations or data that were not published in peer-reviewed journals. The initial electronic search identified 2317 documents. After evaluating the abstracts of all these documents, we selected 37 articles that met the criteria (

Figure 2). During selection of articles we were aware of unavoidable biases such as selection bias (differences between baseline characteristics of the groups) and detection bias (differences in how outcomes are measured). Thirteen studies (35%) included adolescent patients (>15 years). The extraction of the pertinent data was standardized using a predefined form. Variables related to diagnostic stability were searched according to the following definitions.

Figure 2. PRISMA Flow chart



8.1.1 Operative definitions:

Diagnostic stability is the measure of the degree to which a diagnosis remains the same during follow up and constitutes a longitudinal validation of the original baseline diagnosis. It is based on the agreement of diagnoses over time and is irrespective of cross-sectional diagnosis at a single point of follow-up (55). In a pioneering 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 (35).

Prospective consistency (conceptually similar to positive predictive value) is the proportion of subjects in a diagnostic category at the first evaluation who received the same diagnosis at their last evaluation. Retrospective consistency (conceptually similar to sensitivity) is the proportion of subjects in a diagnostic category at the last evaluation who were in that same category at baseline (52,65,84). Cohen's kappa inter-rater agreement measures the diagnostic agreement corrected by chance. Diagnostic agreement was interpreted according to Landis and Koch(88) ($k < 0$ absence of agreement, 0.10–0.20 slight, 0.21–0.40 fair, 0.41–0.60 moderate, 0.61–0.80 good, and 0.81–1 excellent).

We used two indices of diagnostic stability:

1. Temporal consistency: the presence or absence of a particular disorder at first and last evaluations. Two indices were considered: prospective consistency and retrospective consistency. Of note, some recent papers use the term diagnostic stability coefficient as a synonymous of prospective consistency (89). Using the

broad ICD-10 F1-F9 categories as diagnoses, we computed prospective consistency comparing diagnoses made at the initial evaluation with those made at the final one, and retrospective consistency comparing diagnoses made at the final evaluation with those made at the initial one.

2. Diagnostic constancy: Since prospective and retrospective consistency were based on only two evaluations, they often do not reflect diagnostic processes based on multiple evaluations characteristic of routine clinical practice (Pierce et al., 2019). We thus included a criterion according to which subjects who received diagnoses of BD in at least 75% of the evaluations were categorized as "stable BD diagnosis".

8.2 Sample

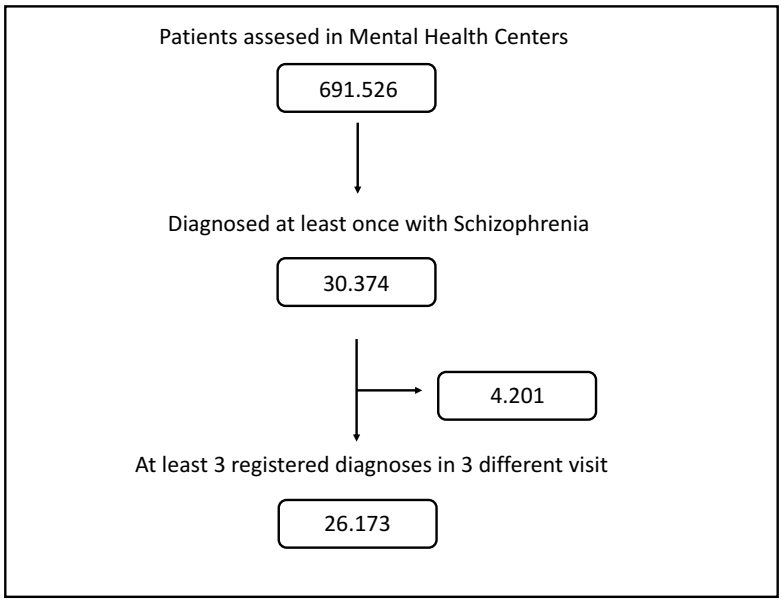
The Madrid Psychiatric Registration System established in 1980 includes all individuals treated in public mental health centers in Madrid (Spain) until 2008. Public mental health centers are part of the National Health Services and provide free medical coverage to a catchment area of about 6,000,000 inhabitants and are funded through taxes. The Madrid Case Registry (*Registro Acumulativo de Casos de la Comunidad de Madrid*) is a naturalistic study of diagnostic stability and consistency over time of the mental disorders in the area. It includes information from all psychiatric visits to public outpatient mental health clinics in the province of Madrid, Spain between 1980 and 2008. The database includes sociodemographic information and clinical diagnoses. From 1980 to 1992, all diagnoses in the registry were coded according to the 9th Revision of the International Classification of Diseases (ICD-9). Since 1992, diagnoses have been assigned according to the 10th Revision of the ICD (ICD-10). ICD-9 codes were converted to ICD-10 codes using the

guidelines published by the World Health Organization (WHO, 1993). The treating clinician (a psychiatrist or clinical psychologist) entered the diagnostic codes at every follow-up visit. A maximum of 2 diagnoses per patient per visit were recorded. Diagnostic counts in this study include comorbidities (for instance, a F1-F3 diagnosis would count both as F1 and F3).

8.2.1 Subsample for Hipothesis I: Shift changes in schizophrenia

For this study, we selected patients who met the following inclusion criteria: (I) diagnosed with schizophrenia (ICD-10 category F20) during at least one visit and (II) at least three registered diagnoses by psychiatrists or clinical psychologists (in three different visits to the outpatient clinics). In the resulting subset of patients (n=26,163) with a diagnosis of schizophrenia during at least one outpatient visit; we examined the diagnoses given to these patients during previous visits to public mental health clinics (i.e., prior to the diagnosis of schizophrenia). All methods were performed in accordance with the relevant guidelines and regulations and the Institutional Review Board of “Hospital 12 de Octubre” and “Fundación Jiménez Diaz” approved this study. See figure 3.

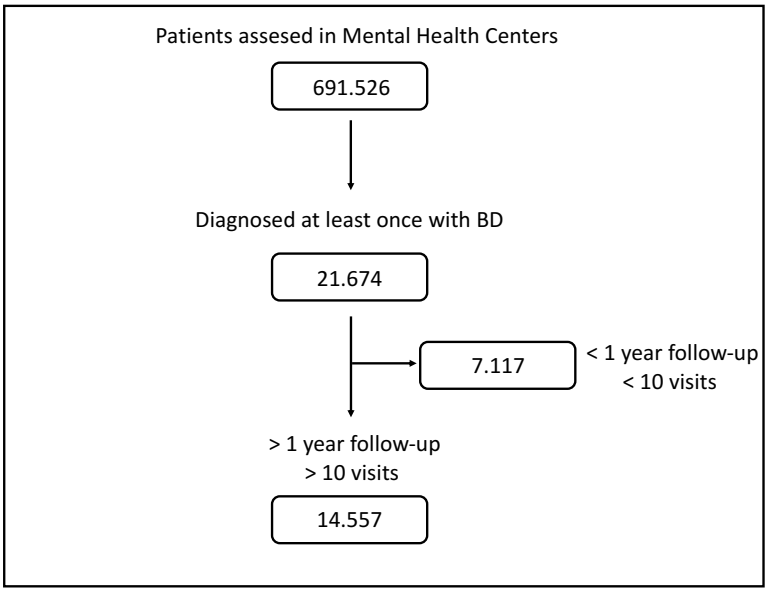
Figure 3. Sample Selection is Schizophrenia



8.2.2 Subsample por Hipotesis II: Shifg changes in BD

For this study, we selected patients who met the following inclusion criteria: adults aged ≥ 18 years who, (I) received a BD diagnosis in at least one visit and (II) undertook at least 10 visits over the study period (minimal adequacy of care, defined as having ≥4 outpatient visits in the last year and use of psychotropic medication, or ≥8 outpatient visits (with or without a medication - definition used in prior studies) (91,92). Out of a total population of 691,526 patients that were evaluated among 30 years, 14,557 met these inclusion criteria. This study was overseen by the Institutional Review Board at Instituto de Investigación Sanitaria-Fundación Jiménez Díaz. The RECORD guidelines were followed to report findings (93). See figure 4.

Figure 4. Sample selection in BD



8.3 Data analysis

8.3.1 Shift changes in schizophrenia

The probability of maintaining or changing the diagnosis of schizophrenia in the outpatient mental health visits was calculated in the 48 months following the initial diagnosis of schizophrenia. Probabilities were computed considering all the assessments made in one-month time lapses after the initial diagnosis. For each month, the total number of diagnoses was added and then divided proportionally according to their distribution. If a diagnosis was missing (due to longer delay between visits), the diagnosis recorded in the previous month was used instead.

8.3.1.1 Continuous-time hidden Markov model of longitudinal diagnostic shifts

In order to build a graph describing the sequence of diagnoses over time, a statistical model is needed that incorporates: I) the frequencies of the transitions among mental disorders, II) the time lags between consecutive psychiatric visits, and III) the diagnostic uncertainty due to missing information in some of the records. We used a novel technique based on a continuous-time hidden Markov model (CT-HMM) to build the graphical model (94). This method is based on a Hidden Markov model (HMM), which posits that starting from the current state of a stochastic system or process, it is possible to establish a description of its future probability, assuming that the measures are performed regularly in time. Nevertheless, in clinical practice that is not the case due to irregular or missed visits. Taking into account these cases, CT-HMM incorporates that both changes between hidden states and the appearance of new features can occur at any time (95).

The model makes use of the following assumptions: I) the different patients are instances of the same stochastic process to be modeled; II) the process is stationary, so that the intensity of the interaction between two diseases does not depend on the age of the subject; III) a patient stays in the same state (disease) until the time instant of the following medical claim; IV) the Markov property holds, meaning that the future clinical history of a patient only depends on his/her present state, and is independent of the past. In this model, each element q_{ij} in the matrix of parameters describes the strength of the relation between diseases *i* and *j* (96). The time lags between clinical events are also considered, and are represented as the sum of each row's elements $q_i=\sum_j q_{ij}$, which is high

if the average transition time between *i* and the next clinical event is short. Finally, records with incomplete or lost diagnoses are treated probabilistically as uncertain observations under some underlying "hidden" disorder. This uncertainty was reflected by the $b_{i(k)}$ parameters, which described the probability of a certain diagnosis *k* if the real underlying diagnosis was *i*.

8.3.2 Shift changes in BD

We used Chi square and Fisher exact test to test sociodemographic differences between people with stable and non-stable BD diagnosis. We then applied a multivariable logistic model to examine predictors of diagnostic shift including the significant variables of univariate analysis as covariates (gender, marital status, educational level, employment status, occupation, type of cohabitation and background; described in Table 1), selecting the final model with a progressive elimination method (the likelihood ratio was used as criteria for model fit).

Survival analyses were used to estimate the time from the beginning of the follow-up to the first diagnosis of BD. Mantel-cox was used to test the time difference of first BD diagnosis between people with stable and non-stable BD diagnosis.

RESULTS

9. RESULTS

9.1 Sample description

9.1.1 Shift changes in schizophrenia

A total of 26.163 patients were diagnosed with schizophrenia. About half of the patients in the sample were male (n = 13,941; 53.3%). Mean age at the first assessment was 37.6 years (SD = 15.5) and mean age at the first diagnosis of schizophrenia was 39.3 years (SD = 14.9). At the time of their first visit to the mental health centers the patients were generally single (n = 15,741; 62.5%), and half of them were living with their family of origin (50.5%). The patients generally had low educational attainment: 46.0% had completed less than fifth grade (n = 10,997). Only one out of every four patients was working (n = 5,964; 26.5%) while most other patients were unemployed (n = 4,820; 21.5%), home-makers (n = 3,587; 16.0%), disabled (n = 2712; 12.1%), or studying (n = 2325; 10.4%). The following variables pre- sented missing data over 2%: marital status (n = 965), educational level (n = 2,277), and working status (n = 3,737).

Patients with schizophrenia diagnoses made 1,455,063 visits to mental health centers (mean ± SD = 96.6 ± 174.7). The total number of visits with non-schizophrenia diagnoses prior to schizophrenia was 279,245, with an average of 18.03 visits per patient (SD = 28.81)

9.1.2 Shift changes in BD

A total of 14,557 patients were diagnosed with BD. Out of the sample, 63.9% were female (n=9,134). Regarding marital status, the majority of the sample (51%) were married. Concerning education, 32.2% had a primary school education, and only 9.2% had a higher education (degree or bachelor's degree) as opposed to 12.8% who were illiterate or had no education. Of the sample, 29.6% had an active job, 9.5% had a work disability (temporary or permanent), 12% received a salary (retirement or income) and 24% were engaged in housework.

These patients received 848,147 psychiatric and/or psychological consultations. The mean follow-up time for these patients was 3,295.9 days (standard deviation [SD] 1,967.6 days), the mean number of visits was 58.3 (SD 66.7), and the median was 38 visits. Socio-demographic data is shown in appendix.

9.2 Transitions

9.2.1 Shift changes in schizophrenia

9.2.1.1 Previous diagnoses.

In the sample, 56.7% of individuals (14,883/26,163) had received another diagnosis prior to being diagnosed with schizophrenia. Table 2 describes the diagnoses received in the first outpatient assessment at the mental health centers. Table 3 details the most frequent diagnoses on a per-visit basis prior to the diagnosis of schizophrenia. When comorbidities were included, the most frequent diagnostic categories prior to

schizophrenia followed the order of diagnoses reported in table 3: 'mood disorders' (F3; 31.3%), 'neurotic, stress-related and somato- form disorders' (F4; 21.2%), non-schizophrenia diagnoses within the category of 'schizophrenia, schizotypal and delusional disorders' (F2; 19.2%), and 'disorders of adult personality and behavior' (F6; 12.3%). With respect to psychotic disorders specifically, 7.2% of patients were previously diagnosed of persistent delusional disorders (F22), 5.0% of acute and transient psychotic disorders (F23), 3.7% of unspecified nonorganic psychosis (F29), and 3.1% of schizoaffective disorders (F25).

Table 2. Most frequent diagnoses at initial assessment. Only diagnoses made at least in 100 visits are listed

ICD-10 diagnoses	Frequency	Percentage
F20 Schizophrenia	11330	43.3
F4 Neurotic, stress-related and somatoform disorders	4516	17.3
F3 Mood disorders	3402	13.0
F22 Persistent delusional disorders	1068	4.1
F6 Disorders of adult personality and behavior	1034	3.9
F1 Mental and behavioral disorders due to psychoactive substance use	981	3.7
F23 Acute and transient psychotic disorders	776	3.0
F29 Unspecified nonorganic psychosis	656	2.5
F9 Behavioral and emotional disorders with onset usually occurring in childhood and adolescence	345	1.3
F25 Schizoaffective disorders	276	1.1
F0 Organic mental disorders	263	1.0
F7 Mental retardation	211	0.8
F5 Behavioral syndromes associated with physiological disturbances and physical factors	205	0.8
F8 Disorders of psychological development	145	0.6
Total	24465	96.3

Table 3. Most frequent diagnoses until schizophrenia. Only diagnoses made at least in 1000 visits are listed.

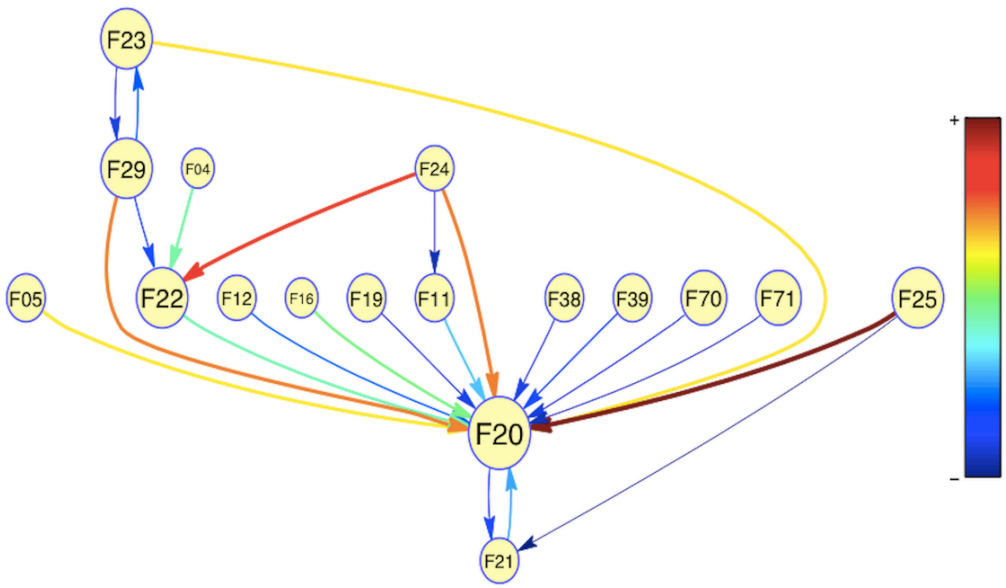
ICD-10 diagnoses	Frequency	Percentage
F3 Mood disorders	77461	27.7
F4 Neurotic, stress-related and somatoform disorders	50468	18.1
F6 Disorders of adult personality and behavior	24257	8.7
F22 Persistent delusional disorders	17660	6.3
F23 Acute and transient psychotic disorders	11907	4.9
F1 Mental and behavioral disorders due to psychoactive substance use	11639	4.3
F29 Unspecified nonorganic psychosis	8632	4.2
F25 Schizoaffective disorders	7488	3.1
F0 Organic mental disorders	4661	2.7
F9 Behavioral and emotional disorders with onset usually occurring in childhood and adolescence	4563	1.7
F7 Mental retardation	3468	1.6
F3-F6	3147	1.2
F3-F4	2789	1.1
F5 Behavioral syndromes associated with physiological disturbances and physical factors	2762	1.0
F4-F6	2647	1.0
F8 Disorders of psychological development	1886	0.9
F1-F6	1528	0.7
F1-F3	1423	0.5
Total	252127	90.3

9.2.1.2 Transition to schizophrenia

We then examined the probability of progression of these diagnoses to schiz- ophrenia, considering also time proximity (figure 5). As might be expected, the strongest associations were between the psychotic-spectrum diagnoses and schizophrenia. From the highest to the lowest probability, schizoaffec- tive (F25) disorders, induced delusional disorders (F24), unspecified nonorganic psychosis (F29), acute and transient psychotic disorders (F23), persistent delusional disorders (F22), and schizotypal disorders (F21)

were all connected with a subsequent diagnosis of schizophrenia. The transitions to schizophrenia were indirect in some cases, usually through other psychotic disorders (for instance, F24 to F22 to F20).

Figure 5 Probabilistic links of ICD-10 diagnoses converging into schizophrenia (F20).



The size of the circles indicates the frequency of the diagnoses in our sample. The color and width of the arrows describe the strength of the interactions according to the model. F04/F05: Organic amnesic syndrome/Delirium, not induced by alcohol and other psychoactive substances; F11/12/F16/F19: Mental and behavioral disorders due to use of opioids/cannabinoids/hallucinogens/multiple drug use and use of other psychoactive substances; F21: Schizotypal disorder; F22: Persistent delusional disorders; F23: Acute and transient psychotic disorders; F24: Induced delusional disorder; F25: Schizoa ective disorders; F29: Unspeci ed nonorganic psychosis; F38: Other mood disorders; F39: Unspeci ed mood disorder; F70/F71: Mild/moderate mental retardation.

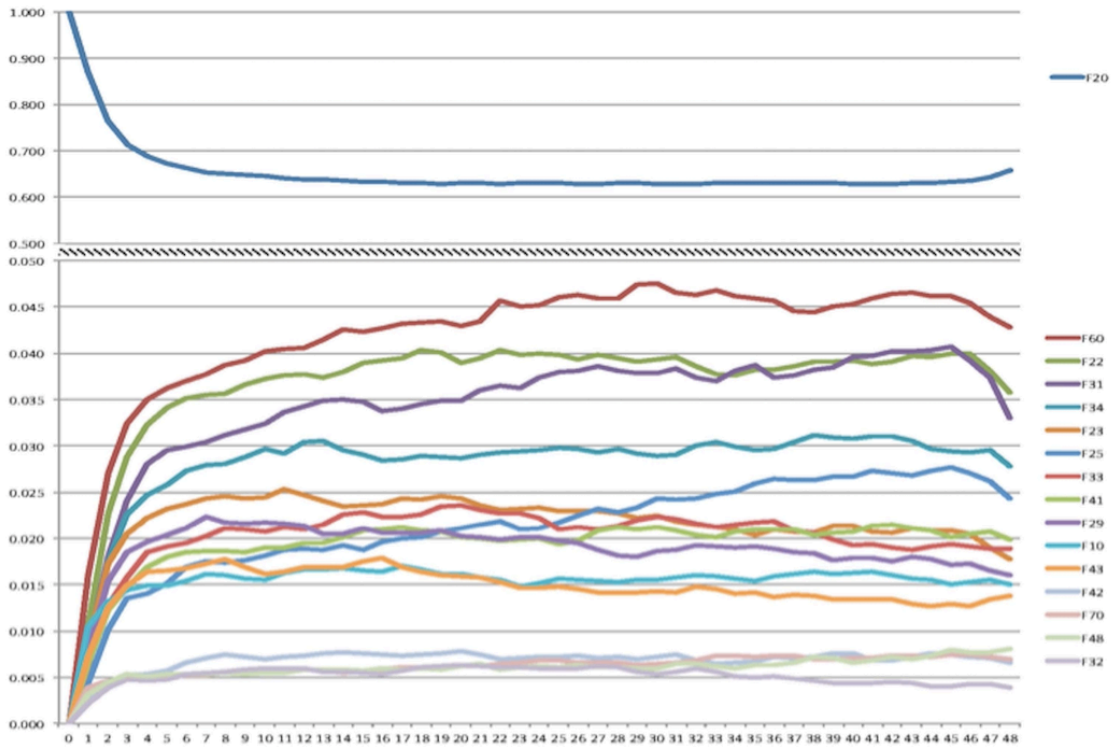
A probabilistic link from unspecified or other affective disorders (F38 and F39) towards schizophrenia was also represented, but not from bipolar disorder or major depression. Patients with alcohol, cannabis and multiple drug use disorders (F11, F12 and F19) were consequently diagnosed schizophrenia, but a direct transition was particularly frequent

for the fewer subjects with hallucinogen use disorders (F16). Some organic mental disorders appeared also in our model. The diagnosis of delirium not induced by psychoactive substances (F05) showed a high probability of direct transition into schizophrenia, while the less frequent organic amnesic syndrome (F04) usually progressed to persistent delusional disorders before. Mild and moderate cases of mental retardation (F70 and F71) were also directly linked to schizophrenia in our model. Of note, several diagnostic categories, such as anxiety (F4) and personality disorders (F6), were not represented in the graphical model.

9.2.1.3 Evolution/stability

Diagnostic shift from schizophrenia was more commonly toward the following diagnoses, represented by the average percentage in the first 48 months: personality disorders (F60: 4.2%), delusional disorders (F22: 3.7%), bipolar disorder (F31: 3.5%), persistent mood disorders (F34: 2.8%), acute and transient psychotic disorders (F23: 2.2%) or schizoaffective disorder (F25: 2.1%). However, the majority (64.5%) of the patients with an initial diagnosis of schizophrenia continued to receive the same diagnosis in subsequent assessments (figure 6). Patients who had a diagnostic shift from schizophrenia to a non-schizophrenia diagnosis did so generally in the first six months after the diagnosis of schizophrenia had been made. After that time interval the rates of each diagnostic category remained stable.

Figure 6. Diagnostic evolution of the first diagnosis of schizophrenia (F20) in the following 48 months.



The upper section shows the probability (0.5–1) of maintaining the F20 diagnosis. The lower section shows the probability (0–0.05) of changing this diagnosis.

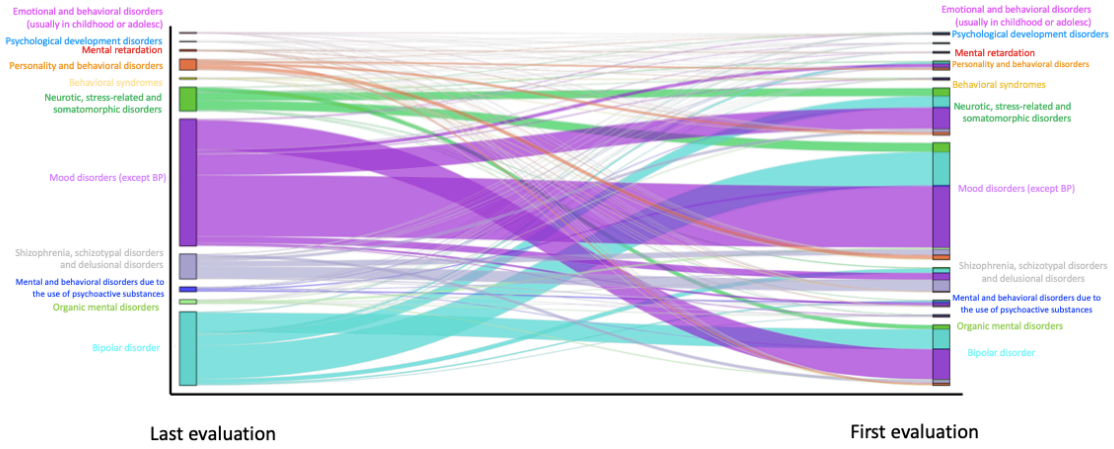
9.2.2 Shift changes in BD

9.2.2.1 Prospective consistency in Psychiatric Diagnoses

Our consistency comparisons included 15,082 diagnoses made at the initial evaluation and 15,507 at the final one. Figure 7 depicts retrospective diagnostic shifts.

The greatest prospective consistency was found among subjects diagnosed with mood/affective disorders (F3 category): 77.7%of patients diagnosed with F3 in the initial evaluation received a diagnosis under the same category at the final evaluation.

Figure 7. Retrospective consistency in Psychiatric Diagnoses



This graph is an "alluvial diagram". On the left side we see the final diagnoses, the width of each bar represents the number of patients with such diagnosis. On the right side we see the initial diagnoses also in proportion. Focusing on how a particular color bar in the left side (final diagnosis) splits into several other bars on the right side we can trace the proportions of the different initial diagnoses that converge to the same final diagnose.

We also found a high prospective consistency in patients diagnosed with Schizophrenia, schizotypal disorders and delusional disorders (F2):60% of these patients also received the same diagnosis. On the contrary, patients diagnosed with Mental and behavioral disorders due to the use of psychoactive substances (F1) had a low prospective consistency 30,5%.

9.2.2.2 Prospective consistency of BD diagnoses

A total 3,988 patients received a BD diagnosis in their first visit; 5,396 received it in their final visit; and 2,026 received it in both visits. Prospective and retrospective consistencies were, respectively, 50.8% and 37.5%. Cohen's kappa between first and last BD diagnoses was found to be low (k=0.17).

A category that led to diagnostic shift was Schizophrenia, schizotypal disorders and delusional disorders (F2), in 8.3% (n=449) of patients finally diagnosed with BD received an initial F2 diagnosis; and in 7.5% (n=301) of cases as the final diagnosis of those who were diagnosed with BD at the beginning.

With regards to non-bipolar affective disorders, out of 8,141 patients initially diagnosed with Mood/affective disorders (F3), 51% (n=4,153) had a non-BD diagnosis. In the final evaluation, 9,717 patients were assigned a Mood/affective disorders diagnosis (F3), of which 5,396 were bipolar (F31).

One in five patients (21%, n= 1,135) initially diagnosed with Neurotic disorders, stress-related disorders and somatomorphic disorders (F4) were diagnosed with BD in the last visit. Conversely, 10.7% (n=429) of patients diagnosed with BD at the first evaluation ended up with a Neurotic, stress-related and somatomorphic disorders.

Diagnoses of Personality and behavioral disorders in adults (F6) were initially assigned to 3.6% (n=195) of those with a final diagnosis of BD. Conversely, in the last evaluation, Personality and behavioral disorders in adults' diagnoses amounted to 4.2% (n=167) of those initially categorized as BD. The remaining diagnostic categories appeared in less than 3% of initial or final BD diagnoses.

9.2.2.3 Diagnostic stability of BD diagnoses

Out of the total sample of 14,557 patients, only 18.6% (n = 2,718) were categorized as stable BD diagnosis (e.g. retained the BD diagnosis in >75% of clinical encounters). We summarize the findings in figure 8.

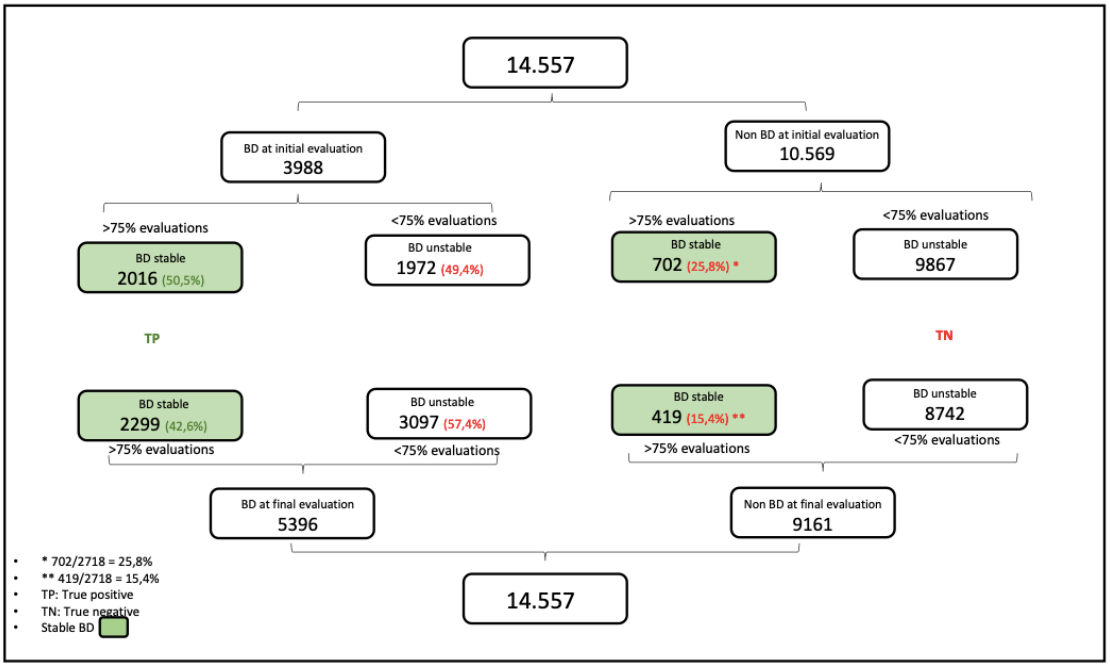
.

Among these 2,718 "stable" patients diagnosed with BD, the mean time from the first therapeutic contact with the Mental Healthcare Center to the first time the patient was diagnosed with BD was 318.1 days (95% CI: 188.5-347.8). The average time from the first therapeutic contact within the Mental Healthcare Center to the last time the patient was diagnosed with BD was 7,386.7 days. The median was 7,429 days (95% CI: 7,068.2-7,705.2). There is a difference between the time needed to make the first diagnosis of BD between those who kept a stable diagnosis (median 0 days) and those who did not (median 966 days) (Mantel-Cox test, chi-square=2,852.10; p<0.0001). If only patients with a stable BD diagnosis are taken into account, they represent 0.4% of the sample. Taking into account the total number of patients evaluated, about 3.1% had a diagnosis of BD at some point during follow-up.

More than 50% of the sample has been evaluated by the same psychiatrist at least 66% of the time. Among stable BD patients, 77% of the times have been seen by the same psychiatrist and among the BD unstable patients 64% of the times have been seen by the same psychiatrist (T student=27.081, df=14.083, p<0.001). Stable BD patients who have not been evaluated by the same psychiatrist at the first and last evaluation are 14.2% of the time and stable BD patients who have been seen by the same psychiatrist at the first and the last evaluation are 28.1% with an OR 2,375 (95% CI: 2,174 – 2,594).

The agreement between the first diagnosis with BD stable group compared to the last diagnosis with BD stable group (BD stable – first diagnosis: k=0.492, p=<0,000; BD stable – last diagnosis: k=0.420, p=<0,000, respectively).

Figure 8. BD diagnoses related to follow-up time



9.2.2.4 Factors related to the diagnostic stability of BD

The United Nations, for statistical purposes, defines persons between the ages of 15 and 24 as youth. Following this definition when performing the analysis on the stable BD group (when retained the BD diagnosis in >75% of clinical encounters), greater stability was found if the diagnosis was made after hospitalization OR 1,932 (95% CI: 1,682 - 2,219), if the age of onset was >25 years OR 4,318 (95% CI: 2,527 – 7,377), if the diagnosis of BD was made at age < 24 years OR 6,133 (95% CI: 3,477 – 10,817), if >65% of the visits were held by the same psychiatrist OR 2,246 (95% CI: 1,978 – 2,550), and if the patient had been assessed by the same psychiatrist in the first and last assessments OR 1,667 (95% CI: 1,475 – 1,883). Hosmer and Lemeshow Test: Chi-square= 10,620, df=7, P=0,156.

DISCUSSION

10. DISCUSSION

10.1 Shift changes in schizophrenia

This is a naturalistic study that describes which diagnoses the patients received before and after being diagnosed with schizophrenia for the first time. We found that although the most frequent prior diagnoses for those patients were anxiety (F4) and mood disorders (F3), direct transitions to schizophrenia usually came from psychotic spectrum disorders. Furthermore, we also found that the initial diagnosis of schizophrenia is less likely to change if it is confirmed some months after its onset.

Indeed, psychotic symptoms have been noted as the best predictor of progression to schizophrenia among individuals at high risk (42). Schizoaffective disorders are boundary diagnoses placed halfway between bipolar disorder and schizophrenia (34). Schizotypal disorders and non-specified psychotic disorders are frequently used as a proxy for schizophrenia (39,48,97). Therefore, it is not surprising that these diagnoses showed the most probable transitions to schizophrenia (figure 5). The uncommon diagnosis of induced delusional disorder also led to schizophrenia in the short term according to our model. The literature shows that a second diagnostic look sometimes reveals that a patient with supposedly induced delusional disorder shares the same genetically driven form of paranoid schizophrenia as an affected relative (98). However, there are only a few case reports of this clinical entity (99), and more studies are needed.

In line with the high prevalence of prodromal depressive symptomatology (44), many

patients with schizophrenia were previously diagnosed with anxiety or mood disorders. However, direct progression from affective or anxiety disorders did not appear as a probable transition to schizophrenia, implying that most patients received other diagnoses before eventually receiving a diagnosis of schizophrenia. In fact, anxiety and affective symptoms could be reactive expressions to prodromal and psychotic symptoms (100). Personality disorders were the third most common diagnoses after mood and anxiety disorders, and likewise they were not directly followed by schizophrenia. This might reflect the uncertainty of the clinicians, who try to avoid direct transitions between conflicting diagnoses and perhaps some caution when it comes to diagnose a chronic and severe disorder such as schizophrenia. In other words, the high frequency of less severe previous diagnoses could be due to a conservative approach dealing with diagnostic uncertainty in early stages of a mental illness.

The direct transition between alcohol or drug use disorders represented in our model agrees with the mounting evidence of a causal influence of psychoactive substances in schizophrenia, particularly cannabis (101). On the other hand, one study found that externalizing disorders were more frequent in the childhood and adolescence of patients with schizophrenia compared with those with bipolar disorder or depression (102). However, behavioral and emotional disorders with onset usually occurring in childhood and adolescence (F9) were rarely registered before schizophrenia (1.3%) and the transition from F9 to schizophrenia was uncommon in our sample.

We also found that most patients retained the diagnosis of schizophrenia in the four years after its onset (65%). Although schizophrenia has a high diagnostic stability, according to

previous works, prospective consistencies can vary from 52–100% (19,62,85,103,104). In our sample approximately 35% of the patients diagnosed with schizophrenia, presented a change in their diagnosis (prospective consistency of 65%). This finding is consistent with the 70% temporal consistency that we have previously reported for schizophrenia in another sample, mostly assessed in outpatient facilities (19). Overall, the results suggest that clinical assessment, which appears to be the most accurate diagnostic procedure for psychotic disorders (105), repeatedly maintains the diagnosis of schizophrenia once it is confirmed. Moreover, the time lapse for its confirmation agrees with the 6-month current diagnostic criteria in DSM-IV, which has been described as excessively restrictive as it might lead to the detection of chronic subjects with a worse prognosis(3,59,106). When not confirmed, the most frequent alternative diagnoses are personality, affec- tive or non-schizophrenia psychotic disorders.

Even though the diagnosis of schizophrenia is more common in adolescence and early adulthood (107,108), the risk of developing psychosis persists, and the age-incidence relationship seems to be altered by gender (109). The evidence to date shows that men have an increased incidence at the end of the second decade and the beginning of the third, subsequently presenting lower rates, which remain thereafter. However, in women the first peak presents later on, with a smaller second peak in middle age (107) and at 65 years, possibly a third peak (110). In our cohort the observed age of onset fits on the upper limit reported in the literature (20–40 years at disease onset (111)). More than 50% of Kraepelin patients had onset of symptoms between ages 30 to 40, and over 20% between ages 40 to 50 (112).

Novel approaches, supported by recent epidemiological and clinical research (113,114), try to consider the differential weight of dimensional traits of schizophrenia. In this study we expected to find a diagnostic pattern before schizophrenia that would correspond to the four symptom dimensions described in dimensional models (49). Prior affective disorders and non-schizophrenia psychotic disorders were frequent in our sample, although the former seem to precede the latter in most cases. The other two symptom dimensions proposed in patients with schizophrenia (negative and cognitive) were rarely translated into specific diagnoses. This might suggest either that clinicians disregarded negative and cognitive symptoms as being part of a more severe clinical picture (affective or psychotic) or that an asymmetrical model of dimensions involved in schizophrenia would fit epidemiological data more closely. It should be noted that diagnostic classifications have evolved over time. The diagnostic criteria of ICD 9 are based on Schneider´s first rank symptoms and overlook negative and cognitive symptoms. This is less true for ICD 10 criteria.

This study examines clinical practice in Spain in real-world conditions, as it evaluates the follow-up of a large sample of patients with schizophrenia consecutively recruited in a 30-year interval. The representativeness of the study is enhanced by the free access and wide coverage of public medical care in our country. Moreover, since residential changes to other provinces are infrequent in Spain (<2% per year) (115) and visits to other mental health centers of the region would be included in the registry, our data is likely to reflect the real-world diagnostic pathways of the patients. However, some patients had probably received the first diagnosis of schizophrenia before being included in the registry or

outside the system (e.g., in a private consultation). Thus, our results likely represent not only incidence cases but also prevalent cases that have increased the mean age at first diagnosis. However, other studies have also reported similar mean age at first diagnosis in schizophrenia. Despite this weakness, our results still provide an estimate of the shift patterns in the diagnosis of schizophrenia among patients who were retained in the public mental health system of Madrid. On the other hand, as structured diagnostic instruments were not used, it is unclear to what extent the clinical picture is changing immediately prior to the diagnosis of schizophrenia or in the early phases of illness, or whether clinicians vary in their ability to recognize the disorder. However, clinical evaluation seems so far to be the best diagnostic tool for schizophrenia (105), providing reliable and valid diagnoses when performed by mental health specialists (116).

The evolution of diagnoses before and after that of schizophrenia indicates frequent initial misdiagnoses or comorbidity with affective, anxiety and personality disorders. Nevertheless, a diagnosis of schizophrenia is usually reached from psychotic spectrum disorders or directly assigned, and once reached it is confirmed in the following six months for two of every three patients.

10.2 Shift changes in BD

The present study addresses the issue of diagnostic stability of BD in outpatient settings and contributes to the knowledge about the temporal consistency of BD and the usual diagnostic changes that occur during its evolution.

DISCUSSION

The results showed a limited number of stable BD diagnoses and notably lower than previous studies. Some methodological reasons could explain the differences with previous studies, especially the low number of evaluations and the shorter follow-up period used in previous studies. To be sure of the diagnosis of stable BD is needed at least between 6-12 months. The administrative prevalence (the proportion of the population in a defined area -the community of Madrid in this study- who are receiving services) of BD in this psychiatric sample is 0.4%, lower than usually reported in clinical and non-clinical samples (17,117,118). However, this could be related to the fact that diagnoses were made in outpatient settings and as reported in previous studies (3,57) a higher diagnostic stability is observed when the diagnosis is made after a discharge from hospital. A more detailed study of the factors that influence the stability of BD and a better knowledge of the course of diagnoses throughout its evolution are proposed as future lines of research.

The natural evolution of BD is prone to a high variability, however, the central symptoms of affective episodes are not present as frequently and the presence of comorbid disorders, which is quite common, leads to misdiagnoses during daily clinical practice (119,120). The present study found that the stability of BD was low and even lower than in previous studies (3,54,57,59,69), with the three different indices used.

In the first evaluation 27.4% of the patients received a diagnosis of BD and only 18.3% of the total sample was considered stable according to the criteria established in this study. Additionally, confusion around the usual differential diagnoses of BD was found. All these conclusions are detailed and discussed in the following sections.

It is important to emphasize that the increased specificity of the diagnostic criteria for BD in ICD-10 versus ICD-9 may have somewhat influenced our results.

It is essential to clarify, that there has been no deinstitutionalization in Spain and the registry was conceived as a tool during the psychiatric reform. On note, this is one of the few epidemiological studies on this issue conducted outside Scandinavia and where multiple types of stability measures are used and this database has been used in previous works (65,121).

10.2.1 BD diagnosis at the first evaluation

Diagnostic shifts in BD are especially frequent at first contact with the physician, misleading initial symptoms due to substance abuse, depressive or psychotic symptoms. The greatest prospective consistency was found in Mood/affective disorders category (F3), since the sample was selected among patients with at least one BD diagnosis. While only 27.4% of subjects were diagnosed with BD at the first evaluation, the rest were diagnosed at least once during subsequent evaluations. Similar results were presented in a previous study *(69)*, where it was noted that these figures were consistent with the high prevalence of misdiagnosis (48 and 69%) found in naturalistic research using self-administered questionnaires in general practitioner consultations *(122,123)* and also in studies in which diagnoses were based on the application of DSM-IV criteria *(124)*.

However, in our sample, 50.5% of patients who were diagnosed with BD at the first evaluation remained stable in three quarters of the evaluations. This fact is not consistent with the figure reported by Chen et al. *(54)*, who noted that 70% of the subjects with an

initial diagnosis of BD did not change to a different category during evolution. On the other hand, the percentage of patients with a stable diagnosis of BD (n=2,718) who were correctly diagnosed in the first evaluation (n=2,016) increases in our sample to 74.2%. These results support the hypothesis of the diagnostic difficulty of BD in the first evaluations.

10.2.2 BD diagnosis at the last evaluation

The latest evaluation showed an increase in the number of diagnoses of BD (37.1% of the sample) and, of those, 42.6% had been stable throughout the study. On the other hand, 84.6% of patients with stable diagnoses (n=2,718) were accurately diagnosed in their last visit (n=2,299).

This result may reflect a progressive increase in diagnostic stability throughout the evaluations (in our case a minimum of ten), which is congruent with the idea that routine reassessment could improve the chances of a successful diagnostic process. However, Schwartz et al. in 2000 (3) reported that the retrospective consistency of BD was 85% when comparing 6-month and 24-month diagnoses, but was reduced to 73% when comparing initial and 24-month diagnoses. This would mean that consistency rates for some diagnoses decreased as the follow-up period increased. In any case, the retrospective consistency of our study's BD (37.5%) is low compared to other studies that measured it (58.4-94.4%), similar to that presented by Baca-García et al. in 2007 (38%) (65) and higher than Weeke et al. in 1984 (20%) (4). The low retrospective consistency might be explained by the fact that this is a longitudinal study based on data retrieved

from community mental health centers, hence conducted in a real-world scenario and in general population, and not in a specific BD unit where patients start off after being already correctly diagnosed.

10.2.3 Diagnostic stability of BD

To our knowledge, this has been the largest longitudinal study, 14,557 patients over 30 years of study, which has evaluated the diagnostic stability of BD under ecological conditions. In 2005, Kessing (69) mentioned that no study had investigated the diagnostic stability of the most common ICD-10 psychiatric diagnoses administered under ecological clinical conditions. This is the case of our study, which has shown a low stability of the ICD-10 BD categories, measured by temporal consistency and diagnostic constancy, with findings considerably lower than in previous studies. The reasons for these differences in diagnostic temporal stability are not clear, but may be due to the large sample size, the extensive duration of follow-up, the high number of evaluations, diagnostic criteria or sociodemographic variables.

Time consistency showed low results with a prospective consistency of 50.8% and a retrospective consistency of 37.5%. It should be noted that the kappa value was low (kappa=0.17) between the first and last diagnosis. However, since kappa values take into account stable positive cases and stable negative cases, but also cases that remit and new cases, low kappa values can be observed if a high number of new or remitting cases occur (125) and therefore do not necessarily reflect a lack of diagnostic stability.

DISCUSSION

The results of our study showed that only 18.3% of patients were diagnosed with BD in 75% of the evaluations. In these patients with stable diagnosis, the mean number of evaluations until the first diagnosis of BD was 3.7 and an average time of 318.1 days. These values were increased to 21.2 evaluations and 1,511.2 days within the group with a non-stable diagnosis. Thus, patients with a stable diagnosis of BD were diagnosed earlier (less than one year) and needed fewer evaluations than those without a stable diagnosis (somewhat more than 4 years until the BD diagnosis was made).

In our study, patients with stable BD diagnosis achieved diagnostic stability at 7,386.7 days (slightly more than 20 years) and 279 follow-up visits. And patients with a non-stable diagnosis had their diagnosis withdrawn at 2,929 days (approximately 8 years) and after 55 evaluations. These results could be in line with previous reports by Hirschfeld in 2003 and Baldesarini in 1999, who reported a delay in correct diagnosis of about 8-10 years from the onset of the disease (122,126). In our study, the data suggest that both consolidating and withdrawing the diagnosis of BD is a task that requires many years of follow-up and a numerous evaluation, while consolidating needed about 14 visits/year, withdrawing entailed less than 7 visits/year. This fact may reflect that patients with a stable diagnosis of BD are more complex and require more health care, than those for whom this diagnosis is withdrawn.

Although it was not the main objective of our study, four variables were not only included, but were predictive of diagnostic stability of BD: marital status, educational level, work situation and personal history of psychiatric care. In another preliminary study (19) four variables related to the stability of bipolar diagnosis were found: sex, age ≥ 40 years,

number of psychiatric consultations and outpatient Mental Health Centers. In any case, more studies focusing on these variables are needed.

The higher consistency rates found by other authors (3,54,57,59,69) may have been influenced by a number of drawbacks that diminish the generalizability of these studies.

10.2.4 Diagnostic shifts in BD

Patients with a stable BD diagnosis had some diagnostic fluctuation that included the typical differential diagnoses of BD. Our study found high rates of misdiagnosis of BD with other affective disorders: 44.5% of patients who were diagnosed at the first evaluation of a non-bipolar affective disorder were eventually diagnosed at the last follow-up visit with BD; and 51% of patients who were initially diagnosed with BD were no longer diagnosed at the last evaluation. Previous studies concur that the high rates of misdiagnosis derive from confusion with unipolar depression *(122,123)*, especially in cases where the BD debuts with one or more depressive episodes. As for neurotic and anxiety disorders, the percentage of these diagnoses at the beginning is high (21.03%) in patients who are ultimately diagnosed with BD. Other less frequent diagnostic shifts occurred with the spectrum of schizophrenia (7.5% at first evaluation and 8.3% at the last evaluation) and with personality disorders (4.2% baseline and 3.6% final).

Many factors may be involved in the unstable progression of a psychiatric diagnosis. Schwartz mentioned that diagnostic changes over time may reflect the evolution of a disease, the emergence of new information, or the unreliability of measurements (3). The relative lack of stability in diagnoses over time in this study may be due to disease

progression or reflect weaknesses inherent in clinical evaluations. There could be other ways to determine the validity of stable diagnoses, such as the use of prescriptions for mood stabilizers, this being a limitation of the study.

The results of this paper raise concerns about psychiatric research findings, especially in studies with short follow-up periods for chronic conditions that may not allow enough time to reach an accurate diagnosis or in studies that do not take into account the context.

Our study has limitations. Our study has limitations. First, we did not consider differences between type I and II bipolar disorder, despite clinical and prognostic implications, because codes were mostly recorded following ICD-9, where type II bipolar disorder could not be specified. Second, diagnoses in our database are recorded following the independent judgement of clinicians rather than alternative assessments such as research scales. This, however, enhances external validity of our results as they likely reflect the course of illness from a real clinical practice perspective. While diagnostic scales can reduce measurement error, they (i) require specific training and result too time-consuming to be used routinely and (ii) are mostly validated in the context of highly selected samples of patients, for research purposes. In conclusion, as noted previously in the literature (127), our results should be considered an externally valid representation of patterns of real clinical diagnostic change, which has important implications for treatment planning, rather than patterns in the prevalence of the disorder. Also, a limitation was that the form filled out at each visit consisted of socio-demographic data and psychiatric diagnosis, leaving out other relevant data.

A point to keep in mind is that whether diagnostic changes in our data (e.g., patients whose diagnoses changed from or to bipolar disorder) reflect misdiagnosis, the natural history of the phenotypical presentation of these patients' disease, or a mix of both cannot be clarified using this data source. Accordingly, conclusions regarding over or underdiagnosing of BD based on our results should be made with caution. The study is also limited by the possible existence of uncontrolled pathways of psychiatric care, but may more accurately reflect the real clinical practice, perhaps revealing the poor accuracy of clinical evaluation systems in usual practice.

CONCLUSIONES

11. CONCLUSIONES

1. Los diagnósticos más frecuentes antes del diagnóstico de la esquizofrenia son los trastornos de ansiedad y del estado de ánimo, sin embargo las transiciones directas a la esquizofrenia provienen de trastornos del espectro psicótico.
2. Antes y después del diagnóstico de la esquizofrenia son frecuentes los diagnósticos erróneos o la comorbilidad con los trastornos afectivos, de ansiedad y de personalidad.
3. La estabilidad diagnóstica en el trastorno bipolar es un área poco estudiada. Hay que recordar que la estabilidad es uno de los criterios de Guze y Robbis para validar un diagnóstico.
4. La literatura señala una alta estabilidad diagnóstica para el trastorno bipolar. No obstante, la práctica clínica diaria tiene limitaciones para valorar la estabilidad especialmente en lo referente a su evolución interfasica.
5. La estabilidad diagnóstica en el trastorno bipolar se alcanza por encima de los 10 años de evolución.
6. Los pacientes con trastorno bipolar más estables son diagnosticados en las primeras visitas.
7. Se observa una mayor estabilidad en el trastorno bipolar cuando la edad de inicio es mayor 25 años, cuando el diagnóstico de trastorno bipolar se realiza a una edad menor 24 años o si el diagnóstico se realiza tras hospitalización.

8. La baja estabilidad diagnóstica detectada en este estudio debe tenerse en cuenta a la hora de comparar con los ensayos clínicos y epidemiológicos, donde las muestras son más pequeñas y el seguimiento es más corto.
9. Proponemos un perfil clínico de alto riesgo para la inestabilidad diagnóstica en el trastorno bipolar: sexo femenino, síntomas psicóticos, cambios en el tratamiento, antecedentes familiares de trastorno afectivo y tardía de inicio de las manifestaciones psiquiátricas.
10. La mejora en el diagnóstico del trastorno bipolar podría lugar a intervenciones terapéuticas más tempranas mejorando el pronóstico.

CONCLUSIONS

12. CONCLUSIONS

1. The most frequent diagnoses before schizophrenia were anxiety and mood disorders. However, direct transitions to schizophrenia come from psychotic spectrum disorders.
2. Misdiagnosis or comorbidity with affective, anxiety and personality disorders are frequent before and after the diagnosis of schizophrenia.
3. Diagnostic stability in bipolar disorder is an understudied area, despite the fact that it represents the best means of validating the diagnosis of bipolar disorder.
4. High diagnostic stability is frequently reported for bipolar disorder; however, daily clinical practice is limited in its ability to assess what occurs between affective phases.
5. A delay of more than 10 years was found to reach diagnostic stability in bipolar disorder.
6. Patients with a stable bipolar disorder are diagnosed at the first visits.
7. Greater stability in bipolar disorder is observed if the age of onset is after 25 years, if a diagnosis is made before than 24 years or if the diagnosis is made after hospitalization.
8. The low diagnostic stability detected in this study should be considered when comparing with clinical and epidemiological trials, where samples are smaller and follow-up is shorter.

CONCLUSIONS

9. We propose a high-risk clinical profile for diagnostic instability in bipolar disorder: female sex, psychotic symptoms, changes in treatment, family history of affective disorder and late age of onset of psychiatric manifestations.
10. Improved diagnosis of bipolar disorder will lead to earlier therapeutic interventions.

TABLES

13. TABLES

FIGURES

14. FIGURES

BIBIOGRAFY

15. BIBLIOGRAFY

1. Chang WC, Chan S, Chung D. Diagnostic Stability of Functional Psychosis: a Systematic Review. Hong Kong Journal of Psychiatry. 2009 Mar 1;19.

2. Fischer BA, Buchanan RW. Schizophrenia in adults: clinical manifestations, course, assessment, and diagnosis [Internet]. 2011. Available from: www. uptodate.com

3. 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 Jun;57(6):593–600.

4. Weeke A. Admission pattern and diagnostic stability among unipolar and bipolar manic-depressive patients. Acta Psychiatr Scand. 1984 Dec;70(6):603–13.

5. Harvey PD, Sabbag S, Prestia D, Durand D, Twamley EW, Patterson TL. Functional Milestones and Clinician Ratings of Everyday Functioning in People with Schizophrenia: Overlap Between Milestones and Specificity of Ratings. J Psychiatr Res. 2012 Dec;46(12):1546–52.

6. Janoutová J, Janácková P, Serý O, Zeman T, Ambroz P, Kovalová M, et al. Epidemiology and risk factors of schizophrenia. Neuro Endocrinol Lett. 2016;37(1):1–8.

7. Lehman AF, Lieberman JA, Dixon LB, McGlashan TH, Miller AL, Perkins DO, et al. Practice guideline for the treatment of patients with schizophrenia, second edition. Am J Psychiatry. 2004 Feb;161(2 Suppl):1–56.

8. Murray CJL, Lopez AD, Organization WH, Bank W, Health HS of P. The Global burden of disease : a comprehensive assessment of mortality and disability from diseases, injuries, and risk factors in 1990 and projected to 2020 : summary [Internet]. World Health Organization; 1996 [cited 2021 May 6]. Available from: https://apps.who.int/iris/handle/10665/41864

9. Chen R, Liou T-H, Chang K-H, Yen C-F, Liao H-F, Chi W-C, et al. Assessment of functioning and disability in patients with schizophrenia using the WHO Disability Assessment Schedule 2.0 in a large-scale database. Eur Arch Psychiatry Clin Neurosci. 2018 Feb;268(1):65–75.

10. Harrow M, Herbener ES, Shanklin A, Jobe TH, Rattenbury F, Kaplan KJ. Followup of psychotic outpatients: dimensions of delusions and work functioning in schizophrenia. Schizophr Bull. 2004;30(1):147–61.

11. Lundin L, Flyckt L. [Schizophrenia past and present--the perception of long term prognoses have changed]. Lakartidningen. 2015 Oct 13;112.

12. Sabbag S, Twamley EW, Vella L, Heaton RK, Patterson TL, Harvey PD. Predictors of the Accuracy of Self Assessment of Everyday Functioning in People with Schizophrenia. Schizophr Res. 2012 May;137(1–3):190–5.

13. Smith K. Trillion-dollar brain drain. Nature. 2011 Oct 4;478(7367):15.

14. Ferrari AJ, Stockings E, Khoo J-P, Erskine HE, Degenhardt L, Vos T, et al. The prevalence and burden of bipolar disorder: findings from the Global Burden of Disease Study 2013. Bipolar Disord. 2016;18(5):440–50.

15. Hayes JF, Marston L, Walters K, King MB, Osborn DPJ. Mortality gap for people with bipolar disorder and schizophrenia: UK-based cohort study 2000-2014. Br J Psychiatry. 2017 Sep;211(3):175–81.

16. Moore PJ, Little MA, McSharry PE, Geddes JR, Goodwin GM. Forecasting depression in bipolar disorder. IEEE Trans Biomed Eng. 2012 Oct;59(10):2801–7.

17. Hidalgo-Mazzei D, Undurraga J, Reinares M, Bonnín C del M, Sáez C, Mur M, et al. The real world cost and health resource utilization associated to manic episodes: The MANACOR study. Rev Psiquiatr Salud Ment. 2015 Jun;8(2):55–64.

18. 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 Aug;15(4):425–34.

19. 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.

20. Whitty P, Clarke M, McTigue O, Browne S, Kamali M, Larkin C, et al. Diagnostic stability four years after a first episode of psychosis. Psychiatr Serv. 2005 Sep;56(9):1084–8.

21. Palomar-Ciria N, Cegla-Schvartzman F, Lopez-Morinigo J-D, Bello HJ, Ovejero S, Baca-García E. Diagnostic stability of schizophrenia: A systematic review. Psychiatry Res. 2019 Sep;279:306–14.

22. Woo BKP, Sevilla CC, Obrocea GV. Factors influencing the stability of psychiatric diagnoses in the emergency setting: review of 934 consecutively inpatient admissions. Gen Hosp Psychiatry. 2006 Oct;28(5):434–6.

23. Daradkeh TK. Stability of psychiatric diagnoses in clinical practice. Int J Soc Psychiatry. 1996;42(3):207–12.

24. Daradkeh T, El-Rufaie O, Younis Y, Ghubash R. The diagnostic stability of ICD-10 psychiatric diagnoses in clinical practice. Eur Psychiatry. 1997;12(3):136–9.

25. Akiskal HS, Akiskal KK, Lancrenon S, Hantouche EG, Fraud J-P, Gury C, et al. Validating the bipolar spectrum in the French National EPIDEP Study: overview of the phenomenology and relative prevalence of its clinical prototypes. J Affect Disord. 2006 Dec;96(3):197–205.

26. Kim B, Wang HR, Son JI, Kim CY, Joo YH. Bipolarity in depressive patients without histories of diagnosis of bipolar disorder and the use of the Mood Disorder Questionnaire for detecting bipolarity. Compr Psychiatry. 2008 Oct;49(5):469–75.

27. Rybakowski JK, Suwalska A, Lojko D, Rymaszewska J, Kiejna A. Bipolar mood disorders among Polish psychiatric outpatients treated for major depression. J Affect Disord. 2005 Feb;84(2–3):141–7.

28. Lopez-Castroman J. Errors, evolution and stability of bipolar disorder diagnosis. In: Plunkett JM, editor. Bipolar disorder: causes, diagnosis and treatment. New York, USA: Nova Science Publishers; 2011. p. 354.

29. Parker G. Borderline personality disorder and bipolar disorder: commentary on Paris and Black. J Nerv Ment Dis. 2015 Jan;203(1):13–4.

30. Zimmerman M. Is underdiagnosis the main pitfall in diagnosing bipolar disorder? No. BMJ. 2010 Feb 22;340:c855.

31. Liddle PF. The symptoms of chronic schizophrenia. A re-examination of the positive-negative dichotomy. Br J Psychiatry. 1987 Aug;151:145–51.

32. Schroder J, Buchsbaum MS, Siegel BV, Geider FJ, Lohr J, Tang C, et al. Cerebral metabolic activity correlates of subsyndromes in chronic schizophrenia. Schizophr Res. 1996 Mar;19(1):41–53.

33. Birnbaum R, Weinberger DR. Genetic insights into the neurodevelopmental origins of schizophrenia. Nat Rev Neurosci. 2017 Dec;18(12):727–40.

34. Nasrallah H, Tandon R, Keshavan M. Beyond the facts in schizophrenia: closing the gaps in diagnosis, pathophysiology, and treatment. Epidemiol Psychiatr Sci. 2011 Dec;20(4):317–27.

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

36. Bromet EJ, Kotov R, Fochtmann LJ, Carlson GA, Tanenberg-Karant M, Ruggero C, et al. Diagnostic shifts during the decade following first admission for psychosis. Am J Psychiatry. 2011 Nov;168(11):1186–94.

37. Chang WC, Pang SLK, Chung DWS, Chan SSM. Five-year stability of ICD-10 diagnoses among Chinese patients presented with first-episode psychosis in Hong Kong. Schizophr Res. 2009 Dec;115(2–3):351–7.

38. Fusar-Poli P, Cappucciati M, Rutigliano G, Heslin M, Stahl D, Brittenden Z, et al. Diagnostic Stability of ICD/DSM First Episode Psychosis Diagnoses: Meta-analysis. Schizophr Bull. 2016 Nov;42(6):1395–406.

39. Salvatore P, Baldessarini RJ, Tohen M, Khalsa H-MK, Sanchez-Toledo JP, Zarate CA, et al. McLean-Harvard International First-Episode Project: Two-Year Stability of ICD-10 Diagnoses in 500 First-Episode Psychotic Disorder Patients. J Clin Psychiatry. 2011 Feb;72(2):183–93.

40. an der Heiden W, Häfner H. The epidemiology of onset and course of schizophrenia. Eur Arch Psychiatry Clin Neurosci. 2000;250(6):292–303.

41. Häfner H, Maurer K, Trendler G, an der Heiden W, Schmidt M. The early course of schizophrenia and depression*. Eur Arch Psychiatry Clin Neurosci. 2005 Jun;255(3):167–73.

42. Miller PM, Lawrie SM, Byrne M, Cosway R, Johnstone EC. Self-rated schizotypal cognitions, psychotic symptoms and the onset of schizophrenia in young people at high risk of schizophrenia. Acta Psychiatr Scand. 2002 May;105(5):341–5.

43. Møller P, Husby R. The initial prodrome in schizophrenia: searching for naturalistic core dimensions of experience and behavior. Schizophr Bull. 2000;26(1):217–32.

44. Addington J, Cornblatt BA, Cadenhead KS, Cannon TD, McGlashan TH, Perkins DO, et al. At clinical high risk for psychosis: outcome for nonconverters. Am J Psychiatry. 2011 Aug;168(8):800–5.

45. Rosen JL, Miller TJ, D'Andrea JT, McGlashan TH, Woods SW. Comorbid diagnoses in patients meeting criteria for the schizophrenia prodrome. Schizophr Res. 2006 Jul;85(1–3):124–31.

46. Sara GE, Burgess PM, Malhi GS, Whiteford HA, Hall WC. The impact of cannabis and stimulant disorders on diagnostic stability in psychosis. J Clin Psychiatry. 2014 Apr;75(4):349–56.

47. Murray RM, Dutta R. The right answer for the wrong reasons? World Psychiatry. 2007 Jun;6(2):93–4.

48. Carpenter WT, Bustillo JR, Thaker GK, Os J van, Krueger RF, Green MJ. The psychoses: Cluster 3 of the proposed meta-structure for DSM-V and ICD-11: Paper 4 of 7 of the thematic section: 'A proposal for a meta-structure for DSM-V and ICD-11.' Psychological Medicine. 2009 Dec;39(12):2025–42.

49. van Os J, Kenis G, Rutten BPF. The environment and schizophrenia. Nature. 2010 Nov 11;468(7321):203–12.

50. Bromet EJ, Naz B, Fochtmann LJ, Carlson GA, Tanenberg-Karant M. Long-term diagnostic stability and outcome in recent first-episode cohort studies of schizophrenia. Schizophr Bull. 2005 Jul;31(3):639–49.

51. Ruggero CJ, Carlson GA, Kotov R, Bromet EJ. Ten-year diagnostic consistency of bipolar disorder in a first-admission sample. Bipolar Disord. 2010 Feb;12(1):21–31.

52. Addington J, Chaves A, Addington D. Diagnostic stability over one year in first-episode psychosis. Schizophr Res. 2006 Sep;86(1–3):71–5.

53. 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 Dec;175:537–43.

54. Chen YR, Swann AC, Johnson BA. Stability of diagnosis in bipolar disorder. J Nerv Ment Dis. 1998 Jan;186(1):17–23.

55. 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 Aug;151(8):1200–8.

56. 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(3):187–92.

57. 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 Oct;66(10):1239–46.

58. 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 May;38(5):535–9.

59. Veen ND, Selten J-P, Schols D, Laan W, Hoek HW, van der Tweel I, et al. Diagnostic stability in a Dutch psychosis incidence cohort. Br J Psychiatry. 2004 Dec;185:460–4.

60. Cegla-Schvartzman FB, Ovejero S, López-Castroman J, Baca-García E. Diagnostic Stability in Bipolar Disorder: A Narrative Review. Harv Rev Psychiatry. 2019 Feb;27(1):3–14.

61. Ratheesh A, Cotton SM, Betts JK, Chanen A, Nelson B, Davey CG, et al. Prospective progression from high-prevalence disorders to bipolar disorder: Exploring characteristics of pre-illness stages. J Affect Disord. 2015 Sep 1;183:45–8.

62. Amini H, Alaghband-rad J, Omid A, Sharifi V, Davari-Ashtiani R, Momeni F, et al. Diagnostic stability in patients with first-episode psychosis. Australas Psychiatry. 2005 Dec;13(4):388–92.

63. Baldwin P, Browne D, Scully PJ, Quinn JF, Morgan MG, Kinsella A, et al. Epidemiology of first-episode psychosis: illustrating the challenges across diagnostic boundaries through the Cavan-Monaghan study at 8 years. Schizophr Bull. 2005 Jul;31(3):624–38.

64. 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 Jun;27(3):189–93.

65. 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 Jun;115(6):473–80.

66. Hollister LE. Stability of psychiatric diagnoses among acutely ill patients. Psychopathology. 1992;25(4):204–8.

67. 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 May;38(5):535–9.

68. Cegla-Schvartzman FB, Ovejero S, López-Castroman J, Baca-García E. Diagnostic Stability in Bipolar Disorder: A Narrative Review. Harv Rev Psychiatry. 2019 Feb;27(1):3–14.

69. Kessing LV. Diagnostic stability in bipolar disorder in clinical practise as according to ICD-10. J Affect Disord. 2005 Apr;85(3):293–9.

70. Heslin M, Lomas B, Lappin JM, Donoghue K, Reininghaus U, Onyejiaka A, et al. Diagnostic change 10 years after a first episode of psychosis. Psychol Med. 2015 Oct;45(13):2757–69.

71. Heslin M, Lappin JM, Donoghue K, Lomas B, Reininghaus U, Onyejiaka A, et al. Ten-year outcomes in first episode psychotic major depression patients compared with schizophrenia and bipolar patients. Schizophr Res. 2016 Oct;176(2–3):417–22.

72. Jørgensen P, Mortensen PB. Admission pattern and diagnostic stability of patients with functional psychoses in Denmark during a two-year observation period. Acta Psychiatr Scand. 1988 Sep;78(3):361–5.

73. Kim JS, Baek JH, Choi JS, Lee D, Kwon JS, Hong KS. Diagnostic stability of first-episode psychosis and predictors of diagnostic shift from non-affective psychosis to bipolar disorder: a retrospective evaluation after recurrence. Psychiatry Res. 2011 Jun 30;188(1):29–33.

74. Angst J, Sellaro R, Stassen HH, Gamma A. Diagnostic conversion from depression to bipolar disorders: results of a long-term prospective study of hospital admissions. J Affect Disord. 2005 Feb;84(2–3):149–57.

75. Bukh JD, Andersen PK, Kessing LV. Rates and predictors of remission, recurrence and conversion to bipolar disorder after the first lifetime episode of depression--a prospective 5-year follow-up study. Psychol Med. 2016 Apr;46(6):1151–61.

76. Coryell W, Endicott J, Maser JD, Keller MB, Leon AC, Akiskal HS. Long-term stability of polarity distinctions in the affective disorders. Am J Psychiatry. 1995 Mar;152(3):385–90.

77. Dudek D, Siwek M, Zielińska D, Jaeschke R, Rybakowski J. Diagnostic conversions from major depressive disorder into bipolar disorder in an outpatient setting: results of a retrospective chart review. J Affect Disord. 2013 Jan 10;144(1–2):112–5.

78. Gilman SE, Dupuy JM, Perlis RH. Risks for the transition from major depressive disorder to bipolar disorder in the National Epidemiologic Survey on Alcohol and Related Conditions. J Clin Psychiatry. 2012 Jun;73(6):829–36.

79. James A, Wotton CJ, Duffy A, Hoang U, Goldacre M. Conversion from depression to bipolar disorder in a cohort of young people in England, 1999-2011: A national record linkage study. J Affect Disord. 2015 Oct 1;185:123–8.

80. Li C-T, Bai Y-M, Huang Y-L, Chen Y-S, Chen T-J, Cheng J-Y, et al. Association between antidepressant resistance in unipolar depression and subsequent bipolar disorder: cohort study. Br J Psychiatry. 2012 Jan;200(1):45–51.

81. Nakamura K, Iga J, Matsumoto N, Ohmori T. Risk of bipolar disorder and psychotic features in patients initially hospitalised with severe depression. Acta Neuropsychiatr. 2015 Apr;27(2):113–8.

82. Østergaard SD, Straszek S, Petrides G, Skadhede S, Jensen SOW, Munk-Jørgensen P, et al. Risk factors for conversion from unipolar psychotic depression to bipolar disorder. Bipolar Disord. 2014 Mar;16(2):180–9.

83. Woo YS, Shim IH, Wang H-R, Song HR, Jun T-Y, Bahk W-M. A diagnosis of bipolar spectrum disorder predicts diagnostic conversion from unipolar depression to bipolar disorder: a 5-year retrospective study. J Affect Disord. 2015 Mar 15;174:83–8.

84. Alavi F, Nakhaee N, Sabahi A. Diagnostic stability of psychiatric disorders in re-admitted psychiatric patients in Kerman, Iran. Glob J Health Sci. 2014 Jul 14;6(5):294–300.

85. Atwoli L, Ndambuki D, Owiti P, Manguro G, Omulimi N. Short-term diagnostic stability among re-admitted psychiatric in-patients in Eldoret, Kenya. African Journal of Psychiatry. 2012;15(2):114–8.

86. Barrigon ML, Cegla-Schvartzman F. Sex, Gender, and Suicidal Behavior. Curr Top Behav Neurosci. 2020;46:89–115.

87. Beller EM, Glasziou PP, Altman DG, Hopewell S, Bastian H, Chalmers I, et al. PRISMA for Abstracts: reporting systematic reviews in journal and conference abstracts. PLoS Med. 2013;10(4):e1001419.

88. Landis JR, Koch GG. The measurement of observer agreement for categorical data. Biometrics. 1977 Mar;33(1):159–74.

89. Pierce K, Gazestani VH, Bacon E, Barnes CC, Cha D, Nalabolu S, et al. Evaluation of the Diagnostic Stability of the Early Autism Spectrum Disorder Phenotype in the General Population Starting at 12 Months. JAMA Pediatr. 2019 01;173(6):578–87.

90. 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.

91. Carmona R, Cook BL, Baca-García E, Chavez L, Alvarez K, Iza M, et al. Comparison of Mental Health Treatment Adequacy and Costs in Public Hospitals in Boston and Madrid. J Behav Health Serv Res. 2019 Jul;46(3):464–74.

92. Wang PS, Demler O, Kessler RC. Adequacy of treatment for serious mental illness in the United States. Am J Public Health. 2002 Jan;92(1):92–8.

93. Benchimol EI, Smeeth L, Guttmann A, Harron K, Moher D, Petersen I, et al. The REporting of studies Conducted using Observational Routinely-collected health Data (RECORD) Statement. PLOS Medicine. 2015 Oct 6;12(10):e1001885.

94. Leiva J, Artes A, Baca E. Workshop on Personalized Medicine. 2011; Granada, Spain.

95. Liu Y-Y, Li S, Li F, Song L, Rehg JM. Efficient Learning of Continuous-Time Hidden Markov Models for Disease Progression. Adv Neural Inf Process Syst. 2015;28:3599–607.

96. Kijima M. Markov Processes for Stochastic Modeling. Chapman and Hall; 1997. 341 p.

97. Owens DGC, Johnstone EC. Precursors and prodromata of schizophrenia: findings from the Edinburgh High Risk Study and their literature context. Psychol Med. 2006 Nov;36(11):1501–14.

98. Reif A, Pfuhlmann B. Folie a deux versus genetically driven delusional disorder: case reports and nosological considerations. Compr Psychiatry. 2004 Apr;45(2):155–60.

99. Mouchet-Mages S, Gourevitch R, Loo H. [Folie a deux: update of an old concept regarding two cases]. Encephale. 2008;34:31–7.

100. Yung AR, Buckby JA, Cosgrave EM, Killackey EJ, Baker K, Cotton SM, et al. Association between psychotic experiences and depression in a clinical sample over 6 months. Schizophr Res. 2007 Mar;91(1–3):246–53.

101. Macleod J, Oakes R, Copello A, Crome I, Egger M, Hickman M, et al. Psychological and social sequelae of cannabis and other illicit drug use by young people: a systematic review of longitudinal, general population studies. Lancet. 2004 May 15;363(9421):1579–88.

102. Rubino IA, Frank E, Nanni RC, Pozzi D, Scalea TL di, Siracusano A. A Comparative Study of Axis I Antecedents before Age 18 of Unipolar Depression, Bipolar Disorder and Schizophrenia. PSP. 2009;42(5):325–32.

103. Forrester A, Owens DG, Johnstone EC. Diagnostic stability in subjects with multiple admissions for psychotic illness. Psychol Med. 2001 Jan;31(1):151–8.

104. Munk-Jørgensen P. The schizophrenia diagnosis in Denmark. A register-based investigation. Acta Psychiatr Scand. 1985 Sep;72(3):266–73.

105. Perälä J, Suvisaari J, Saarni SI, Kuoppasalmi K, Isometsä E, Pirkola S, et al. Lifetime prevalence of psychotic and bipolar I disorders in a general population. Arch Gen Psychiatry. 2007 Jan;64(1):19–28.

106. Maj M. Critique of the DSM-IV operational diagnostic criteria for schizophrenia. Br J Psychiatry. 1998 Jun;172:458, 460.

107. Häfner H, Maurer K, Löffler W, Riecher-Rössler A. The influence of age and sex on the onset and early course of schizophrenia. Br J Psychiatry. 1993 Jan;162:80–6.

108. Weinberger DR. Implications of normal brain development for the pathogenesis of schizophrenia. Arch Gen Psychiatry. 1987 Jul;44(7):660–9.

109. Angermeyer MC, Kühn L. Gender differences in age at onset of schizophrenia. An overview. Eur Arch Psychiatry Neurol Sci. 1988 Sep;237(6):351–64.

110. Häfner H, Riecher A, Maurer K, Löffler W, Munk-Jørgensen P, Strömgren E. How does gender influence age at first hospitalization for schizophrenia? A transnational case register study. Psychological Medicine. 1989 Nov;19(4):903–18.

111. Chen L, Selvendra A, Stewart A, Castle D. Risk factors in early and late onset schizophrenia. Compr Psychiatry. 2018 Jan;80:155–62.

112. Harris MJ, Jeste DV. Late-onset schizophrenia: an overview. Schizophr Bull. 1988;14(1):39–55.

113. Esterberg ML, Compton MT. The psychosis continuum and categorical versus dimensional diagnostic approaches. Curr Psychiatry Rep. 2009 Jun;11(3):179–84.

114. Wiles NJ, Zammit S, Bebbington P, Singleton N, Meltzer H, Lewis G. Self-reported psychotic symptoms in the general population: results from the longitudinal study of the British National Psychiatric Morbidity Survey. Br J Psychiatry. 2006 Jun;188:519–26.

115. Instituto Nacional de Estadistica. Estadística de variaciones residenciales. Serie 1998–2011 [Internet]. 2012. Available from: www.ine.es

116. Harvey PD, Heaton RK, Carpenter WT, Green MF, Gold JM, Schoenbaum M. Diagnosis of schizophrenia: consistency across information sources and stability of the condition. Schizophr Res. 2012 Sep;140(1–3):9–14.

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

118. 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 Dec;33(12):587–95.

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

120. 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 Apr;1(2):65–74.

121. Lopez-Castroman J, Leiva-Murillo JM, Cegla-Schvartzman F, Blasco-Fontecilla H, Garcia-Nieto R, Artes-Rodriguez A, et al. Onset of schizophrenia diagnoses in a large clinical cohort. Sci Rep. 2019 Jul 8;9(1):9865.

122. Hirschfeld RMA, 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 Feb;64(2):161–74.

123. 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 Aug;31(4):281–94.

124. Ghaemi SN, Sachs GS, Chiou AM, Pandurangi AK, Goodwin K. Is bipolar disorder still underdiagnosed? Are antidepressants overutilized? J Affect Disord. 1999 Mar;52(1–3):135–44.

125. 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.

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

127. Moreno C, Laje G, Blanco C, Jiang H, Schmidt AB, Olfson M. National trends in the outpatient diagnosis and treatment of bipolar disorder in youth. Arch Gen Psychiatry. 2007 Sep;64(9):1032–9.

APENDIX

Diagnostic Stability in Bipolar Disorder: A Narrative Review

Fanny B. Cegla-Schvartzman, MD, Santiago Ovejero, MD, Jorge López-Castroman, MD, PhD, and Enrique Baca-García, MD, PhD

Learning objectives: After participating in this activity, learners should be better able to:
- Evaluate diagnostic stability in bipolar disorder
- Analyze the factors contributing to diagnostic stability

Objective: Diagnostic stability is the degree to which a diagnosis remains unchanged during follow-up. It is an important measure of predictive validity in bipolar disorder (BD). In this study, we review the literature concerning diagnostic stability in BD, analyze the factors contributing to diagnostic stability, and describe the implications of diagnostic boundaries and diagnostic delay.

Methods: A comprehensive literature search of MEDLINE and EMBASE databases was conducted, including all studies published from 1980 to 2016, to evaluate the diagnostic stability of BD. Thirty-seven articles were included: 6 focusing mainly on BD, 18 on psychotic disorders, 10 on depression, and 3 on diagnostic stability in psychiatric disorders in general. Data analysis was performed in standardized fashion using a predefined form.

Results: Despite a high variability of the methodological approaches taken, an acceptable degree of diagnostic stability was found. The most common criteria for evaluating diagnostic stability were prospective consistency and retrospective consistency. The mean prospective and retrospective consistencies were 77.4% and 67.6%, respectively. A large majority of studies were performed in Europe or in North America (67.5%), compared to 21.6% in Asia and only 10.8% in Africa, Oceania, and South America. Extreme ages, female gender, psychotic symptoms, changes to treatment, substance abuse, and family history of affective disorder have been related to diagnostic instability.

Conclusions: Several factors appear to have a negative impact on the diagnostic stability, but the evidence is insufficient to draw any robust conclusions. Nevertheless, despite variable prospective and retrospective consistencies, the overall diagnostic stability is good. Standardized methods need to be used to obtain more accurate assessments of stability.

Keywords: bipolar disorder, diagnostic change, diagnostic stability, prospective consistency, retrospective consistency

According to the 2013 Global Burden of Disease study, the number of cases of bipolar disorder (BD) increased by 50% in the 20 years prior to publication, reaching 50 million cases and thus making this disease the sixteenth leading cause of years lived with disability in 2013.[1] BD can have a major impact on an individual's ability to function and is associated with long-term risk of suicide.[2]

Diagnosis of psychiatric disorders, including BD, relies principally on cross-sectional assessment of clinical presentation. BD is a chronic and severe mental disorder characterized by recurrent episodes of depressed, elevated, and mixed mood. Moreover, as it is considered a lifelong illness, BD is likely to remain, once diagnosed, stable over time.[3] In clinical practice, however, a significant variation in diagnosis is observed.[4]

From the Department of Psychiatry, Hospital Universitario Fundación Jiménez Díaz, Madrid (Drs. Cegla-Schvartzman, Ovejero, and Baca-García); University of Montpellier, Centre Hospitalier Universitaire de Nîmes, and French National Institute of Health and Medical Research (INSERM) Unit 1061, Montpellier (Dr. López-Castroman); Departments of Psychiatry, Universidad Autónoma (Madrid), Hospital Universitario Rey Juan Carlos (Móstoles), Hospital General de Villalba (Madrid), Hospital Universitario Infanta Elena (Valdemoro) (all Spain) (Dr. Baca-García); Centro de Investigación en Salud Mental (CIBERSAM), Carlos III Institute of Health, Madrid (Dr. Baca-García); Universidad Católica del Maule, Talca, Chile (Dr. Baca-García); Department of Psychiatry, Columbia University (Dr. Baca-García).

Supported, in part, by Instituto de Salud Carlos III grant nos. ISCIII PI13/02200 and PI16/01852, Delegación del Gobierno para el Plan Nacional de Drogas grant no. 20151073, and American Foundation for Suicide Prevention grant no. LSRG-1-005-16 (all Dr. Baca-García).

Correspondence: Enrique Baca-García, MD, PhD, Department of Psychiatry, Fundación Jiménez Díaz, Av. Reyes Católicos 2, 28040 Madrid, Spain. Email: ebaca@fjd.es

Supplemental digital contents are available for this article. Direct URL citations appear in the printed text and are provided in the HTML and PDF versions of this article on the journal's Web site (www.harvardreviewofpsychiatry.org).

CME *Harvard Review of Psychiatry* offers CME for readers who complete questions about featured articles. Questions can be accessed from the *Harvard Review of Psychiatry* website (www.harvardreviewofpsychiatry.org) by clicking the CME tab. Please read the featured article and then log into the website for this educational offering. If you are already online, click here to go directly to the CME page for further information.

DOI: 10.1097/HRP.0000000000000187

In other branches of medicine, diagnoses are often supported by the identification of the underlying biological processes, whereas in psychiatry, diagnoses are based primarily on clinical syndromes.[5] In the absence of objective biological symptomatology, stability over time is the best means of validating BD diagnoses, and it can also be used to predict the course of the disorder.[6] For this reason, diagnostic stability may be used as a basis for the therapeutic management of BD patients. Conversely, the absence of stability in a diagnosis may have serious implications.[7]

The importance of obtaining accurate diagnoses of BD is of paramount importance. Underdiagnosis leads to delayed or ineffective treatment,[8] and overdiagnosis may also have adverse personal and social consequences, including unnecessary exposure to the risks of medication, missed opportunities for treatment of other conditions, and effects on work participation.[9] The rate at which bipolar patients receive inaccurate diagnoses in mental health facilities is estimated to be as high as 20%–60%.[10–12]

Though the number of studies focusing on diagnostic stability has increased in recent years, there remains a paucity of this type of research. Most of the published studies on diagnostic stability are focused on the first episode of psychosis and particularly on the subsequent consistency of schizophrenia diagnoses.[13] Other studies have used a meta-analytic approach to calculate the interrater reliability of BD diagnosis, as in the rigorous work by Santelmann and colleagues.[14,15] Studies focusing on BD diagnostic stability over time are scarcer. Diagnostic validity in psychiatric disorders requires stability over time, and because of the potential treatment implications, changes in diagnosis are a major issue to be considered.[16,17] In the case of BD, the literature is inconsistent. Some studies suggest moderate to high levels of temporal diagnostic stability of BD.[17–20] The initial diagnosis of BD, however, is often problematic. Diagnostic delays of 8 to 10 years are common, and studies consisting of repeated longitudinal evaluations have called into question the stability of this diagnosis in actual practice.[21–24]

Regardless of the great clinical and research implications of the issue of diagnostic stability of BD, this remains an understudied area.[5] Descriptions of the transition from diagnosis of depression and psychotic episodes to BD are of great interest and can be carried out by measuring the diagnostic stability of BD arising from such diagnoses. The aim of this study is to review the literature concerning diagnostic stability in BD, discuss the factors contributing to diagnostic stability, and describe the implications of diagnostic boundaries and diagnostic delay.

METHODS

Literature Search Strategy

A comprehensive literature search was conducted using the MEDLINE and EMBASE databases to search for texts published from 1980 to 2016. The keywords used for the search included the following: BD, manic-depression, mania, or bipolar spectrum AND diagnostic stability, diagnostic change, diagnostic consistency, diagnostic shift, diagnostic concordance, diagnostic conversion, and diagnostic progression. This keyword search was limited to the title and abstracts. The abstracts of the retrieved articles were then checked by applying the eligibility criteria. The PRISMA guidelines were followed to report findings.[25]

Inclusion Criteria and Retrieved Articles

Articles were included if they were published in English or Spanish, provided that they included adolescent (over 15 years of age) and adult cases, described a diagnostic shift from depression or psychotic episodes to BD, and informed the assessment of diagnostic stability for BD. Our definition of BD was not overly inclusive; we required that studies evaluate BD, psychotic episodes, or major depression (which included visits to the emergency department, inpatient hospitalization, and outpatient visits). Samples containing pediatric patients or pregnant women were not included. One study with patients younger than 15 years was included due to the low number of such patients (n = 20) in the total sample (n = 69,792). No studies in which the diagnoses were not performed by specialized mental health personnel were included. Cross-sectional, retrospective, and prospective studies were included. Citations within identified articles were included as additional sources. We omitted studies that included fewer than ten patients, uncompleted studies, conference abstracts, doctoral theses, and data not published in peer-reviewed journals. The initial electronic search identified 2317 documents. After evaluating the abstracts, we selected 37 articles that met the criteria (Supplemental Figure 1, available online at http://links.lww.com/HRP/A74). When selecting the articles, we were aware of unavoidable biases such as selection bias (differences between baseline characteristics of the groups) and detection bias (differences in how outcomes are measured). Thirteen studies (35%) included adolescent patients (>15 years). The selected studies are summarized in Table 1. The extraction of the pertinent data was standardized using a checklist based on PRISMA guidelines. Variables related to diagnostic stability were searched according to definitions appearing below.

Operative Definitions

Diagnostic stability is the measure of the degree to which a diagnosis remains unchanged during follow-up, providing a longitudinal means of validating the baseline diagnosis. It is based on the agreement of diagnoses over time and is irrespective of cross-sectional diagnosis at a single point of follow-up.[30] In a pioneering article published in 1970, Robins and Guze[16] mention diagnostic stability as one of the necessary criteria to verify the presence of a psychiatric syndrome, and for the first time the authors established a relationship with the predictive value of psychiatric diagnoses.

Prospective consistency (conceptually similar to positive predictive value) is the proportion of subjects in a diagnostic

Table 1

Narrative Review of Studies Examining Diagnostic Stability in Bipolar Disorders

Authors & year	Sample size/country	Study design & follow-up time	Inclusion criteria	Instruments	Results (95% confidence interval)
Studies focused on bipolar disorder					
Weeke et al. (1984)[4]	3062 Denmark	Retrospective 7 years	At least one manic-depressive diagnosis at admission Reassessed in second admission	ICD-8	RC: 20% (18.6–21.4)
Chen et al. (1998)[18]	936 (>18 years) USA	Prospective 7 years	At least four admissions Initial and final diagnosis schizophrenia or BD	DSM-III-R	PC: 70% (69.9–72.9) RC: 60% (56.9–63.1)
Kessing et al. (2005)[26]	4116 Denmark	Retrospective 9 years	Manic episode or BD Ten contact periods of assessment	ICD-10	PC: 68.8% (67.1–70.7)
Baca-Garcia et al. (2007)[21]	1153 Spain	Prospective 12 years	One BD diagnosis in at least ten assessments Multiple settings	ICD-10	PC: 49% (46.1–51.9) RC: 38% (35.1–40.8)
Ruggero et al. (2010)[17]	195 USA	Prospective 10 years	First admission due to psychosis At least one BD diagnosis in four assessments Consensus diagnosis	DSM-IV	PC: 79.6% (72.1–87.2) RC: 74.8% (66.8–82.7)
Ratheesh et al. (2015)[27]	52 Australia	Prospective 1 year	Meet Bipolar at Risk criteria (15–25 years, subthreshold manic symptoms and subthreshold depression in combination with either cyclothymic features or family history of BD)	DSM-IV	Diagnostic change to BD: 7.7% (0.4–14.9)
Studies focused on psychotic episodes					
Jorgensen et al. (1988)[28]	1136 Denmark	Retrospective 2 years	First admission with psychoses	ICD-8	PC: 72.9% (68.5–77.4)
Marneros at al. (1991)[29]	355 Germany	Historical prospective 25.2 years (mean follow-up)	One of the following episodes at least once: schizophrenic, melancholic, manic, manic-depressive mixed, schizodepressive, schizomanic and schizomanic-depressive mixed	DSM-III	PC: 62% (29.0–96.0)
Hollister et al. (1992)[24]	162 USA	Retrospective 3 years	Patients admitted four or more times	DSM-III-R	PC: 52.2% (37.7–66.6)
Fennig et al. (1994)[30]	278 USA	Prospective 6 months	First-admission psychotic patients Consensus diagnosis Baseline and 6-month assessments	DSM-III-R	PC: 85.7% (75.7–94.7) RC: 81.9% (71.7–90.8)

Table 1

Continued

Authors & year	Sample size/country	Study design & follow-up time	Inclusion criteria	Instruments	Results (95% confidence interval)
Studies focused on psychotic episodes					
Daradkeh et al. (1997)[31]	107 United Arab Emirates	Retrospective 2 years	>1 admission due to psychoses	ICD-10	PC: 87% (73.2–100.7)
Amin et al. (1999)[32]	168 UK	Prospective 3 years	First-episode psychosis Consensus diagnosis	DSM-III-R ICD-10	DSM-III-R PC: 78% (61.4–95.1) ICD-10 PC: 91% (77.9–103.0)
Schwartz et al. (2000)[3]	547 USA	Prospective 2 years	First-admission psychotic patients Baseline, 6-month, and 24-month assessments	DSM-IV	PC: 83% (76.8–89.2) RC: 84.8% (78.8–90.8)
Amini et al. (2005)[33]	48 Iran	Prospective 1 year	First episode of psychosis	DSM-IV ICD-10	PC: 100% RC: 94.4% (83.9–105.0)
Baldwin et al. (2005)[34]	194 Ireland	Prospective 6 months	First episode of psychosis	DSM-IV	PC: 97% (90.2–103.1)
Rufino et al. (2005)[22]	59 Brazil	Prospective 15 months	First psychotic episode Emergency setting 12-month minimal follow-up	DSM-IV	PC: 22.6% (7.9–37.3)
Schimmelmann et al. (2005)[35]	492 Australia	Retrospective 18 months	First-episode psychosis admitted patients 18-month reassessment	DSM-IV	PC: 83.2% (75.9–90.5) RC: 89.2% (84.3–96.3)
Whitty et al. (2005)[6]	147 Ireland	Prospective 4 years	First episode of psychosis Reassessed after four years	DSM-IV	PC: 80% (62.5–97.5)
Chang et al. (2009)[20]	166 China	Prospective 5 years	Young people with first-episode psychosis Consensus diagnosis	ICD-10	PC: 100% RC: 73.1% (56.0–90.1)
Bromet et al. (2011)[36]	470 USA	Prospective 10 years	First admission due to psychosis Consensus diagnosis	DSM-III-R DSM-IV	PC: 69.5% (60.2–78.7) RC: 58.4% (49.3–67.5)
Kim et al. (2011)[37]	150 Korea	Retrospective 15 years	At least one relapsed psychotic episode with admission Consensus diagnosis	DSM-IV	PC: 86.8% (76.1–97.6) RC: 64.7% (51.1–77.8)
Salvatore et al. (2011)[38]	517 USA	Prospective 2 years	Patients hospitalized in a first psychotic episode	ICD-10	Mania with psychosis PC: 99% (97.0–100.9) Mixed affective episode PC: 94.9% (91.4–98.3)

(Continued on next page)

Table 1

Continued

Authors & year	Sample size/country	Study design & follow-up time	Inclusion criteria	Instruments	Results (95% confidence interval)
Studies focused on psychotic episodes					
Heslin et al. (2015)[39]	505 UK	Prospective 10 years	First episode of psychosis Reassessment ten years later Consensus diagnosis	ICD-10 DSM-IV-TR	PC: 76.4% (65.1–87.6) RC: 67.7% (56.1–79.4)
Heslin et al. (2016)[40]	360 UK	Prospective 10 years	First episode of psychosis Reassessment ten years later Consensus diagnosis	ICD-10	PC: 97.1% (93.2–101.0) Diagnostic change to BD: 9.7% (2.9–16.6)
Studies focused on depression					
Coryell et al. (1995)[41]	932 USA	Prospective 5–10 years	In- and outpatients treated for affective disorders Older than 17 years IQ >70	RDC	Diagnostic change to BD: 10.2% (7.2–13.3)
Angst et al. (2005)[42]	406 Switzerland	Prospective 26 years	Inpatients with mania, endogenous depression, endoreactive depression, manic-depressive disorder, affective disorder with psychotic features including schizoaffective disorder	ICD-9	Diagnostic change to BD: 39.2% (33.7–44.6)
Gilman et al. (2012)[43]	6214 USA	Prospective 3 years	Diagnosis of MDD in the National Epidemiologic Survey on Alcohol and Related Conditions	DSM-IV	Diagnostic change to BD: 3.9% (3.4–4.4)
Li et al. (2012)[44]	2 cohorts: 1485 (2000–07); 2459 (2003–07) Taiwan	Prospective 8 years	All adult patients from the Nationwide Health Insurance database diagnosed with MDD by psychiatrists	ICD-9	Diagnostic change to BD 2000–07: 10.0% (8.5–11.5) Diagnostic change to BD 2003–07: 12.1% (10.8–13.4)
Dudek et al. (2013)[45]	122 Poland	Retrospective 30 years	Age >18 years at onset First established diagnosis of depression	ICD-9 ICD-10	Diagnostic change to BD: 32.8% (24.4–41.1)
Østergaard et al. (2014)[46]	8588 Denmark	Prospective 12 years	Patients assigned a diagnosis of psychotic depression from data from Danish registers	ICD-10	Diagnostic change to BD: 7.1% (6.5–7.6)
James et al. (2015)[47]	69,792 UK	Retrospective 4–12 years	Clinical diagnosis of depression from a linked data set of English national Hospital Episode Statistics	ICD-10	Diagnostic change to BD: 5.65% (5.4–5.8)

(Continued on next page)

Table 1

Continued

Authors & year	Sample size/country	Study design & follow-up time	Inclusion criteria	Instruments	Results (95% confidence interval)
Studies focused on depression					
Nakamura et al. (2015)[48]	89 Japan	Retrospective 9 years	Patients who were hospitalized for severe depression, both with and without psychotic episodes	ICD-10	Diagnostic change to BD: 12.3% (5.5–19.1)
Woo et al. (2015)[49]	250 Korea	Retrospective 5 years	Medical records of patients with a diagnosis of MDD without prior history of mania or hypomania	DSM-IV	Diagnostic change to BD: 18.4% (13.5–23.2)
Bukh et al. (2016)[50]	301 Denmark	Prospective 5 years	First-episode depression	ICD-10	Diagnostic change to BD: 8.6% (5.5–11.8)
Studies focused on psychiatric disorders in general					
Tsuang et al. (1981)[23]	445 USA	Prospective 30–40 years	Patients admitted for schizophrenia and affective disorders Consensus diagnosis	Feighner criteria	Interview form PC: 56.0% (35.5%–75.5%) Notes PC: 80.4% (64.0% to 96.8%)
Atwoli et al. (2012)[19]	114 Kenya	Prospective Turnaround time = 35 days	All admissions with at least one previous psychiatric admission	DSM-IV-TR	PC: 91.4% (82.1–100.7) RC: 69.6% (56.2–82.9)
Alavi et al. (2014)[51]	485 Iran	Retrospective 12 years	All admissions with at least one previous psychiatric admission	DSM-IV-TR	PC: 71% (66.9–74.9) RC: 69.4% (65.4–73.6)

BD, bipolar disorder; DSM, *Diagnostic and Statistical Manual of Mental Disorders*; ICD, *International Classification of Diseases*; MDD, major depressive disorder; PC, prospective consistency; RC, retrospective consistency; RDC, Research Diagnostic Criteria.

category at first evaluation who received the same diagnosis at last evaluation. *Retrospective consistency* (conceptually similar to sensitivity) is the proportion of subjects in a diagnostic category at the last evaluation who were in that same category at baseline.[21,51,52] Cohen's kappa for interrater agreement measures the diagnostic agreement corrected by chance. Diagnostic agreement was interpreted according to Landis and Koch[53] (k < 0 absence of agreement, .10–.20 slight, .21–.40 fair, .41–.60 moderate, .61–.80 good, and .81–1 excellent).

RESULTS

In this review, we will briefly describe and compare the studies that have evaluated diagnostic shifts from depression or psychotic episodes to BD and that have informed the assessment of diagnostic stability for BD in the last 36 years (from 1980 to 2016). We selected 37 studies focused on diagnostic stability: 6 of them are mainly focused on BD,[4,17,18,21,26,27] 18 on psychotic disorders,[3,6,20,22,24,28–40] 10 on depression,[41–50] and 3 on diagnostic stability in psychiatric disorders in general.[19,23,51]

Most of the studies are prospective in design (n = 25; 67.5%), and 12 (32.4%) are retrospective. The main diagnostic criteria used were the *Diagnostic and Statistical Manual of Mental Disorders* (n = 17; 46%) and *International Classification of Diseases* (n = 15; 40.5%); three studies (8.1%) used both; one (2.7%) used the Research Diagnostic Criteria; and one (2.7%) used the Feighner criteria. Of the 37 included studies, 40 different assessment instruments were used to classify patients and symptoms. The most widely used criteria for diagnostic stability were prospective and retrospective consistency, which were used by 27 studies (73%). Less common measurement tools included the proportion of diagnostic change, used by 12 studies (32%), and Cohen's kappa for interrater agreement, used in 7 (19%).

Examining all 37 studies, we found a mean prospective consistency of 77.4% and a retrospective consistency of 67.6%. Because of the variability in samples sizes and in the relative importance of each group, we calculated the weighted mean for prospective and retrospective consistency in each diagnostic group, controlled by sample size (Table 2). By contrast, studies focusing on depression revealed a trend of diagnostic shifts to BD, with the consequence that prospective and retrospective consistencies cannot be assessed. Prospective and retrospective consistencies are summarized in Figure 1. A sizable majority of studies were performed in Europe or in North America (n = 25; 67.5%), whereas 21.6% (n = 8) were performed in Asia and 10.8% (n = 4) in Africa, Oceania, and South America. The sample sizes varied from 48 to 69,792 patients. Four studies (10.8%) had fewer than 100 patients; 25 (67.5%) had between 100 and 1000 patients; 7 (20%) had between 1000 and 10,000 patients; and 1 (2.7%) had more than a 10,000 patients. In describing the results of these studies, we focus on the most common criteria for diagnostic stability: prospective and retrospective consistency. A detailed summary can be found on Table 1.

Table 2

Weighted Means Controlled by Sample Size for Prospective and Retrospective Consistency in Each Diagnostic Group

Diagnostic group	Prospective consistency (95% CI)	Retrospective consistency (95% CI)
Bipolar disorders	65.7 (64.5–66.9)	32.9 (31.6–34.2)
Psychotic episodes	80.4 (79.4–81.4)	75.7 (74.1–77.3)
Psychiatric disorders in general	70.9 (68.6–73.2)	69.4 (65.8–73.1)

CI, confidence interval.

Studies Focused on BD

Six studies focused mainly on the diagnostic stability of BD throughout the evolution of the disease. These studies have highly variable consistencies, with an average retrospective consistency of 39.9% (95% confidence interval [CI], 38.7–41.3) and an average prospective consistency of 66.8% (95% CI, 65.4–68.1).

An analysis of the six studies reveals that two are retrospective studies, and the other four are prospective. The two retrospective studies (Weeke [1984][4] and Kessing [2005][26]) were conducted with large samples (n = 3062 and 4116, respectively) and for periods of seven and nine years, respectively. In both, the inclusion criteria required at least one manic-depressive diagnosis or one diagnosis of BD. In the study by Weeke,[4] 20% of the registry sample was retrospectively classified as bipolar. Kessing[26] had 56.2% BD diagnoses at first contact, with a 30% change during follow-up. The retrospective consistency varied from 20% to 30%. Kessing and colleagues[26] found that diagnostic delay was especially frequent among the young, though also in female patients.

Four studies (Chen [1998],[18] Baca-García [2007],[21] Ruggero [2010],[17] and Ratheesh [2015][27]) prospectively examined the stability of BD diagnoses. Two of these studies were conducted with large samples (Chen with n = 936 and Baca-García with n = 1153), whereas the populations included in Ruggero[17] (n = 195) and Ratheesh[27] (n = 52) were substantially smaller. Three of the studies lasted between 7 and 12 years, and only one lasted a year. The inclusion criteria were similar for three studies (at least one BD diagnosis or one manic-depressive diagnosis in several assessments), and one study evaluated the progression from other highly prevalent mental disorders to BD. The prospective consistency found in these studies was higher than the one found in retrospective studies focused on BD, with a difference varying between 49% (Baca-García) and 79.6% (Ruggero).

Ruggero and colleagues[17] found a high prospective consistency (79.6%). Two factors, however, may lead to inconsistent diagnoses among patients followed over several years:

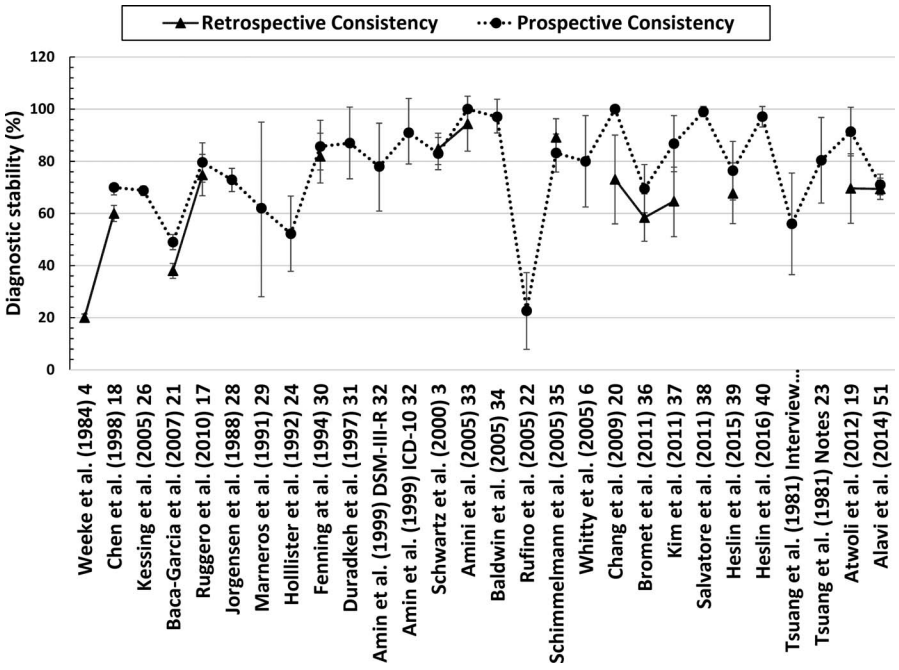


Figure 1. Diagnostic stability over time in studies including bipolar disorder patients. Error bars show confidence intervals at 95%.

(1) changes in the course of the underlying psychopathology, and (2) assessment errors. Even when using optimal assessment practices, a complex clinical presentation may make it difficult to accurately detect BD, resulting in increased odds of misdiagnosis over time. Baca-Garcia and colleagues[21] found a high prevalence of misdiagnosis and diagnostic shift from other psychiatric disorders to BD; additionally, temporal consistency was lower than in other studies (prospective and retrospective consistencies were 49% and 38%, respectively), and only 23% of patients received a BD diagnosis during >75% of the follow-up assessments.

Studies Focused on Psychotic Episodes

Eighteen studies examined the transition from psychotic episodes to BD, 13 of which are prospective, and 5 retrospective. Most of them focused on first psychotic episodes. The five retrospective studies were conducted by Kim[37] in 2011, Schimmelman[35] in 2005, Daradkeh[31] in 1997, Hollister[24] in 1992, and Jorgensen[28] in 1988. The samples varied from 107 to 1136 patients; four had follow-up periods of between 18 months and 3 years; and one had a follow-up period of 15 years.[37] For these retrospective studies, the average prospective and retrospective consistencies were 76.4% (95% CI, 73.4–79.5) and 76.9% (95% CI, 69.5–83.3%), respectively.

The other 13 studies were prospective in design, beginning with the first episode of psychosis. Two had a sample of 100 or fewer patients, and the other 11 ranged between 100 and 550. A study by Heslin[40] was excluded from analysis because it covered a similar population to the one described in another publication.[39] For these prospective studies, the average prospective and retrospective consistencies were 79.0% (95% CI, 75.9–82.2) and 76.7% (95% CI, 72.6–80.7), respectively. Bromet and colleagues[36] found that changes in symptoms and treatment were independent factors for diagnostic shifts.

Rufino and colleagues[22] found that psychotic disorder diagnoses made in an emergency room (ER) setting had high sensitivity but low specificity, and that BD showed the highest specificity. When the first diagnosis was made in an ER setting, kappa values were low when compared to the longitudinal follow-up diagnosis. By contrast, the agreement rates between diagnoses at ER discharge compared to follow-up diagnosis were satisfactory, with a kappa value of 0.57.

Studies Focused on Depression

Ten studies examined the evolution of the diagnosis from depression to BD. Six studies were prospective, and four retrospective. The four retrospective studies (Dudek [2013],[45] Nakamura [2015],[48] Woo [2015],[49] and James [2015][47]) were carried out with samples of 250, 89, 122, and 69,792 patients, with follow-up periods of 5, 9, 30, and 4 to 12 years, respectively. All reported the percentage of patients with diagnoses that shifted toward BD, which varied from 5.65%[47] to 32.8%.[45] James and colleagues[47] showed that 5.6% of depressed patients were eventually given a diagnosis of BD. The change to BD was more frequent in females, patients with higher age, and psychotic depression.

Within the prospective studies, sample sizes varied largely, ranging from 301[50] to 8588[46] patients. Most had a follow-up period between 3 and 12 years.[41–44,46,50] Angst's 2005 study[42] had a >20-year follow-up. All of them also reported

the percentage of patients that shifted their diagnoses toward BD, which varies from 3.9%[43] to 39.2%.[42]

Among patients with depression, a family history of affective disorder,[41,42] multiple depressive episodes,[42,46] psychotic symptoms,[41] treatment resistance,[44] and early age of onset[41,42,46] were found to increase the risk for conversion to BD. Other factors related to conversion to BD were living alone and receiving a disability pension.[46]

Boundaries of BD and Diagnostic Delay

Blacker and Tsuang[54] described BD as one of the most robust diagnostic entities in psychiatry, but some uncertainties are recognized. They identified a number of contested boundaries of the disorder, raising the question whether improved diagnostic criteria may be necessary for a number of research and clinical purposes.[54] The reasons that could demonstrate the distinct phenomenology of bipolar disorder include the following: its occurrence across history[55] and cultures,[56] its patterns of inheritance,[57,58] and its clear disturbance of physiologic function.[59]

The boundaries of BD remain unclear. Sara and colleagues[60] reported that diagnostic practice has changed in recent years, suggesting that diagnostic boundaries of BD are expanding in the Australian population, resulting in an 8% increase in BD prevalence—a trend that is accompanied by diagnostic change to other conditions. The change is observed in sensitive areas such as subclinical mood disorders, personality and affective disorders, and other mental pathologies such as psychotic events, substance use, and anxiety.

The time from disease onset to diagnosis is variable, though most authors suggest that significant delays are almost universal. Investigators from Harvard Medical School found that the mean time from the onset of BD symptoms to diagnosis was 9.6 years.[61] Hirschfeld and colleagues[62] found that at initial presentation, approximately 70% of patients were misdiagnosed and one-third received proper diagnosis only after 10 years.

DISCUSSION

Fifty million patients currently live with BD, a disorder ranked as the sixteenth leading cause of years lived with disability in 2013.[1] The diagnosis of BD relies principally on a cross-sectional evaluation of clinical features, though an accurate diagnosis can be reached only longitudinally. Diagnostic stability over time is the best source of evidence to validate the diagnosis of BD and predict its prognosis.[6] In addition, incorrect diagnoses have been reported in 20%–60% of cases,[10–12] and the retrospective diagnostic stability is reported as around 20%–38% in large series (>1000 patients).[4,21] Understanding the factors that determine diagnostic stability in BD may help improve prognosis.

To describe the factors related to diagnostic stability in BD, we performed a narrative review of the literature. The results were categorized based on the disorder preceding the BD diagnosis—mainly psychoses and affective disorders. Furthermore, data were grouped by study design for the sake of homogeneity and comparability.

There is no standard criterion to examine the diagnostic stability in BD; at least four different criteria are used, including prospective consistency, retrospective consistency, Cohen's kappa for interrater agreement, and proportion of diagnostic change. Taking into account only those that are used most frequently, here we observed very high variability. Symptom changes, the effects of treatments, the reinterpretation of clinical information, and the low reliability of diagnostic measures are some of the well-recognized factors of diagnostic stability.[3] Correct BD diagnoses take time to be established. Clinical polarities, social withdrawal, agitation, fluctuating symptoms, and concomitant substance abuse are commonly found in BD and may confound the clinical impression.[30,36] Due to the intrinsic nature of BD, there is no certainty of an accurate diagnosis.[18] Methodological changes are required to improve the way that diagnostic stability is measured. For example, a greater consensus on the use of data-collection instruments could improve the external validity of the results. In addition, other concepts such as *diagnostic reliability* have been used and could add value, bringing higher validity from a methodological and clinical perspective; interrater reliability is relevant because it measures the degree of agreement between two psychiatrists on the same diagnosis for the same patient.[14,15]

It is worth noting the large number of scales and the different diagnostic criteria used by the researchers, as this factor increased the variability of results and limited the ability to perform comparisons between studies. Since the methodological approaches, samples sizes, number of longitudinal evaluations, and population types for the 37 studies selected are so diverse, they cannot be compared directly. Additionally, the geographical distribution of the studies illustrates the importance of studying this topic outside Europe and North America to avoid cultural biases when interpreting results.

One of the greatest difficulties to be considered in prospective studies is the loss to follow-up, which alters prospective consistency. A low frequency of BD can be explained by transiently diagnosed patients with major depressive disorder who may evolve to BD.[63] The rate of BD misdiagnosis has been estimated at around 80% in a community sample[64] and 40% in sample of inpatients.[65]

In an attempt to assess the overall diagnostic stability for BD, we have calculated average prospective and retrospective consistencies, when available. We found a prospective consistency of between 75% and 80% and a retrospective consistency between 65% and 70%, which can be considered good diagnostic stability. This is consistent with two large meta-analyses by Santelmann and colleagues, which demonstrated that the reliability of BD diagnoses was good (kappa = .77; 95% CI, 0.73–0.82)[15] and excellent (kappa = 0.82; 95% CI, 0.77–0.86)[14] according to Landis and Koch's conventions for kappa values. Our assertion concerning consistency is biased, however, by the variability of the follow-up

times, which ranged from 35 days to 30 years (mean = 15.2 ± 6.7 years), methodological differences, and different clinical situations (selection and detection bias). Studies focused on patients with BD presented prospective consistencies between 50% and 80%, whereas studies focusing on psychotic episodes show prospective consistencies that reached 100%. This high diagnostic stability is derived from prospective studies with very small samples. When the studies with samples greater than 1000 patients were analyzed, a prospective consistency of around 65%–70% was found.

The fact that prospective stability is higher than retrospective stability is a reflection of daily clinical practice, where multiple manifestations or criteria sets are needed for the diagnosis of BD, resulting in delayed therapeutic interventions. Improving the detection of BD is of utmost importance for the clinician.

Recognizing the factors of diagnostic stability (and change) is crucial, as it could aid in understanding why diagnoses shift, thus providing clinicians with information on when to be alert to possible changes and consequently when to modify treatment.[39] In the 37 studies analyzed, we found several factors that influenced diagnostic stability and diagnostic delay. Younger age and female gender were risk factors for longer diagnostic delays[26] and for diagnostic changes from unipolar depression to BD.[47] Psychotic symptoms, changes in treatment, family history of affective disorder, early and older age of onset were also related to diagnosis instability. Another factor that should be mentioned is the presence of substance abuse. In several studies[26,27,51] it was reported as a modifier of clinical presentation, therefore leading to a misdiagnosis. Substance-related disorders are highly prevalent among psychiatric patients,[51] and in BD the rate is as high as 42%.[66]

In the light of our results, we suggest a high-risk clinical profile for diagnostic instability characterized by female gender, psychotic symptoms, changes in treatment, family history of affective disorder, and early or late age of onset.

The main limitation of this review concerns the selective use of assessment instruments and the criteria for diagnostic stability used on studies included. This means that the results are not comparable. As with all narrative reviews, some works may not have been included, because the search terms did not match any in the description or MESH terms, and causal analysis or meta-analytic assessment could not be performed.

Although a high diagnostic stability for BD has been reported in the literature, it is important to highlight that in daily clinical practice, the intervals between affective episodes is not yet completely understood, especially in cases of borderline personality disorder and type II BD. Additional work is needed to enhance diagnostic stability and to develop tools to quantify the risk of instability.

CONCLUSIONS

Based on our narrative review, we may tentatively draw some preliminary conclusions:

1. Diagnostic stability in BD is an understudied area, even though it represents the best means of validating the diagnosis of BD.
2. Improving the diagnosis of BD will result in early therapeutic interventions.
3. We propose a high-risk clinical profile for diagnostic instability: female gender, psychotic symptoms, changes in treatment, family history of affective disorder, and early or late age of onset of psychiatric manifestations.
4. High diagnostic stability for BD is frequently reported; daily practice is limited, however, in its ability to account for what occurs between affective episodes; additional research is needed.
5. To assess diagnostic changes, methods with higher methodological validity, such as Cohen's kappa for interrater agreement, should be used.

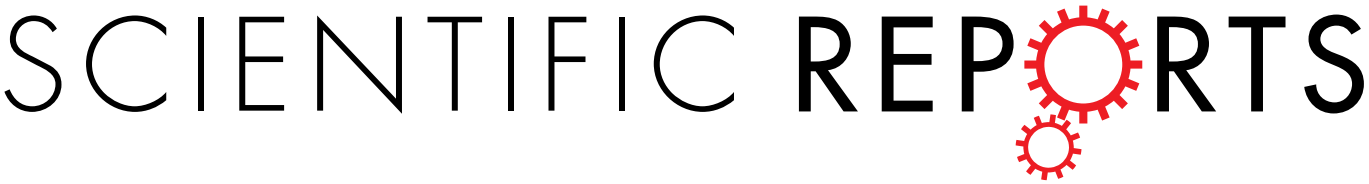
Declaration of interest: The authors report no conflicts of interest. The authors alone are responsible for the content and writing of the article.

REFERENCES

1. Ferrari AJ, Stockings E, Khoo JP, et al. The prevalence and burden of bipolar disorder: findings from the Global Burden of Disease Study 2013. Bipolar Disord 2016;18:440–50.
2. Moore PJ, Little MA, McSharry PE, Geddes JR, Goodwin GM. Forecasting depression in bipolar disorder. IEEE Trans Biomed Eng 2012;59:2801–7.
3. Schwartz JE, Fennig S, Tanenberg-Karant M, et al. Congruence of diagnoses 2 years after a first-admission diagnosis of psychosis. Arch Gen Psychiatry 2000;57:593–600.
4. Weeke A. Admission pattern and diagnostic stability among unipolar and bipolar manic-depressive patients. Acta Psychiatr Scand 1984;70:603–13.
5. Chang W. Diagnostic stability of functional psychosis: a systematic review. Hong Kong J Psychiatry 2009;19:30–41.
6. Whitty P, Clarke M, McTigue O, et al. Diagnostic stability four years after a first episode of psychosis. Psychiatr Serv 2005;56:1084–8.
7. Lopez-Castroman J. Errors, evolution and stability of bipolar disorder diagnosis. In: Plunkett JM, ed. Bipolar disorder: causes, diagnosis and treatment. New York: Nova Science, 2011.
8. Parker G. Borderline personality disorder and bipolar disorder: commentary on Paris and Black. J Nerv Ment Dis 2015;203:13–4.
9. Zimmerman M. Is underdiagnosis the main pitfall in diagnosing bipolar disorder? No. BMJ 2010;340:c855.
10. Rybakowski JK, Suwalska A, Lojko D, Rymaszewska J, Kiejna A. Bipolar mood disorders among Polish psychiatric outpatients treated for major depression. J Affect Disord 2005;84:141–7.
11. Akiskal HS, Akiskal KK, Lancrenon S, et al. Validating the bipolar spectrum in the French National EPIDEP Study: overview of the phenomenology and relative prevalence of its clinical prototypes. J Affect Disord 2006;96:197–205.
12. Kim B, Wang HR, Son JI, Kim CY, Joo YH. Bipolarity in depressive patients without histories of diagnosis of bipolar disorder and the use of the Mood Disorder Questionnaire for detecting bipolarity. Compr Psychiatry 2008;49:469–75.
13. Bromet EJ, Naz B, Fochtmann LJ, Carlson GA, Tanenberg-Karant M. Long-term diagnostic stability and outcome in recent first-episode cohort studies of schizophrenia. Schizophr Bull 2005;31:639–49.

14. Santelmann H, Franklin J, Busshoff J, Baethge C. Interrater reliability of schizoaffective disorder compared with schizophrenia, bipolar disorder, and unipolar depression—systematic review and meta-analysis. Schizophr Res 2016;176:357–63.
15. Santelmann H, Franklin J, Busshoff J, Baethge C. Test-retest reliability of schizoaffective disorder compared with schizophrenia, bipolar disorder, and unipolar depression—a systematic review and meta-analysis. Bipolar Disord 2015;17:753–68.
16. Robins E, Guze SB. Establishment of diagnostic validity in psychiatric illness: its application to schizophrenia. Am J Psychiatry 1970;126:983–7.
17. Ruggero CJ, Carlson GA, Kotov R, Bromet EJ. Ten-year diagnostic consistency of bipolar disorder in a first-admission sample. Bipolar Disord 2010;12:21–31.
18. Chen YR, Swann AC, Johnson BA. Stability of diagnosis in bipolar disorder. J Nerv Ment Dis 1998;186:17–23.
19. Atwoli L, Ndambuki D, Owiti P, Manguro G, Omulimi N. Short-term diagnostic stability among re-admitted psychiatric in-patients in Eldoret, Kenya. Afr J Psychiatry (Johannesbg) 2012;15:114–8.
20. Chang WC, Pang SL, Chung DW, Chan SS. Five-year stability of ICD-10 diagnoses among Chinese patients presented with first-episode psychosis in Hong Kong. Schizophr Res 2009;115: 351–7.
21. Baca-Garcia E, Perez-Rodriguez MM, Basurte-Villamor I, et al. Diagnostic stability and evolution of bipolar disorder in clinical practice: a prospective cohort study. Acta Psychiatr Scand 2007; 115:473–80.
22. Rufino AC, Uchida RR, Vilela JA, Marques JM, Zuardi AW, Del-Ben CM. Stability of the diagnosis of first-episode psychosis made in an emergency setting. Gen Hosp Psychiatry 2005;27:189–93.
23. 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;38:535–9.
24. Hollister LE. Stability of psychiatric diagnoses among acutely ill patients. Psychopathology 1992;25:204–8.
25. Beller EM, Glasziou PP, Altman DG, et al. PRISMA for Abstracts: reporting systematic reviews in journal and conference abstracts. PLoS Med 2013;10:e1001419.
26. Kessing LV. Diagnostic stability in bipolar disorder in clinical practise as according to ICD-10. J Affect Disord 2005;85: 293–9.
27. Ratheesh A, Cotton SM, Betts JK, et al. Prospective progression from high-prevalence disorders to bipolar disorder: exploring characteristics of pre-illness stages. J Affect Disord 2015;183: 45–8.
28. Jorgensen P, Mortensen PB. Admission pattern and diagnostic stability of patients with functional psychoses in Denmark during a two-year observation period. Acta Psychiatr Scand 1988; 78:361–5.
29. 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;241:187–92.
30. 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.
31. Daradkeh T, El-Rufaie O, Younis Y, Ghubash R. The diagnostic stability of ICD-10 psychiatric diagnoses in clinical practice. Eur Psychiatry 1997;12:136–9.
32. 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.
33. Amini H, Alaghband-rad J, Omid A, et al. Diagnostic stability in patients with first-episode psychosis. Australas Psychiatry 2005; 13:388–92.
34. Baldwin P, Browne D, Scully PJ, et al. Epidemiology of first-episode psychosis: illustrating the challenges across diagnostic boundaries through the Cavan-Monaghan study at 8 years. Schizophr Bull 2005;31:624–38.
35. 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;66:1239–46.
36. Bromet EJ, Kotov R, Fochtmann LJ, et al. Diagnostic shifts during the decade following first admission for psychosis. Am J Psychiatry 2011;168:1186–94.
37. Kim JS, Baek JH, Choi JS, Lee D, Kwon JS, Hong KS. Diagnostic stability of first-episode psychosis and predictors of diagnostic shift from non-affective psychosis to bipolar disorder: a retrospective evaluation after recurrence. Psychiatry Res 2011; 188:29–33.
38. Salvatore P, Baldessarini RJ, Tohen M, et al. McLean-Harvard International First-Episode Project: two-year stability of ICD-10 diagnoses in 500 first-episode psychotic disorder patients. J Clin Psychiatry 2011;72:183–93.
39. Heslin M, Lomas B, Lappin JM, et al. Diagnostic change 10 years after a first episode of psychosis. Psychol Med 2015;45: 2757–69.
40. Heslin M, Lappin JM, Donoghue K, Keller MB, Leon AC, Akiskal HS. Ten-year outcomes in first episode psychotic major depression patients compared with schizophrenia and bipolar patients. Schizophr Res 2016;176:417–22.
41. Coryell W, Endicott J, Maser JD, Keller MB, Leon AC, Akiskal HS. Long-term stability of polarity distinctions in the affective disorders. Am J Psychiatry 1995;152:385–90.
42. Angst J, Sellaro R, Stassen HH, Gamma A. Diagnostic conversion from depression to bipolar disorders: results of a long-term prospective study of hospital admissions. J Affect Disord 2005; 84:149–57.
43. Gilman SE, Dupuy JM, Perlis RH. Risks for the transition from major depressive disorder to bipolar disorder in the National Epidemiologic Survey on Alcohol and Related Conditions. J Clin Psychiatry 2012;73:829–36.
44. Li CT, Bai YM, Huang YL, et al. Association between antidepressant resistance in unipolar depression and subsequent bipolar disorder: cohort study. Br J Psychiatry 2012;200:45–51.
45. Dudek D, Siwek M, Zielinska D, Jaeschke R, Rybakowski J. Diagnostic conversions from major depressive disorder into bipolar disorder in an outpatient setting: results of a retrospective chart review. J Affect Disord 2013;144:112–5.
46. Ostergaard SD, Straszek S, Petrides G, et al. Risk factors for conversion from unipolar psychotic depression to bipolar disorder. Bipolar Disord 2014;16:180–9.
47. James A, Wotton CJ, Duffy A, Hoang U, Goldacre M. Conversion from depression to bipolar disorder in a cohort of young people in England, 1999–2011: a national record linkage study. J Affect Disord 2015;185:123–8.
48. Nakamura K, Iga J, Matsumoto N, Ohmori T. Risk of bipolar disorder and psychotic features in patients initially hospitalised with severe depression. Acta Neuropsychiatr 2015;27:113–8.
49. Woo YS, Shim IH, Wang HR, Song HR, Jun TY, Bahk WM. A diagnosis of bipolar spectrum disorder predicts diagnostic conversion from unipolar depression to bipolar disorder: a 5-year retrospective study. J Affect Disord 2015;174:83–8.
50. Bukh JD, Andersen PK, Kessing LV. Rates and predictors of remission, recurrence and conversion to bipolar disorder after the first lifetime episode of depression—a prospective 5-year follow-up study. Psychol Med 2016;46:1151–61.
51. Alavi F, Nakhaee N, Sabahi A. Diagnostic stability of psychiatric disorders in re-admitted psychiatric patients in Kerman, Iran. Glob J Health Sci 2014;6:294–300.

52. Addington J, Chaves A, Addington D. Diagnostic stability over one year in first-episode psychosis. Schizophr Res 2006;86:71–5.
53. Landis JR, Koch GG. The measurement of observer agreement for categorical data. Biometrics 1977;33:159–74.
54. Blacker D, Tsuang MT. Contested boundaries of bipolar disorder and the limits of categorical diagnosis in psychiatry. Am J Psychiatry 1992;149:1473–83.
55. Berrios GE. The history of mental symptoms: descriptive psychopathology since the nineteenth century. Cambridge; New York: Cambridge University Press, 1996.
56. Weissman MM, Bland RC, Canino GJ, et al. Cross-national epidemiology of major depression and bipolar disorder. JAMA 1996;276:293–9.
57. Tsuang MT, Winokur G, 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.
58. 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.
59. Goodwin FK, Jamison KR, Ghaemi SN. Manic-depressive illness: bipolar disorders and recurrent depression. 2nd ed. New York: Oxford University Press, 2007.
60. Sara GE, Malhi GS. Trends in diagnosis of bipolar disorder: have the boundaries changed? Aust N Z J Psychiatry 2015;49:1021–8.
61. Baethge C, Tondo L, Bratti IM, et al. Prophylaxis latency and outcome in bipolar disorders. Can J Psychiatry 2003;48:449–57.
62. 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;64:161–74.
63. De Fruyt J, Demyttenaere K. Bipolar (spectrum) disorder and mood stabilization: standing at the crossroads? Psychother Psychosom 2007;76:77–88.
64. Hirschfeld RM, Calabrese JR, Weissman MM, et al. Screening for bipolar disorder in the community. J Clin Psychiatry 2003;64:53–9.
65. Ghaemi SN, Sachs GS, Chiou AM, Pandurangi AK, Goodwin K. Is bipolar disorder still underdiagnosed? Are antidepressants overutilized? J Affect Disord 1999;52:135–44.
66. Hunt GE, Malhi GS, Cleary M, Lai HM, Sitharthan T. Prevalence of comorbid bipolar and substance use disorders in clinical settings, 1990–2015: systematic review and meta-analysis. J Affect Disord 2016;206:331–49.



OPEN

Onset of schizophrenia diagnoses in a large clinical cohort

Jorge Lopez-Castroman[1,2], José Miguel Leiva-Murillo[3], Fanny Cegla-Schvartzman[4], Hilario Blasco-Fontecilla[5], Rebeca Garcia-Nieto[4], Antonio Artes-Rodriguez[3], Consuelo Morant-Ginestar[6], Philippe Courtet[1,7], Carlos Blanco[8], Fuensanta Aroca[9] & Enrique Baca-García[4,8,10,11,12,13,14,15]

Received: 6 December 2018
Accepted: 21 June 2019
Published online: 08 July 2019

We aimed to describe the diagnostic patterns preceding and following the onset of schizophrenia diagnoses in outpatient clinics. A large clinical sample of 26,163 patients with a diagnosis of schizophrenia in at least one outpatient visit was investigated. We applied a Continuous Time Hidden Markov Model to describe the probability of transition from other diagnoses to schizophrenia considering time proximity. Although the most frequent diagnoses before schizophrenia were anxiety and mood disorders, direct transitions to schizophrenia usually came from psychotic-spectrum disorders. The initial diagnosis of schizophrenia was not likely to change for two of every three patients if it was confirmed some months after its onset. When not confirmed, the most frequent alternative diagnoses were personality, affective or non-schizophrenia psychotic disorders. Misdiagnosis or comorbidity with affective, anxiety and personality disorders are frequent before and after the diagnosis of schizophrenia. Our findings give partial support to a dimensional view of schizophrenia and emphasize the need for longitudinal assessment.

Schizophrenia and schizoaffective disorder have a lifetime prevalence of about 1% and are among the leading causes of disability[1–5]. Due to its early onset and its deteriorating course, schizophrenia causes an immense economic burden. A large majority of patients with schizophrenia are unemployed, and impairments in functioning across social, vocational and residential domains remain severe even during periods of remission from active psychosis[6–10], resulting in costs estimated at $62 billion in the US in 2002 and increasing three times its value a decade later. The economic burden of schizophrenia is estimated at $155.7 billion for 2013 including excess direct health care costs of $37.7 billion[11,12]. Approximately 80% of patients with schizophrenia relapse within 5 years of the first episode and many do not fully recover[13]. Moreover, the cognitive deficits and lack of insight that are core features of schizophrenia impair the patients' ability to recognize their disability or the symptoms that precede a relapse[14]. The disorder is therefore a permanent source of anguish for patients and their families and is associated with an increased risk for suicide and general medical conditions[3,15]. Although significant advances have been made in the understanding of the illness during the last 130 years (since Kraepelin's original classification in 1887) the underpinnings of its etiology remain unknown. There are no biomarkers that could be regularly used in clinical practice for the diagnosis of schizophrenia[16,17]. Schizophrenia is also a very heterogeneous illness, as originally described by Bleuler in 1911 and later confirmed in several studies[18,19]. Thus, longitudinal validation provides one of the most direct evidences of diagnostic validity[20].

[1]INSERM Unit 1061 and University of Montpellier, Montpellier, France. [2]Nimes University Hospital, Nimes, France. [3]Department of Signal Theory and Communications, Telecommunication Engineering Faculty, Carlos III University, Edif. Torres Quevedo, Avda. Universidad 30, 28911, Leganes, Madrid, Spain. [4]Department of Psychiatry Hospital Universitario Fundación Jiménez Díaz, Madrid, Spain. [5]Department of Psychiatry, IDIPHIPSA-Puerta de Hierro University Hospital, CIBERSAM, UAM, Madrid, Spain. [6]Department of Mental Health, Madrid Regional Health Council, Madrid, Spain. [7]Department of Emergency Psychiatry and Post-Acute Care, CHRU Montpellier F-34000, Montpellier, France. [8]Department of Psychiatry at the New York State Psychiatric Institute and Columbia University, New York, USA. [9]Instituto de Matemáticas, UNAM, Mexico city, Mexico. [10]Department of Psychiatry, Universidad Autónoma, Madrid, Spain. [11]Department of Psychiatry, Hospital Universitario Rey Juan Carlos, Móstoles, Spain. [12]Department of Psychiatry, Hospital General de Villalba, Madrid, Spain. [13]Department of Psychiatry, Hospital Universitario Infanta Elena, Valdemoro, Spain. [14]CIBERSAM (Centro de Investigación en Salud Mental), Carlos III Institute of Health, Madrid, Spain. [15]Universidad Católica del Maule, Talca, Chile. Jorge López-Castroman, José Miguel Leiva-Murillo and Fanny Cegla-Schvartzman contributed equally. Correspondence and requests for materials should be addressed to J.L.-C. (email: jorgecastroman@gmail.com)

Several studies have investigated the long-term diagnostic stability of schizophrenia. Meta-analytical evidence shows high prospective diagnostic stability in schizophrenia spectrum, but most of the literature was focused on small samples of first-episode psychosis or individuals at high-risk of schizophrenia[21–24]. There is also limited knowledge of the different diagnoses received by the patients around the onset of schizophrenia. Based on two epidemiological studies, An der Heiden *et al.* (2000) reported that the first psychotic episode in schizophrenia was preceded in 75% of patients by an average of 5 years of prodromic symptoms, usually negative and depressive symptoms[25,26]. Other authors have also noted the presence of prior behavioral and affective abnormalities in patients with schizophrenia[27,28]. Likewise, prodromal samples studied prospectively were frequently diagnosed with depression, anxiety or substance use disorders, particularly cannabis, before making a transition to psychosis[29–31].

On the other hand, the current scientific paradigm is challenging the long-standing categorical perspective[32], in favor of a more dimensional conceptualization of psychoses[33]. According to the dimensional model, an extended phenotype of schizophrenia in the general population (vulnerability) would underlie the less common clinical phenotype of schizophrenia[34]. High levels of severity in different symptom dimensions would lead to clinical assessment, identification of correlated symptoms in other dimensions and finally, the diagnosis of schizophrenia. We might expect prior diagnoses corresponding to the different dimensions that have been proposed (negative, affective, psychotic, and cognitive) when studying a large sample of patients with schizophrenia.

Aims of The Study

To examine the diagnostic evolution of patients with schizophrenia before and after this diagnosis is made for the first time in public mental health facilities. Therefore, we investigated a large clinical sample to identify which diagnoses preceded and followed that of schizophrenia in those patients who remained in treatment. We hypothesized that the early onset of symptoms in different dimensions would be reflected in correspondingly different diagnostic pathways leading to the diagnosis of schizophrenia.

Method

Sample. The Madrid Psychiatric Registration System established in 1980 includes all individuals treated in public mental health centers in Madrid (Spain) until 2008. Public mental health centers are part of the National Health Services and provide free medical coverage to a catchment area of about 6,000,000 inhabitants and are funded through taxes. The Madrid Case Registry (*Registro Acumulativo de Casos de la Comunidad de Madrid*) is a naturalistic study of diagnostic stability and consistency over time of the mental disorders in the area. It includes information from all psychiatric visits to public outpatient mental health clinics in the province of Madrid, Spain between 1980 and 2008. The database includes sociodemographic information and clinical diagnoses. From 1980 to 1992, all diagnoses in the registry were coded according to the 9th Revision of the International Classification of Diseases (ICD-9). Since 1992, diagnoses have been assigned according to the 10th Revision of the ICD (ICD-10). ICD-9 codes were converted to ICD-10 codes using the guidelines published by the World Health Organization (WHO, 1993). The treating clinician (a psychiatrist or clinical psychologist) entered the diagnostic codes at every follow-up visit. A maximum of 2 diagnoses per patient per visit were recorded. Diagnostic counts in this study include comorbidities (for instance, a F1-F3 diagnosis would count both as F1 and F3). A detailed description of the database can be obtained elsewhere[35]. For this study we selected patients who met the following inclusion criteria: 1) diagnosed with schizophrenia (ICD-10 category F20) during at least one visit and 2) at least three registered diagnoses by psychiatrists or clinical psychologists (in three different visits to the outpatient clinics). In the resulting subset of patients (n = 26,163) with a diagnosis of schizophrenia during at least one outpatient visit; we examined the diagnoses given to these patients during previous visits to public mental health clinics (i.e., prior to the diagnosis of schizophrenia). All methods were performed in accordance with the relevant guidelines and regulations. The Institutional Review Board of "Hospital 12 de Octubre" and "Fundación Jiménez Diaz" approved this study.

Data analysis. The probability of maintaining or changing the diagnosis of schizophrenia in the outpatient mental health visits was calculated in the 48 months following the initial diagnosis of schizophrenia. Probabilities were computed considering all the assessments made in one-month time lapses after the initial diagnosis. For each month the total number of diagnoses was added and then divided proportionally according to their distribution. If a diagnosis was missing (due to a longer delay between visits), the diagnosis recorded in the previous month was used instead.

Continuous-time hidden Markov model of longitudinal diagnostic shifts. In order to build a graph describing the sequence of diagnoses over time, a statistical model is needed that incorporates: (i) the frequencies of the transitions among mental disorders, (ii) the time lags between consecutive psychiatric visits, and (iii) the diagnostic uncertainty due to missing information in some of the records. We used a novel technique based on a continuous-time hidden Markov model (CT-HMM) to build the graphical model[36]. This method is based on a Hidden Markov model (HMM), which posits that starting from the current state of a stochastic system or process, it is possible to establish a description of its future probability, assuming that the measures are performed regularly in time. Nevertheless, in clinical practice that is not the case due to irregular or missed visits. Taking into account these cases, CT-HMM incorporates that both changes between hidden states and the appearance of new features can occur at any time[37].

The model makes use of the following assumptions: (i) the different patients are instances of the same stochastic process to be modeled; (ii) the process is stationary, so that the intensity of the interaction between two diseases does not depend on the age of the subject; (iii) a patient stays in the same state (disease) until the time instant of the following medical claim; (iv) the Markov property holds, meaning that the future clinical history of

ICD-10 diagnoses	Frequency	Percentage
F20 Schizophrenia	11330	43.3
F4 Neurotic, stress-related and somatoform disorders	4516	17.3
F3 Mood disorders	3402	13.0
F22 Persistent delusional disorders	1068	4.1
F6 Disorders of adult personality and behavior	1034	3.9
F1 Mental and behavioral disorders due to psychoactive substance use	981	3.7
F23 Acute and transient psychotic disorders	776	3.0
F29 Unspecified nonorganic psychosis	656	2.5
F9 Behavioral and emotional disorders with onset usually occurring in childhood and adolescence	345	1.3
F25 Schizoaffective disorders	276	1.1
F0 Organic mental disorders	263	1.0
F7 Mental retardation	211	0.8
F5 Behavioral syndromes associated with physiological disturbances and physical factors	205	0.8
F8 Disorders of psychological development	145	0.6
Total	24465	96.3

Table 1. Most frequent diagnoses at initial assessment. Only diagnoses made at least in 100 visits are listed.

a patient only depends on his/her present state, and is independent of the past. In this model, each element q_{ij} in the matrix of parameters describes the strength of the relation between diseases *i* and *j*[38]. The time lags between clinical events are also considered, and are represented as the sum of each row's elements $q_i = \Sigma_j q_{ij}$, which is high if the average transition time between *i* and the next clinical event is short. Finally, records with incomplete or lost diagnoses are treated probabilistically as uncertain observations under some underlying "hidden" disorder. This uncertainty was reflected by the $b_{i(k)}$ parameters, which described the probability of a certain diagnosis *k* if the real underlying diagnosis was *i*.

Results

Sample description. About half of the patients in the sample were male (n = 13,941; 53.3%). Mean age at the first assessment was 37.6 years (SD = 15.5) and mean age at the first diagnosis of schizophrenia was 39.3 years (SD = 14.9). At the time of their first visit to the mental health centers the patients were generally single (n = 15,741; 62.5%), and half of them were living with their family of origin (50.5%). The patients generally had low educational attainment: 46.0% had completed less than fifth grade (n = 10,997). Only one out of every four patients was working (n = 5,964; 26.5%) while most other patients were unemployed (n = 4,820; 21.5%), homemakers (n = 3,587; 16.0%), disabled (n = 2712; 12.1%), or studying (n = 2325; 10.4%). The following variables presented missing data over 2%: marital status (n = 965), educational level (n = 2,277), and working status (n = 3,737).

Patients with schizophrenia diagnoses made 1,455,063 visits to mental health centers (mean ± SD = 96.6 ± 174.7). The total number of visits with non-schizophrenia diagnoses prior to schizophrenia was 279,245, with an average of 18.03 visits per patient (SD = 28.81).

Previous diagnoses. In the sample, 56.7% of individuals (14,883/26,163) had received another diagnosis prior to being diagnosed with schizophrenia. Table 1 describes the diagnoses received in the first outpatient assessment at the mental health centers. Table 2 details the most frequent diagnoses on a per-visit basis prior to the diagnosis of schizophrenia.

When comorbidities were included, the most frequent diagnostic categories prior to schizophrenia followed the order of diagnoses reported in Table 2: 'mood disorders' (F3; 31.3%), 'neurotic, stress-related and somatoform disorders' (F4; 21.2%), non-schizophrenia diagnoses within the category of 'schizophrenia, schizotypal and delusional disorders' (F2; 19.2%), and 'disorders of adult personality and behavior' (F6; 12.3%). With respect to psychotic disorders specifically, 7.2% of patients were previously diagnosed of persistent delusional disorders (F22), 5.0% of acute and transient psychotic disorders (F23), 3.7% of unspecified nonorganic psychosis (F29), and 3.1% of schizoaffective disorders (F25).

Transition to schizophrenia. We then examined the probability of progression of these diagnoses to schizophrenia, considering also time proximity (Fig. 1). As might be expected, the strongest associations were between the psychotic-spectrum diagnoses and schizophrenia. From the highest to the lowest probability, schizoaffective (F25) disorders, induced delusional disorders (F24), unspecified nonorganic psychosis (F29), acute and transient psychotic disorders (F23), persistent delusional disorders (F22), and schizotypal disorders (F21) were all connected with a subsequent diagnosis of schizophrenia. The transitions to schizophrenia were indirect in some cases, usually through other psychotic disorders (for instance, F24 to F22 to F20). A probabilistic link from unspecified or other affective disorders (F38 and F39) towards schizophrenia was also represented, but not from bipolar disorder or major depression. Patients with alcohol, cannabis and multiple drug use disorders (F11, F12 and F19) were consequently diagnosed schizophrenia, but a direct transition was particularly frequent for the fewer subjects with hallucinogen use disorders (F16). Some organic mental disorders appeared also in our model. The diagnosis of delirium not induced by psychoactive substances (F05) showed a high probability of direct transition into schizophrenia, while the less frequent organic amnesic syndrome (F04) usually progressed

ICD-10 diagnoses	Frequency	Percentage
F3 Mood disorders	77461	27.7
F4 Neurotic, stress-related and somatoform disorders	50468	18.1
F6 Disorders of adult personality and behavior	24257	8.7
F22 Persistent delusional disorders	17660	6.3
F23 Acute and transient psychotic disorders	11907	4.9
F1 Mental and behavioral disorders due to psychoactive substance use	11639	4.3
F29 Unspecified nonorganic psychosis	8632	4.2
F25 Schizoaffective disorders	7488	3.1
F0 Organic mental disorders	4661	2.7
F9 Behavioral and emotional disorders with onset usually occurring in childhood and adolescence	4563	1.7
F7 Mental retardation	3468	1.6
F3-F6	3147	1.2
F3-F4	2789	1.1
F5 Behavioral syndromes associated with physiological disturbances and physical factors	2762	1.0
F4-F6	2647	1.0
F8 Disorders of psychological development	1886	0.9
F1-F6	1528	0.7
F1-F3	1423	0.5
Total	252127	90.3

Table 2. Most frequent diagnoses until schizophrenia. Only diagnoses made at least in 1000 visits are listed.

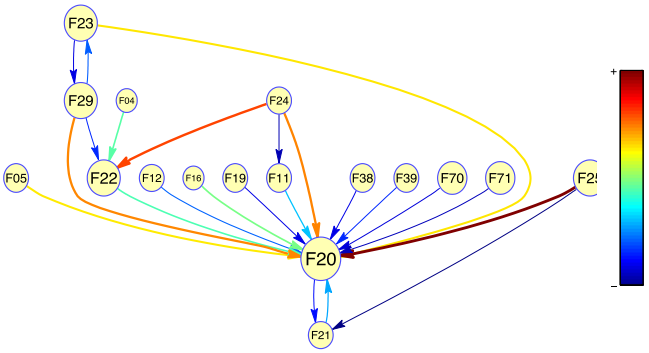


Figure 1. Probabilistic links of ICD-10 diagnoses converging into schizophrenia (F20). The size of the circles indicates the frequency of the diagnoses in our sample. The color and width of the arrows describe the strength of the interactions according to the model. F04/F05: Organic amnesic syndrome/Delirium, not induced by alcohol and other psychoactive substances; F11/12/F16/F19: Mental and behavioral disorders due to use of opioids/cannabinoids/hallucinogens/multiple drug use and use of other psychoactive substances; F21: Schizotypal disorder; F22: Persistent delusional disorders; F23: Acute and transient psychotic disorders; F24: Induced delusional disorder; F25: Schizoaffective disorders; F29: Unspecified nonorganic psychosis; F38: Other mood disorders; F39: Unspecified mood disorder; F70/F71: Mild/moderate mental retardation.

to persistent delusional disorders before. Mild and moderate cases of mental retardation (F70 and F71) were also directly linked to schizophrenia in our model. Of note, several diagnostic categories, such as anxiety (F4) and personality disorders (F6), were not represented in the graphical model.

Evolution/stability. Diagnostic shift from schizophrenia was more commonly toward the following diagnoses, represented by the average percentage in the first 48 months: personality disorders (F60: 4.2%), delusional disorders (F22: 3.7%), bipolar disorder (F31: 3.5%), persistent mood disorders (F34: 2.8%), acute and transient psychotic disorders (F23: 2.2%) or schizoaffective disorder (F25: 2.1%). However, the majority (64.5%) of the patients with an initial diagnosis of schizophrenia continued to receive the same diagnosis in subsequent assessments (Fig. 2). Patients who had a diagnostic shift from schizophrenia to a non-schizophrenia diagnosis did so generally in the first six months after the diagnosis of schizophrenia had been made. After that time interval the rates of each diagnostic category remained stable.

Discussion

This is a naturalistic study that describes which diagnoses the patients received before and after being diagnosed with schizophrenia for the first time. We found that although the most frequent prior diagnoses for those patients were anxiety (F4) and mood disorders (F3), direct transitions to schizophrenia usually came from psychotic-spectrum disorders. Furthermore, we also found that the initial diagnosis of schizophrenia is less likely to change if it is confirmed some months after its onset.

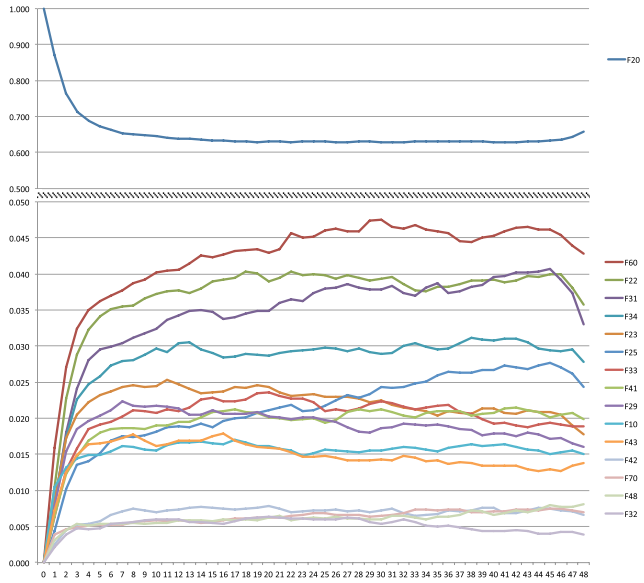


Figure 2. Diagnostic evolution of the first diagnosis of schizophrenia (F20) in the following 48 months. The upper section shows the probability (0.5–1) of maintaining the F20 diagnosis. The lower section shows the probability (0–0.05) of changing this diagnosis.

Indeed, psychotic symptoms have been noted as the best predictor of progression to schizophrenia among individuals at high risk[28]. Schizoaffective disorders are boundary diagnoses placed halfway between bipolar disorder and schizophrenia[16]. Schizotypal disorders and non-specified psychotic disorders are frequently used as a proxy for schizophrenia[23,33,39]. Therefore, it is not surprising that these diagnoses showed the most probable transitions to schizophrenia (Fig. 1). The uncommon diagnosis of induced delusional disorder also led to schizophrenia in the short term according to our model. The literature shows that a second diagnostic look sometimes reveals that a patient with supposedly induced delusional disorder shares the same genetically driven form of paranoid schizophrenia as an affected relative[40]. However, there are only a few case reports of this clinical entity[41], and more studies are needed.

In line with the high prevalence of prodromal depressive symptomatology[29], many patients with schizophrenia were previously diagnosed with anxiety or mood disorders. However, direct progression from affective or anxiety disorders did not appear as a probable transition to schizophrenia, implying that most patients received other diagnoses before eventually receiving a diagnosis of schizophrenia. In fact, anxiety and affective symptoms could be reactive expressions to prodromal and psychotic symptoms[42]. Personality disorders were the third most common diagnoses after mood and anxiety disorders, and likewise they were not directly followed by schizophrenia. This might reflect the uncertainty of the clinicians, who try to avoid direct transitions between conflicting diagnoses and perhaps some caution when it comes to diagnose a chronic and severe disorder such as schizophrenia. In other words, the high frequency of less severe previous diagnoses could be due to a conservative approach dealing with diagnostic uncertainty in early stages of a mental illness.

The direct transition between alcohol or drug use disorders represented in our model agrees with the mounting evidence of a causal influence of psychoactive substances in schizophrenia, particularly cannabis[43]. On the other hand, one study found that externalizing disorders were more frequent in the childhood and adolescence of patients with schizophrenia compared with those with bipolar disorder or depression[44]. However, behavioral and emotional disorders with onset usually occurring in childhood and adolescence (F9) were rarely registered before schizophrenia (1.3%) and the transition from F9 to schizophrenia was uncommon in our sample.

We also found that most patients retained the diagnosis of schizophrenia in the four years after its onset (65%). Although schizophrenia has a high diagnostic stability, according to previous works, prospective consistencies can vary from 52–100%[45–49]. In our sample approximately 35% of the patients diagnosed with schizophrenia, presented a change in their diagnosis (prospective consistency of 65%). This finding is consistent with the 70% temporal consistency that we have previously reported for schizophrenia in another sample, mostly assessed in outpatient facilities[48]. Overall, the results suggest that clinical assessment, which appears to be the most accurate diagnostic procedure for psychotic disorders[50], repeatedly maintains the diagnosis of schizophrenia once it is confirmed. Moreover, the time lapse for its confirmation agrees with the 6-month current diagnostic criteria in DSM-IV, which has been described as excessively restrictive as it might lead to the detection of chronic subjects with a worse prognosis[51–53]. When not confirmed, the most frequent alternative diagnoses are personality, affective or non-schizophrenia psychotic disorders.

Even though the diagnosis of schizophrenia is more common in adolescence and early adulthood[54,55], the risk of developing psychosis persists, and the age-incidence relationship seems to be altered by gender[56]. The evidence to date shows that men have an increased incidence at the end of the second decade and the beginning of the third, subsequently presenting lower rates, which remain thereafter. However, in women the first peak presents

later on, with a smaller second peak in middle age[55] and at 65 years, possibly a third peak[57]. In our cohort the observed age of onset fits on the upper limit reported in the literature (20–40 years at disease onset[58]). More than 50% of Kraepelin patients had onset of symptoms between ages 30 to 40, and over 20% between ages 40 to 50[59].

Novel approaches, supported by recent epidemiological and clinical research[60,61], try to consider the differential weight of dimensional traits of schizophrenia. In this study we expected to find a diagnostic pattern before schizophrenia that would correspond to the four symptom dimensions described in dimensional models[34]. Prior affective disorders and non-schizophrenia psychotic disorders were frequent in our sample, although the former seem to precede the latter in most cases. The other two symptom dimensions proposed in patients with schizophrenia (negative and cognitive) were rarely translated into specific diagnoses. This might suggest either that clinicians disregarded negative and cognitive symptoms as being part of a more severe clinical picture (affective or psychotic) or that an asymmetrical model of dimensions involved in schizophrenia would fit epidemiological data more closely. It should be noted that diagnostic classifications have evolved over time. The diagnostic criteria of ICD 9 are based on Schneider´s first rank symptoms and overlook negative and cognitive symptoms. This is less true for ICD 10 criteria.

This study examines clinical practice in Spain in real-world conditions, as it evaluates the follow-up of a large sample of patients with schizophrenia consecutively recruited in a 30-year interval. The representativeness of the study is enhanced by the free access and wide coverage of public medical care in our country. Moreover, since residential changes to other provinces are infrequent in Spain (<2% per year)[62] and visits to other mental health centers of the region would be included in the registry, our data is likely to reflect the real-world diagnostic pathways of the patients. However, some patients had probably received the first diagnosis of schizophrenia before being included in the registry or outside the system (e.g., in a private consultation). Thus, our results likely represent not only incidence cases but also prevalent cases that have increased the mean age at first diagnosis. However, other studies have also reported similar mean age at first diagnosis in schizophrenia. Despite this weakness, our results still provide an estimate of the shift patterns in the diagnosis of schizophrenia among patients who were retained in the public mental health system of Madrid. On the other hand, as structured diagnostic instruments were not used, it is unclear to what extent the clinical picture is changing immediately prior to the diagnosis of schizophrenia or in the early phases of illness, or whether clinicians vary in their ability to recognize the disorder. However, clinical evaluation seems so far to be the best diagnostic tool for schizophrenia[50], providing reliable and valid diagnoses when performed by mental health-specialists[63].

The evolution of diagnoses before and after that of schizophrenia indicates frequent initial misdiagnoses or comorbidity with affective, anxiety and personality disorders. Nevertheless, a diagnosis of schizophrenia is usually reached from psychotic-spectrum disorders or directly assigned, and once reached it is confirmed in the following six months for two of every three patients.

Data Availability

The datasets used and/or analyzed during the current study are available from the corresponding author on reasonable request.

References

1. Lehman, A. F. *et al*. Practice guideline for the treatment of patients with schizophrenia, second edition. *Am J Psychiatry* **161**, 1–56 (2004).
2. Harvey, P. D. *et al*. Functional milestones and clinician ratings of everyday functioning in people with schizophrenia: overlap between milestones and specificity of ratings. *J Psychiatr Res* **46**, 1546–1552 (2012).
3. Rossler, W., Salize, H. J., van Os, J. & Riecher-Rossler, A. Size of burden of schizophrenia and psychotic disorders. *Eur Neuropsychopharmacol* **15**, 399–409 (2005).
4. Janoutova, J. *et al*. Epidemiology and risk factors of schizophrenia. *Neuro Endocrinol Lett* **37**, 1–8 (2016).
5. Murray, C. & Lopez, A. World Health Organization, World Bank & Harvard School of Public Health. The global burden of disease. *Harvard School of Public Health* (1996).
6. Smith, K. Trillion-dollar brain drain. *Nature* **478**, 15 (2011).
7. Chen, R. *et al*. Assessment of functioning and disability in patients with schizophrenia using the WHO Disability Assessment Schedule 2.0 in a large-scale database. *Eur Arch Psychiatry Clin Neurosci* **268**, 65–75 (2018).
8. Sabbag, S. *et al*. Predictors of the accuracy of self assessment of everyday functioning in people with schizophrenia. *Schizophr Res* **137**, 190–195 (2012).
9. Lundin, L. & Flyckt, L. [Schizophrenia past and present–the perception of long term prognoses have changed]. *Lakartidningen* **112** (2015).
10. Harrow, M. *et al*. Followup of psychotic outpatients: dimensions of delusions and work functioning in schizophrenia. *Schizophr Bull* **30**, 147–161 (2004).
11. Wu, E. Q. *et al*. The economic burden of schizophrenia in the United States in 2002. *J Clin Psychiatry* **66**, 1122–1129 (2005).
12. Cloutier, M. *et al*. The Economic Burden of Schizophrenia in the United States in 2013. *J Clin Psychiatry* **77**, 764–771 (2016).
13. Palmier-Claus, J. E. *et al*. The feasibility and validity of ambulatory self-report of psychotic symptoms using a smartphone software application. *BMC Psychiatry* **12**, 172 (2012).
14. Insel, T. R. Rethinking schizophrenia. *Nature* **468**, 187–193 (2010).
15. Jobe, T. H. & Harrow, M. Long-term outcome of patients with schizophrenia: a review. *Can J Psychiatry* **50**, 892–900 (2005).
16. Nasrallah, H., Tandon, R. & Keshavan, M. Beyond the facts in schizophrenia: closing the gaps in diagnosis, pathophysiology, and treatment. *Epidemiol Psychiatr Sci* **20**, 317–327 (2011).
17. Birnbaum, R. & Weinberger, D. R. Genetic insights into the neurodevelopmental origins of schizophrenia. *Nat Rev Neurosci* **18**, 727–740 (2017).
18. Liddle, P. F. The symptoms of chronic schizophrenia. A re-examination of the positive-negative dichotomy. *Br J Psychiatry Aug* **151**, 145–51 (1987).
19. Schroder, J. *et al*. Cerebral metabolic activity correlates of subsyndromes in chronic schizophrenia. *Schizophrn Res. Mar* **19**(1), 41–53 (1996).
20. Robins, E. & Guze, S. B. Establishment of diagnostic validity in psychiatric illness: its application to schizophrenia. *American Journal of Psychiatry* **126**, 983–987 (1970).
21. Chang, W. C., Pang, S. L., Chung, D. W. & Chan, S. S. Five-year stability of ICD-10 diagnoses among Chinese patients presented with first-episode psychosis in Hong Kong. *Schizophr Res* **115**, 351–357 (2009).
22. Bromet, E. J. *et al*. Diagnostic shifts during the decade following first admission for psychosis. *Am J Psychiatry* **168**, 1186–1194 (2011).

23. Salvatore, P. *et al*. McLean-Harvard International First-Episode Project: two-year stability of ICD-10 diagnoses in 500 first-episode psychotic disorder patients. *J Clin Psychiatry* **72**, 183–193 (2011).
24. Fusar-Poli, P. *et al*. Diagnostic Stability of ICD/DSM First Episode Psychosis Diagnoses: Meta-analysis. *Schizophr Bull* **42**, 1395–1406 (2016).
25. der Heiden, W. & Hafner, H. The epidemiology of onset and course of schizophrenia. *Eur Arch Psychiatry Clin Neurosci* **250**, 292–303 (2000). an.
26. Hafner, H., Maurer, K., Trendler, G., an der Heiden, W. & Schmidt, M. The early course of schizophrenia and depression*. *Eur Arch Psychiatry Clin Neurosci* 255, (167–173 (2005).
27. Moller, P. & Husby, R. The initial prodrome in schizophrenia: searching for naturalistic core dimensions of experience and behavior. *Schizophr Bull* **26**, 217–232 (2000).
28. Miller, P. M., Lawrie, S. M., Byrne, M., Cosway, R. & Johnstone, E. C. Self-rated schizotypal cognitions, psychotic symptoms and the onset of schizophrenia in young people at high risk of schizophrenia. *Acta Psychiatr Scand* **105**, 341–345 (2002).
29. Addington, J. *et al*. At clinical high risk for psychosis: outcome for nonconverters. *Am J Psychiatry* **168**, 800–805 (2011).
30. Rosen, J. L., Miller, T. J., D'Andrea, J. T., McGlashan, T. H. & Woods, S. W. Comorbid diagnoses in patients meeting criteria for the schizophrenia prodrome. *Schizophr Res* **85**, 124–131 (2006).
31. Sara, G. E., Burgess, P. M., Malhi, G. S., Whiteford, H. A. & Hall, W. C. The impact of cannabis and stimulant disorders on diagnostic stability in psychosis. *J Clin Psychiatry* **75**, 349–356 (2014).
32. Murray, R. M. & Dutta, R. The right answer for the wrong reasons? *World Psychiatry* **6**, 93–94 (2007).
33. Carpenter, W. T. *et al*. The psychoses: cluster 3 of the proposed meta-structure for DSM-V and ICD-11. *Psychol Med* **39**, 2025–2042 (2009).
34. van Os, J., Kenis, G. & Rutten, B. P. The environment and schizophrenia. *Nature* **468**, 203–212 (2010).
35. Carballo, J. J. *et al*. Continuity of Depressive Disorders From Childhood and Adolescence to Adulthood: A Naturalistic Study in Community Mental Health Centers. *Prim Care Companion CNS Disord* **13** (2011).
36. Leiva, J., Artes, A. & Baca, E. In *NIPS Workshop on Personalized Medicine* (Granada (Spain), 2011).
37. Liu, Y., Li, S., Li, F., Song, L. & Rehg, J. M. Efficient Learning of Continuous-Time Hidden Markov Models for Disease Progression. *Adv Neural Inf Process Syst* **28**, 3599–3607 (2015).
38. Kijima, M. *Markov Processes for Stochastic Modeling*. (Chapman & Hall, 1997).
39. Owens, D. G. & Johnstone, E. C. Precursors and prodromata of schizophrenia: findings from the Edinburgh High Risk Study and their literature context. *Psychol Med* **36**, 1501–1514 (2006).
40. Reif, A. & Pfuhlmann, B. Folie a deux versus genetically driven delusional disorder: case reports and nosological considerations. *Compr Psychiatry* **45**, 155–160 (2004).
41. Mouchet-Mages, S., Gourevitch, R. & Loo, H. [Folie a deux: update of an old concept regarding two cases]. *Encephale* **34**, 31–37 (2008).
42. Yung, A. R. *et al*. Association between psychotic experiences and depression in a clinical sample over 6 months. *Schizophr Res* **91**, 246–253 (2007).
43. Macleod, J. *et al*. Psychological and social sequelae of cannabis and other illicit drug use by young people: a systematic review of longitudinal, general population studies. *Lancet* **363**, 1579–1588 (2004).
44. Rubino, I. A. *et al*. A comparative study of axis I antecedents before age 18 of unipolar depression, bipolar disorder and schizophrenia. *Psychopathology* **42**, 325–332 (2009).
45. Munk-Jorgensen, P. The schizophrenia diagnosis in Denmark. A register-based investigation. *Acta Psychiatr Scand* **72**, 266–273 (1985).
46. Forrester, A., Owens, D. G. & Johnstone, E. C. Diagnostic stability in subjects with multiple admissions for psychotic illness. *Psychol Med* **31**, 151–158 (2001).
47. Amini, H. *et al*. Diagnostic stability in patients with first-episode psychosis. *Australas Psychiatry* **13**, 388–392 (2005).
48. Baca-Garcia, E. *et al*. Diagnostic stability of psychiatric disorders in clinical practice. *Br J Psychiatry* **190**, 210–216 (2007).
49. Atwoli, L., Ndambuki, D., Owiti, P., Manguro, G. & Omulimi, N. Short-term diagnostic stability among re-admitted psychiatric in-patients in Eldoret, Kenya. *Afr J Psychiatry (Johannesbg)* **15**, 114–118 (2012).
50. Perala, J. *et al*. Lifetime Prevalence of Psychotic and Bipolar I Disorders in a General Population. *Archives General Psychiatry* **64**, 19–28 (2007).
51. Schwartz, J. E. *et al*. Congruence of diagnoses 2 years after a first-admission diagnosis of psychosis. *Archives General Psychiatry* **57**, 593–600 (2000).
52. Maj, M. Critique of the DSM-IV operational diagnostic criteria for schizophrenia. *Br J Psychiatry* **172**(458), 460 (1998).
53. Veen, N. *et al*. Diagnostic stability in a Dutch psychosis incidence cohort. *British Journal of Psychiatry* **185**, 460–464 (2004).
54. Weinberger, D. R. Implications of normal brain development for the pathogenesis of schizophrenia. *Arch Gen Psychiatry* **44**, 660–669 (1987).
55. Hafner, H., Maurer, K., Loffler, W. & Riecher-Rossler, A. The influence of age and sex on the onset and early course of schizophrenia. *Br J Psychiatry* **162**, 80–86 (1993).
56. Angermeyer, M. C. & Kuhn, L. Gender differences in age at onset of schizophrenia. *An overview. Eur Arch Psychiatry Neurol Sci* **237**, 351–364 (1988).
57. Hafner, H. *et al*. How does gender influence age at first hospitalization for schizophrenia? A transnational case register study. *Psychol Med* **19**, 903–918 (1989).
58. Chen, L., Selvendra, A., Stewart, A. & Castle, D. Risk factors in early and late onset schizophrenia. *Compr Psychiatry* **80**, 155–162 (2018).
59. Harris, M. J. & Jeste, D. V. Late-onset schizophrenia: an overview. *Schizophr Bull* **14**, 39–55 (1988).
60. Wiles, N. J. *et al*. Self-reported psychotic symptoms in the general population: results from the longitudinal study of the British National Psychiatric Morbidity Survey. *Br J Psychiatry* **188**, 519–526 (2006).
61. Esterberg, M. L. & Compton, M. T. The psychosis continuum and categorical versus dimensional diagnostic approaches. *Curr Psychiatry Rep* **11**, 179–184 (2009).
62. Instituto Nacional de Estadistica. *Estadística de variaciones residenciales. Serie 1998–2011*, www.ine.es (2012).
63. Harvey, P. D. *et al*. Diagnosis of schizophrenia: consistency across information sources and stability of the condition. *Schizophr Res* **140**, 9–14 (2012).

Acknowledgements

This research was supported by Madrid Regional Government (S2010/BMD-2422 AGES), Community of Madrid (I + D Activities in Biomedicine B2017/BMD-3740, AGES-CM 2CM), European Union Structural Funds. Dr. Lopez-Castroman was supported by the FondaMental Fondation for this study. Financial sources did not have any role on the design and conduction of the study, the collection, management, analysis, and interpretation of the data, or the preparation, review, or approval of the manuscript. The authors do not have any conflict of interest regarding the publication of this manuscript. Dr. Lopez-Castroman and Dr. Baca-Garcia had full access to all of the data in the study and take responsibility for the integrity of the data and the accuracy of the data analysis.

Author Contributions

J.L.C. and E.B.G. conceived and designed the study. J.L.C., J.M.L.M. and F.B.C.S. are responsible for written up and revision of the manuscript and managed the literature searches and analyses. A.A.R., C.M.G. and F.A. are responsible for data reduction, analysis and supervise the data extraction protocol. H.B.F., R.G.N., P.C. and C.B. critically review the manuscript and provide conceptual guidance on study implementation. E.B.G. critically review the manuscript, provide insightful recommendations to data reduction, analysis, interpretation and provide conceptual guidance on study implementation. J.L.C., J.M.L.M. and F.B.C.S. made equal contributions to this study. All authors revised the article critically. All authors read and approved the final manuscript. There is no one else who fulfils the criteria but has not been included as an author.

Additional Information

Competing Interests: The authors declare no competing interests.

Publisher's note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Original Research

Diagnostic Stability in Bipolar Disorder:
A Follow-up Study in 130,000 Patient-Years

Fanny Cegla-Schvartzman, MD[a]; Santiago Ovejero, MD, PhD[a]; Jorge López-Castroman, MD, PhD[b,c]; Nora Palomar-Ciria, MD[d]; Marta Migoya-Borja, MD[a]; Hugo Bello, PhD[e]; Gonzalo Martínez-Alés, MD[f]; and Enrique Baca-García, MD, PhD[a,b,g,h,i,j,k,l,*]

ABSTRACT

Objective: Diagnostic stability is the degree to which a diagnosis remains unchanged during time. Our main objective was to evaluate the diagnostic stability of bipolar disorder (BD) in psychiatric outpatient consultations and determine the socio-demographic variables influencing its stability.

Methods: The Cumulative Register of Cases of the Community of Madrid provided data on all outpatient visits conducted at Madrid's Community Mental Healthcare Centers between 1980–2009. Diagnoses were made according to *ICD-9*/*ICD-10* criteria. Two indices were measured: temporal consistency (maintenance of the diagnosis over time) and diagnostic constancy (presence of BD diagnosis in at least 75% of visits). κ coefficient measured the agreement between diagnoses in the first and last evaluations (prospective and retrospective consistency).

Results: 14,557 patients were diagnosed with BD for at least 1 evaluation and had at least 10 visits and 1 year of follow-up. At first evaluation, 3,988 patients were diagnosed with BD (prospective consistency 50.8%), and at last evaluation 5,396 patients were diagnosed with BD (retrospective consistency 37.5%). A total of 2,026 patients were diagnosed with BD at their first and last evaluations (prospective consistency 18.3%).

Conclusions: This longitudinal study conducted in community mental health centers reflects common diagnostic practices in outpatient settings over a 30-year period (130,000 patient-years). Delay of > 10 years was found to achieve diagnostic stability. Frequent diagnostic shifts were found in relation to BD, the most common being with other affective disorders. Anxiety was also a common misdiagnosis. Greater stability was associated with having been diagnosed after hospitalization, having an age at onset > 25 years, and having an age at diagnosis < 24 years.

J Clin Psychiatry 2021;82(00):20m13764

To cite: Cegla-Schvartzman M, Ovejero S, López-Castroman J, et al. Diagnostic stability in bipolar disorder: a follow-up study in 130,000 patient-years. *J Clin Psychiatry*. 2021;82(00):20m13764.

To share: https://doi.org/10.4088/JCP.20m13764

[a]Department of Psychiatry, Hospital Universitario Fundación Jiménez Díaz, Madrid, Spain
[b]Department of Psychiatry, Centre Hospitalier Universitaire de Nîmes
[c]INSERM Unit 1061 and University of Montpellier, Montpellier, France
[d]Psychiatry Service, Complejo Hospitalario de Soria, Soria, Spain
[e]Department of Applied Mathematics, Universidad de Valladolid, Duques de Soria Campus, Soria, Spain
[f]Department of Epidemiology, Columbia University Mailman School of Public Health, New York, New York
[g]Department of Psychiatry, Universidad Autónoma, Madrid, Spain
[h]Department of Psychiatry, Hospital Universitario Rey Juan Carlos, Móstoles, Spain
[i]Department of Psychiatry, Hospital General de Villalba, Madrid, Spain
[j]Department of Psychiatry, Hospital Universitario Infanta Elena, Valdemoro, Spain
[k]CIBERSAM (Centro de Investigación en Salud Mental), Carlos III Institute of Health, Madrid, Spain
[l]Universidad Católica del Maule, Talca, Chile
**Corresponding author:* Enrique Baca-García, MD, PhD, Department of Psychiatry, Hospital Universitario Fundación Jiménez Díaz, Avda. Reyes Católicos, 2. 28040. Madrid, Spain (ebaca@fjd.es).

Recent estimates suggest a global prevalence of bipolar disorder (BD) of around 45 million cases.[1] Because BD entails substantial functional disability,[2] it is considered a major contributor to the number of years lived with disability worldwide. In addition, BD is associated to an increased risk of all-cause mortality and, in particular, of suicide mortality.[3] As a result, previous research has highlighted the high cost and burden driven by BD in a variety of settings.[4,5]

A common limitation of psychiatric epidemiology studies is that psychiatric diagnoses are based on clinical assessments rather than biological measurements.[6] In the absence of an objective measurement that can serve as a gold standard, interobserver reliability and diagnostic stability over time are key components of the validity of psychiatric diagnoses. Diagnostic stability over time is presumed to be characteristic of psychiatric conditions with a tendency to chronicity and relapses over time, such as BD. However, stability varies markedly across chronic psychiatric disorders. For instance, schizophrenia has been found to be one of the most stable diagnoses.[7–10]

There is a paucity of studies focusing on the diagnostic stability of BD, despite the interest in mental health planning. Most recent studies suggest moderate to high levels of diagnostic stability for BD.[11–20] However, these studies are limited by technical difficulties. For instance, many studies have used only a few evaluation points, 2 or 3 at most, over limited follow-up periods,[21–24] raising concerns about the generalization to wider time periods and suggesting the need for studies including more evaluation points over longer follow-up periods.

The importance of diagnostic stability lies in the fact that those who are misdiagnosed or unstable are inadequately treated, which leads to more hospital admissions and also more suicides.[25,26]

This study estimated the real-world long-term clinical stability of BD diagnoses, using data from repeated outpatient visits, and explored which

Clinical Points

- This longitudinal study was conducted in community mental health centers, in a real-world scenario and in the general population, and reflects conditions in a daily practice over a 30-year period, including both outpatient consultations and hospitalization.
- Frequent diagnostic shifts were found in relation to BD, the most common being with other affective disorders. Anxiety was also a common misdiagnosis.
- Greater stability was observed if age at onset was > 25 years, BD diagnosis was made at age < 24 years, or diagnosis was made after hospitalization.

mental health diagnoses were the most common before and after receiving a BD diagnosis among individuals with no diagnostic stability.

METHODS

Study Setting, Sample, and Measurements

Using the Cumulative Register of Cases of the Community of Madrid, an electronic health care record that includes sociodemographic data and *International Classification of Diseases* (*ICD*) diagnostic codes for all outpatient psychiatric visits held between January 1980–December 2009 at Madrid's Community Mental Healthcare Centers (roughly 14% of Madrid's total population in 1996—5,022,289),[27] we selected all records of adults aged ≥ 18 years who (1) received a BD diagnosis in at least 1 visit and (2) undertook at least 10 visits over the study period (minimal adequacy of care, defined as having ≥ 4 outpatient visits in the last year and use of psychotropic medication, or ≥ 8 outpatient visits with or without a medication, a definition used in prior studies).[28,29] Out of a total population of 691,526 patients that were evaluated during 30 years, 14,557 met these inclusion criteria (Figure 1).

Madrid's Community Mental Healthcare Centers are part of the national health service of Spain, which has universal coverage, is financed by taxes, and has no direct cost for patients. The database includes the entire community of Madrid.

In this register, anonymity was ensured by a numerical coding system based on the assignment of a relational registration number. Accordingly, this study did not require participants' informed consent, in agreement with the Spanish law.[30] This study was overseen by the Institutional Review Board at Instituto de Investigación Sanitaria—Fundación Jiménez Díaz. The RECORD guidelines were followed to report findings.[31]

Diagnoses were made by board-certified psychiatrists in a variety of settings including both outpatient consultations and hospitalization, following guidelines in accordance with the 9th or 10th edition of the *International Classification of Diseases* (*ICD*), taking into account equivalence tables bridging both editions.[32] In addition, records included sociodemographic variables (Table 1).

Data Analysis

We used 2 indices of diagnostic stability:

1. Temporal consistency: the presence or absence of a particular disorder at first and last evaluations. Two indices were considered: prospective consistency and retrospective consistency. Of note, some recent papers use the term *diagnostic stability coefficient* as a synonymous of prospective consistency.[33] Using the broad *ICD-10* F1–F9 categories as diagnoses, we computed prospective consistency comparing diagnoses made at the initial evaluation with those made at the final one, and retrospective consistency comparing diagnoses made at the final evaluation with those made at the initial one.
2. Diagnostic constancy: Since prospective and retrospective consistency were based on only 2 evaluations, they often do not reflect diagnostic processes based on multiple evaluations characteristic of routine clinical practice.[34] We thus included a criterion according to which subjects who received diagnoses of BD in at least 75% of the evaluations were categorized as having a "stable BD diagnosis," since it is a common consensus measure that has been used in previous studies.[35]

We used χ^2 and Fisher exact test to test sociodemographic differences between people with stable and non-stable BD diagnosis. We then applied a multivariable logistic model to examine predictors of diagnostic shift including the significant variables of univariate analysis as covariates (gender, marital status, educational level, employment status, occupation, and type of cohabitation and background; described in Table 1), selecting the final model with a progressive elimination method (the likelihood ratio was used as criteria for model fit).

Survival analyses were used to estimate the time from the beginning of the follow-up to the first diagnosis of BD. Since the follow-up time contributes to the stability of the

Figure 1. Sample Selection

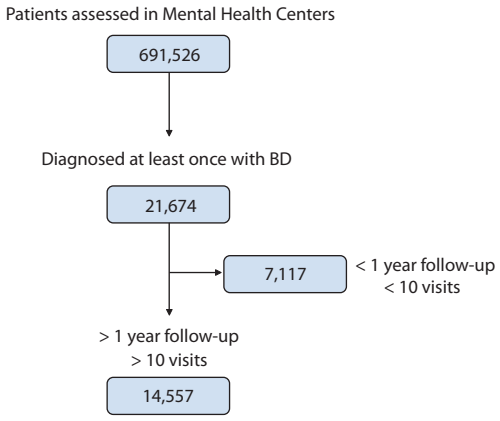


Table 1. Sociodemographic Data

	n	%
Sex		
Female	9,134	63.9
Male	5,161	36.1
Total	14,295	100
Data missing	2	0
Marital status		
Married	7,297	51
Divorced	320	2.2
Single	4,873	34.1
Widower	834	5.8
Separated	568	4
Data missing	405	2.8
Level of education		
Illiterate	365	2.6
No studies	1,458	10.2
Elementary	4,646	32.5
Middle school	2,505	17.5
High school	2,922	20.5
College degree	1,314	9.2
Other	161	1.1
Data missing	926	6.5
Employment status		
Military	18	0.1
Temporal incapacity to work	1,091	7.6
Permanent incapacity to work	271	1.9
Active	4,228	29.6
Looking for first job	197	1.4
Subsidized unemployment	411	2.9
Unsubsidized unemployment	820	5.7
Retirement	1,676	11.7
Rentier	40	0.3
Studying	769	5.4
Work at home	3,427	24
Data missing	1,349	9.4
Employment		
No job	4,549	31.8
Professionals and technicians	1,258	8.8
Management	120	0.8
Administrative	1,116	7.8
Commercial	456	3.2
Hotels and security services	1,385	9.7
Agriculture	124	0.9
Construction industry	763	5.3
Other	4,462	31.2
Armed forces	63	0.4
Data missing	1	0
Residential situation		
Other	586	4.1
Alone	1,216	8.5
Spouse	6,980	48.8
Couple	450	3.1
Family	2,518	17.6
Father only	125	0.9
Mother only	673	4.7
Children	927	6.5
Other family members	506	3.5
Institutionalized	186	1.3
Data missing	130	0.9

diagnosis, the survival predictors were analyzed with the Mantel-Cox model taking into account other covariates and the follow-up time.

RESULTS

Sample Description

A total of 14,557 patients were diagnosed with BD. These patients received 848,147 psychiatric and/or psychological consultations. The mean follow-up time for these patients was 3,295.9 days (standard deviation [SD] = 1,967.6 days), the mean number of visits was 58.3 (SD = 66.7), and the median was 38 visits. Sociodemographic data are shown in Table 1.

Prospective Consistency of Psychiatric Diagnoses

Our consistency comparisons included 15,082 diagnoses made at the initial evaluation and 15,507 at the final one. Figure 2 depicts retrospective diagnostic shifts.

The greatest prospective consistency was found among subjects diagnosed with mood/affective disorders (F3 category): 77.7% of patients diagnosed with F3 in the initial evaluation received a diagnosis under the same category at the final evaluation.

We also found a high prospective consistency in patients diagnosed with schizophrenia, schizotypal disorders, and delusional disorders (F2): 60% of these patients also received the same diagnosis. In contrast, patients diagnosed with mental and behavioral disorders due to the use of psychoactive substances (F1) had a low prospective consistency of 30.5%.

Prospective Consistency of BD Diagnoses

A total of 3,988 patients received a BD diagnosis in their first visit, 5,396 received it in their final visit, and 2,026 received it in both visits. Prospective and retrospective consistencies were, respectively, 50.8% and 37.5%. Cohen κ between first and last BD diagnoses was found to be low (k = 0.17).

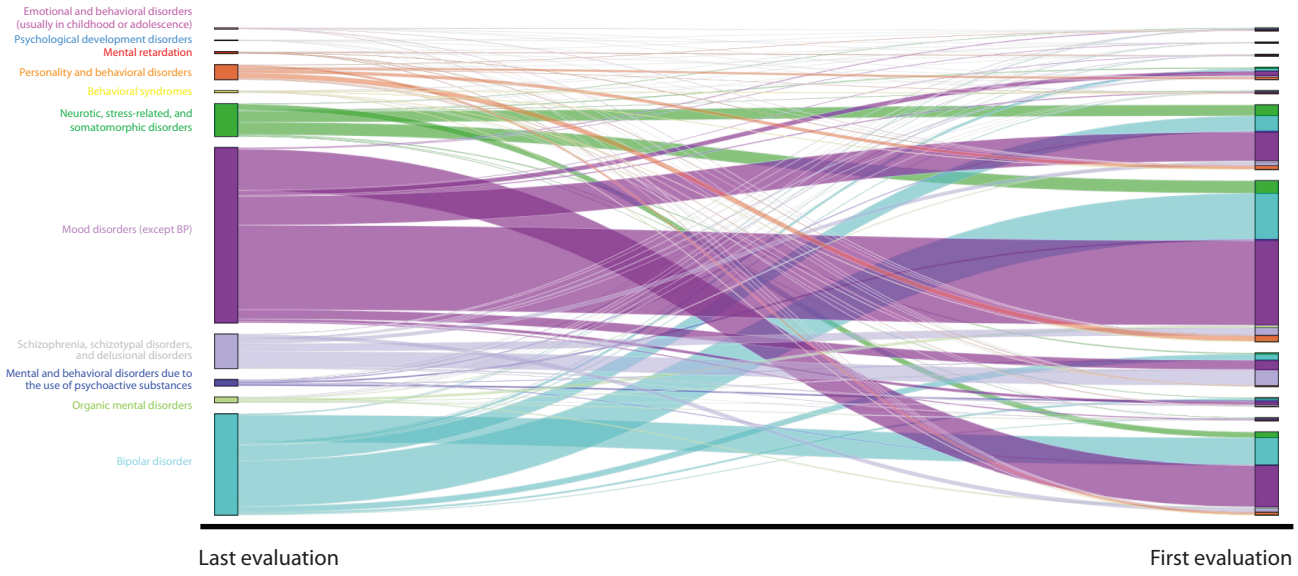
A category that led to diagnostic shift was schizophrenia, schizotypal disorders, and delusional disorders (F2): 8.3% (n = 449) of patients finally diagnosed with BD received an initial F2 diagnosis, and 7.5% (n = 301) of cases who were diagnosed with BD at the beginning had an F2 diagnosis as the final diagnosis.

With regard to non-bipolar affective disorders, out of 8,141 patients initially diagnosed with mood/affective disorders (F3), 51% (n = 4,153) had a non-BD diagnosis. In the final evaluation, 9,717 patients were assigned a mood/affective disorders diagnosis (F3), of which 5,396 were bipolar (F31).

One in 5 patients (21%, n = 1,135) initially diagnosed with neurotic disorders, stress-related disorders, and somato-morphic disorders (F4) were diagnosed with BD in the last visit. Conversely, 10.7% (n = 429) of patients diagnosed with BD at the first evaluation ended up with a diagnosis of neurotic, stress-related, and somato-morphic disorders.

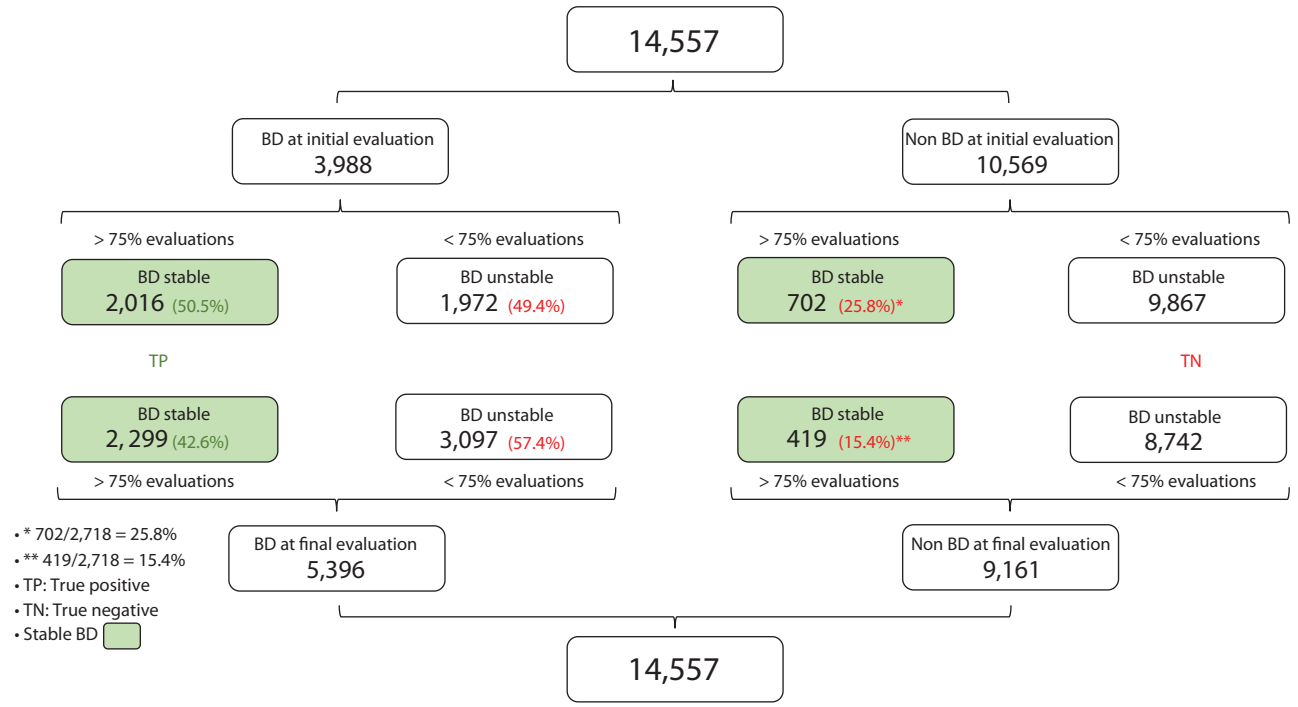
Diagnoses of personality and behavioral disorders in adults (F6) were initially assigned to 3.6% (n = 195) of those with a final diagnosis of BD. Conversely, in the last evaluation, personality and behavioral disorders in adults' diagnoses amounted to 4.2% (n = 167) of those initially categorized as having BD. The remaining diagnostic categories appeared in less than 3% of initial or final BD diagnoses.

Figure 2. Retrospective Consistency in Psychiatric Diagnoses[a]



[a]This graph is an "alluvial diagram." On the left side, the final diagnoses are shown; the width of each bar represents the number of patients with that diagnosis. On the right side, the initial diagnoses are shown, also in proportion. If the reader focuses on how a particular color of bar on the left side (final diagnosis) splits into several other bars on the right side, the proportions of the different initial diagnoses that converge to the same final diagnosis can be traced.

Figure 3. Bipolar Disorder Diagnoses Related to Follow-up Time



Diagnostic Stability of BD Diagnoses

Out of the total sample of 14,557 patients, only 18.6% (n = 2,718) were categorized as having a stable BD diagnosis (eg, retained the BD diagnosis in > 75% of clinical encounters). We summarize the findings in Figure 3.

Among these 2,718 "stable" patients diagnosed with BD, the mean time from the first therapeutic contact with the Mental Healthcare Center to the first time the patient was diagnosed with BD was 318.1 days (95% CI, 188.5–347.8). The average time from the first therapeutic contact within the Mental Healthcare Center to the last time the patient was diagnosed with BD was 7,386.7 days. The median was 7,429 days (95% CI, 7,068.2–7,705.2). There was a difference between the time needed to make the first diagnosis of BD

between those who kept a stable diagnosis (median = 0 days) and those who did not (median = 966 days) (Mantel-Cox test, $\chi^2 = 2{,}852.10$; $P < .0001$). If only patients with a stable BD diagnosis are taken into account, they represent 0.4% of the sample. Taking into account the total number of patients evaluated, about 3.1% had a diagnosis of BD at some point during follow-up.

More than 50% of the sample has been evaluated by the same psychiatrist at least 66% of the time. Among stable BD patients, 77% of the time the patients were seen by the same psychiatrist, and among BD unstable patients, 64% of the time the patients were seen by the same psychiatrist (Student $t_{14.083} = 27.081$, $P < .001$). Among stable BD patients, 14.2% were not evaluated by the same psychiatrist at the first and last evaluation, and among stable BD patients, 28.1% were seen by the same psychiatrist at the first and the last evaluation, with an OR of 2.375 (95% CI, 2.174–2.594).

The agreement found between the first diagnosis with BD stable group compared to the last diagnosis with BD stable group was as follows: BD stable—first diagnosis: k = 0.492, $P \leq .000$; BD stable—last diagnosis: k = 0.420, $P \leq .000$, respectively.

Factors Related to the Diagnostic Stability of BD

The United Nations, for statistical purposes, defines persons between the ages of 15 and 24 years as youth. Following this definition when performing the analysis on the stable BD group (when retained the BD diagnosis in > 75% of clinical encounters), greater stability was found if the diagnosis was made after hospitalization—OR = 1.932 (95% CI, 1.682–2.219), if the age at onset was > 25 years—OR = 4.318 (95% CI, 2.527–7.377), if the diagnosis of BD was made at age < 24 years—OR = 6.133 (95% CI, 3.477–10.817), if > 65% of the visits were held by the same psychiatrist—OR = 2.246 (95% CI, 1.978–2.550), and if the patient had been assessed by the same psychiatrist in the first and last assessments—OR = 1.667 (95% CI, 1.475–1.883) (Hosmer and Lemeshow test: $\chi^2_7 = 10{,}620$, $P = .156$).

DISCUSSION

The present study addresses the issue of diagnostic stability of BD in outpatient settings and contributes to the knowledge about the temporal consistency of BD and the usual diagnostic changes that occur during its evolution.

The results showed a limited number of stable BD diagnoses, notably lower than previous studies. Some methodological reasons could explain the differences with previous studies, especially the low number of evaluations and the shorter follow-up period used in previous studies. To be sure of the diagnosis of stable BD, at least 6–12 months is needed. The administrative prevalence (the proportion of the population in a defined area—the community of Madrid in this study—who are receiving services) of BD in this psychiatric sample is 0.4%, lower than usually reported in clinical and nonclinical samples.[4,36,37] However, this could be related to the fact that diagnoses were made in outpatient settings, and, as reported in previous studies,[16,17] a higher diagnostic stability is observed when the diagnosis is made after a discharge from hospital. A more detailed study of the factors that influence the stability of BD and a better knowledge of the course of diagnoses throughout its evolution are proposed as future lines of research.

The natural evolution of BD is prone to a high variability; however, the central symptoms of affective episodes are not present as frequently, and the presence of comorbid disorders, which is quite common, leads to misdiagnoses during daily clinical practice.[38,39] The present study found that the stability of BD was low, and even lower than in previous studies,[13,16,17,19,40] with the 3 different indices used.

In the first evaluation, 27.4% of the patients received a diagnosis of BD, and only 18.3% of the total sample was considered stable according to the criteria established in this study. Additionally, confusion surrounding the usual differential diagnoses of BD was found. These conclusions are detailed and discussed in the following sections.

It is important to emphasize that the increased specificity of the diagnostic criteria for BD in *ICD-10* versus *ICD-9* may have somewhat influenced our results.

It is essential to clarify that there has been no deinstitutionalization in Spain and that the registry was conceived as a tool during the psychiatric reform. Of note, this is one of the few epidemiologic studies on this issue conducted outside Scandinavia in which multiple types of stability measures were used, and this database has been used in previous works.[35,41]

BD Diagnosis at the First Evaluation

Diagnostic shifts in BD are especially frequent at first contact with the physician, with misleading initial symptoms due to substance abuse, depressive, or psychotic symptoms. The greatest prospective consistency was found in the mood/affective disorders category (F3), since the sample was selected among patients with at least 1 BD diagnosis. While only 27.4% of subjects were diagnosed with BD at the first evaluation, the rest were diagnosed at least once during subsequent evaluations. Similar results were presented in a previous study,[40] where it was noted that these figures were consistent with the high prevalence of misdiagnosis (48% and 69%) found in naturalistic research using self-administered questionnaires in general practitioner consultations[42,43] and also in studies in which diagnoses were based on the application of *DSM-IV* criteria.[44]

However, in our sample, 50.5% of patients who were diagnosed with BD at the first evaluation remained stable in three-quarters of the evaluations. This fact is not consistent with the figure reported by Chen et al,[13] who noted that 70% of the subjects with an initial diagnosis of BD did not change to a different category over time. On the other hand, the percentage of patients with a stable diagnosis of BD (n = 2,718) who were correctly diagnosed in the first evaluation (n = 2,016) increases in our sample to 74.2%. These results support the hypothesis of the diagnostic difficulty of BD in the first evaluations.

BD Diagnosis at the Last Evaluation

The latest evaluation showed an increase in the number of diagnoses of BD (37.1% of the sample), and, of those, 42.6% had been stable throughout the study. On the other hand, 84.6% of patients with stable diagnoses (n = 2,718) were accurately diagnosed in their last visit (n = 2,299).

This result may reflect a progressive increase in diagnostic stability throughout the evaluations (in our case a minimum of 10), which is congruent with the idea that routine reassessment could improve the chances of a successful diagnostic process. However, Schwartz et al in 2000[17] reported that the retrospective consistency of BD was 85% when comparing 6-month and 24-month diagnoses but was reduced to 73% when comparing initial and 24-month diagnoses. This would mean that consistency rates for some diagnoses decreased as the follow-up period increased. In any case, the retrospective consistency of BD in our study (37.5%) is low compared to other studies that measured it (58.4%–94.4%), similar to that presented by Baca-García et al in 2007 (38%),[35] and higher than Weeke in 1984 (20%).[45] The low retrospective consistency might be explained by the fact that this is a longitudinal study based on data retrieved from community mental health centers and hence conducted in a real-world scenario with the general population, and not in a BD-specific unit where patients start off already correctly diagnosed.

Diagnostic Stability of BD

To our knowledge, this is the largest longitudinal study, with 14,557 patients over 30 years of study, that has evaluated the diagnostic stability of BD under ecological conditions. In 2005, Kessing[40] mentioned that no study had investigated the diagnostic stability of the most common *ICD-10* psychiatric diagnoses administered under ecological clinical conditions. This is the case for our study, which has shown a low stability of the *ICD-10* BD categories, measured by temporal consistency and diagnostic constancy, with findings considerably lower than in previous studies. The reasons for these differences in diagnostic temporal stability are not clear but may be due to the large sample size, the extensive duration of follow-up, the high number of evaluations, diagnostic criteria, or sociodemographic variables.

Time consistency showed low results with a prospective consistency of 50.8% and a retrospective consistency of 37.5%. It should be noted that the κ value was low (κ = 0.17) between the first and last diagnosis. However, since κ values take into account stable positive cases and stable negative cases, but also cases that remit and new cases, low κ values can be observed if a high number of new or remitting cases occur[46] and therefore do not necessarily reflect a lack of diagnostic stability.

The results of our study showed that only 18.3% of patients were diagnosed with BD in 75% of the evaluations. In these patients with stable diagnosis, the mean number of evaluations until the first diagnosis of BD was 3.7, with an average time of 318.1 days. These values were increased to 21.2 evaluations and 1,511.2 days within the group with a non-stable diagnosis. Thus, patients with a stable diagnosis of BD were diagnosed earlier (less than 1 year) and needed fewer evaluations than those without a stable diagnosis (somewhat more than 4 years until the BD diagnosis was made).

In our study, patients with a stable BD diagnosis achieved diagnostic stability at 7,386.7 days (slightly more than 20 years) and 279 follow-up visits. Patients with a non-stable diagnosis had their diagnosis withdrawn at 2,929 days (approximately 8 years) and after 55 evaluations. These results could be in line with previous reports by Hirschfeld et al[42] in 2003 and Baldessarini et al[47] in 1999, who reported a delay in correct diagnosis of about 8–10 years from the onset of the disease. In our study, the data suggest that both consolidating and withdrawing the diagnosis of BD are tasks that require many years of follow-up and numerous evaluations; while consolidating required about 14 visits/year, withdrawing entailed fewer than 7 visits/year. This fact may reflect that patients with a stable diagnosis of BD are more complex and require more health care than those for whom this diagnosis is withdrawn.

Although it was not the main objective of our study, 4 variables not only were included but were predictive of diagnostic stability of BD: marital status, educational level, work situation, and personal history of psychiatric care. In another preliminary study,[6] 4 variables related to the stability of bipolar diagnosis were found: sex, age ≥ 40 years, number of psychiatric consultations, and outpatient Mental Health Centers. In any case, more studies focusing on these variables are needed.

The higher consistency rates found by other authors[13,16,17,19,40] may have been influenced by a number of drawbacks that diminish the generalizability of these studies.

Diagnostic Shifts in BD

Patients with a stable BD diagnosis had some diagnostic fluctuation that included the typical differential diagnoses of BD. Our study found high rates of misdiagnosis of BD with other affective disorders: 44.5% of patients who were diagnosed at the first evaluation of a non-bipolar affective disorder were eventually diagnosed at the last follow-up visit with BD, and 51% of patients who were initially diagnosed with BD were no longer diagnosed at the last evaluation. Previous studies concur that the high rates of misdiagnosis derive from confusion with unipolar depression,[42,43] especially in cases in which the BD debuts with 1 or more depressive episodes. As for neurotic and anxiety disorders, the percentage of these diagnoses at the beginning is high (21.03%) in patients who are ultimately diagnosed with BD. Other less frequent diagnostic shifts occurred with the spectrum of schizophrenia (7.5% at first evaluation and 8.3% at the last evaluation) and with personality disorders (4.2% baseline and 3.6% final).

Many factors may be involved in the unstable progression of a psychiatric diagnosis. Schwartz et al[17] mentioned that diagnostic changes over time may reflect the evolution of a disease, the emergence of new information, or the

unreliability of measurements. The relative lack of stability in diagnoses over time in this study may be due to disease progression or reflect weaknesses inherent in clinical evaluations. There could be other ways to determine the validity of stable diagnoses such as the use of prescriptions for mood stabilizers, this being a limitation of the study.

The results of this study raise concerns about psychiatric research findings, especially in studies with short follow-up periods for chronic conditions that may not allow enough time to reach an accurate diagnosis or in studies that do not take into account the context.

Our study has limitations. Our study has limitations. First, we did not consider differences between type I and II bipolar disorder, despite clinical and prognostic implications, because codes were mostly recorded following *ICD-9*, where type II bipolar disorder could not be specified. Second, diagnoses in our database are recorded following the independent judgment of clinicians rather than alternative assessments such as research scales. This, however, enhances external validity of our results, as they likely reflect the course of illness from a real clinical practice perspective. While diagnostic scales can reduce measurement error, they (1) require specific training and are too time-consuming to be used routinely and (2) are mostly validated in the context of highly selected samples of patients, for research purposes. In conclusion, as noted previously in the literature,[48] our results should be considered an externally valid representation of patterns of real clinical diagnostic change, which has important implications for treatment planning, rather than patterns in the prevalence of the disorder. Also, a limitation was that the form filled out at each visit consisted of sociodemographic data and psychiatric diagnosis, leaving out other relevant data.

A point to keep in mind is that the question of whether diagnostic changes in our data (eg, patients whose diagnoses changed from or to bipolar disorder) reflect misdiagnosis, the natural history of the phenotypical presentation of these patients' disease, or a mix of both cannot be clarified using this data source. Accordingly, conclusions regarding over- or underdiagnosis of BD based on our results should be made with caution. The study is also limited by the possible existence of uncontrolled pathways of psychiatric care but may more accurately reflect real clinical practice, perhaps revealing the poor accuracy of clinical evaluation systems in usual practice.

CONCLUSIONS

1. This work reflects real conditions in a daily practice over a 30-year period of observation, including both outpatient consultations and hospitalization.
2. In our sample, the administrative prevalence of stable BD is 0.4%; however, the diagnosed prevalence is 3.1% when all patients diagnosed with BD are included.
3. A delay of > 10 years to achieve diagnostic stability was found.
4. Frequent diagnostic shifts were found in relation to BD, the most common being with other affective disorders. Anxiety was also a common misdiagnosis. There is a 50% diagnostic error rate when BD is diagnosed in the first evaluation.
5. The most diagnostically stable patients are diagnosed at the first visit.
6. Greater stability was observed if age at onset was > 25 years, BD diagnosis was made at age < 24 years, or diagnosis was made after hospitalization (which may explain the low prevalence we found in this study in comparison to other studies in which only hospitalized patients were considered).
7. The low stability detected in this study should be taken into account when evaluating results compared to clinical and epidemiologic trials, in which samples were smaller and followed for less time.

Submitted: November 10, 2020; accepted April 12, 2021.

Published online: Month 00, 2021.

Potential conflicts of interest: None.

Funding/support: This study was partly funded by Carlos III (ISCIII PI16/01852), American Foundation for Suicide Prevention (LSRG-1-005-16), the Madrid Regional Government (B2017/BMD-3740 AGES-CM 2CM; Y2018/TCS-4705 PRACTICO-CM), MINECO/FEDER (ADVENTURE, id. TEC2015-69868-C2-1-R), MCIU Explora Grant aMBITION (id. TEC2017-92552-EXP), and MICINN (CLARA, id. RTI2018-099655-B-I00).

Role of the sponsor: The sponsor was not involved in any of the following: design and conduct of the study; collection, management, analysis, and interpretation of the data; and preparation, review, or approval of the manuscript.

Acknowledgments: The authors thank the health and administrative staff who participated in the data collection for the study. They also thank the reviewers for their comments and observations, which have contributed to the improvement of this work.

REFERENCES

1. Institute for Health Metrics and Evaluation [Internet]. Institute for Health Metrics and Evaluation. Cited September 28, 2020. http://www.healthdata.org/institute-health-metrics-and-evaluation
2. Ferrari AJ, Stockings E, Khoo J-P, et al. The prevalence and burden of bipolar disorder: findings from the Global Burden of Disease Study 2013. *Bipolar Disord*. 2016;18(5):440–450.
3. Hayes JF, Marston L, Walters K, et al. Mortality gap for people with bipolar disorder and schizophrenia: UK-based cohort study 2000–2014. *Br J Psychiatry*. 2017;211(3):175–181.
4. Hidalgo-Mazzei D, Undurraga J, Reinares M, et al. The real world cost and health resource utilization associated to manic episodes: the MANACOR study. *Rev Psiquiatr Salud Ment*. 2015;8(2):55–64.
5. Pini S, de Queiroz V, Pagnin D, et al. Prevalence and burden of bipolar disorders in European countries. *Eur Neuropsychopharmacol*. 2005;15(4):425–434.
6. Baca-Garcia E, Perez-Rodriguez MM, Basurte-Villamor I, et al. Diagnostic stability of psychiatric disorders in clinical practice. *Br J Psychiatry*. 2007;190(3):210–216.
7. Palomar-Ciria N, Cegla-Schvartzman F, Lopez-Morinigo J-D, et al. Diagnostic stability of schizophrenia: a systematic review. *Psychiatry Res*. 2019;279:306–314.
8. Woo BKP, Sevilla CC, Obrocea GV. Factors influencing the stability of psychiatric diagnoses in the emergency setting: review of 934 consecutively inpatient admissions. *Gen Hosp Psychiatry*. 2006;28(5):434–436.
9. Daradkeh TK. Stability of psychiatric diagnoses in clinical practice. *Int J Soc Psychiatry*. 1996;42(3):207–212.
10. Daradkeh T, El-Rufaie O, Younis Y, et al. The diagnostic stability of *ICD-10* psychiatric diagnoses in clinical practice. *Eur Psychiatry*. 1997;12(3):136–139.
11. Addington J, Chaves A, Addington D. Diagnostic stability over one year in first-episode psychosis. *Schizophr Res*. 2006;86(1–3):71–75.
12. Amin S, Singh SP, Brewin J, et al. Diagnostic stability of first-episode psychosis. Comparison of *ICD-10* and *DSM-III-R* systems. *Br J Psychiatry*. 1999;175(6):537–543.

13. Chen YR, Swann AC, Johnson BA. Stability of diagnosis in bipolar disorder. *J Nerv Ment Dis*. 1998;186(1):17–23.
14. 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(8):1200–1208.
15. 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(3):187–192.
16. Schimmelmann BG, Conus P, Edwards J, et al. Diagnostic stability 18 months after treatment initiation for first-episode psychosis. *J Clin Psychiatry*. 2005;66(10):1239–1246.
17. Schwartz JE, Fennig S, Tanenberg-Karant M, et al. Congruence of diagnoses 2 years after a first-admission diagnosis of psychosis. *Arch Gen Psychiatry*. 2000;57(6):593–600.
18. Tsuang MT, Woolson RF, Winokur G, et al. Stability of psychiatric diagnosis. Schizophrenia and affective disorders followed up over a 30- to 40-year period. *Arch Gen Psychiatry*. 1981;38(5):535–539.
19. Veen ND, Selten J-P, Schols D, et al. Diagnostic stability in a Dutch psychosis incidence cohort. *Br J Psychiatry*. 2004;185(6):460–464.
20. Cegla-Schvartzman FB, Ovejero S, López-Castroman J, et al. Diagnostic stability in bipolar disorder: a narrative review. *Harv Rev Psychiatry*. 2019;27(1):3–14.
21. Ratheesh A, Cotton SM, Betts JK, et al. Prospective progression from high-prevalence disorders to bipolar disorder: exploring characteristics of pre-illness stages. *J Affect Disord*. 2015;183:45–48.
22. Amini H, Alaghband-rad J, Omid A, et al. Diagnostic stability in patients with first-episode psychosis. *Australas Psychiatry*. 2005;13(4):388–392.
23. Baldwin P, Browne D, Scully PJ, et al. Epidemiology of first-episode psychosis: illustrating the challenges across diagnostic boundaries through the Cavan-Monaghan study at 8 years. *Schizophr Bull*. 2005;31(3):624–638.
24. Rufino ACTBF, Uchida RR, Vilela JAA, et al. Stability of the diagnosis of first-episode psychosis made in an emergency setting. *Gen Hosp Psychiatry*. 2005;27(3):189–193.
25. Parker G. Borderline personality disorder and bipolar disorder: commentary on Paris and Black. *J Nerv Ment Dis*. 2015;203(1):13–14.
26. Moore PJ, Little MA, McSharry PE, et al. Forecasting depression in bipolar disorder. *IEEE Trans Biomed Eng*. 2012;59(10):2801–2807.
27. Padrón continuo 1996. Total Comunidad de Madrid [Internet]. Available from: http://www.madrid.org/iestadis//fijas/estructu/demograficas/padron/pc96i1.htm
28. Carmona R, Cook BL, Baca-García E, et al. Comparison of mental health treatment adequacy and costs in public hospitals in Boston and Madrid. *J Behav Health Serv Res*. 2019;46(3):464–474.
29. Wang PS, Demler O, Kessler RC. Adequacy of treatment for serious mental illness in the United States. *Am J Public Health*. 2002;92(1):92–98.
30. BOE. BOE - Ley 14/2007 [Internet]. 2007. Available from: https://www.boe.es/boe/dias/2007/07/04/
31. Benchimol EI, Smeeth L, Guttmann A, et al; RECORD Working Committee. The REporting of studies Conducted using Observational Routinely-collected health Data (RECORD) statement. *PLoS Med*. 2015;12(10):e1001885.
32. eCIE-Maps - CIE-9-MC [Internet]. [cited 2020 Sep 28]. Available from: https://eciemaps.mscbs.gob.es/ecieMaps/browser/index_9_mc.html
33. Pierce K, Gazestani VH, Bacon E, et al. Evaluation of the diagnostic stability of the early autism spectrum disorder phenotype in the general population starting at 12 months. *JAMA Pediatr*. 2019;173(6):578–587.
34. 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.
35. Baca-Garcia E, Perez-Rodriguez MM, Basurte-Villamor I, et al. Diagnostic stability and evolution of bipolar disorder in clinical practice: a prospective cohort study. *Acta Psychiatr Scand*. 2007;115(6):473–480.
36. Alonso J, Angermeyer MC, Lépine JP; European Study of the Epidemiology of Mental Disorders (ESEMeD) Project. The European Study of the Epidemiology of Mental Disorders (ESEMeD) project: an epidemiological basis for informing mental health policies in Europe. *Acta Psychiatr Scand suppl*. 2004;109(420):5–7.
37. 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(12):587–595.
38. Blacker D, Tsuang MT. Contested boundaries of bipolar disorder and the limits of categorical diagnosis in psychiatry. *Am J Psychiatry*. 1992;149(11):1473–1483.
39. 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.
40. Kessing LV. Diagnostic stability in bipolar disorder in clinical practise as according to ICD-10. *J Affect Disord*. 2005;85(3):293–299.
41. Lopez-Castroman J, Leiva-Murillo JM, Cegla-Schvartzman F, et al. Onset of schizophrenia diagnoses in a large clinical cohort. *Sci Rep*. 2019;9(1):9865.
42. Hirschfeld RMA, 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(2):161–174.
43. Lish JD, Dime-Meenan S, Whybrow PC, et al. The National Depressive and Manic-depressive Association (DMDA) survey of bipolar members. *J Affect Disord*. 1994;31(4):281–294.
44. Ghaemi SN, Sachs GS, Chiou AM, et al. Is bipolar disorder still underdiagnosed? are antidepressants overutilized? *J Affect Disord*. 1999;52(1-3):135–144.
45. Weeke A. Admission pattern and diagnostic stability among unipolar and bipolar manic-depressive patients. *Acta Psychiatr Scand*. 1984;70(6):603–613.
46. Mattanah JJ, Becker DF, Levy KN, et al. Diagnostic stability in adolescents followed up 2 years after hospitalization. *Am J Psychiatry*. 1995;152(6):889–894.
47. Baldessarini RJ, Tondo L, Hennen J. Treatment delays in bipolar disorders. *Am J Psychiatry*. 1999;156(5):811–812.
48. Moreno C, Laje G, Blanco C, et al. National trends in the outpatient diagnosis and treatment of bipolar disorder in youth. *Arch Gen Psychiatry*. 2007;64(9):1032–1039.

Sex, Gender, and Suicidal Behavior

María Luisa Barrigón and Fanny Cegla-Schvartzman

Contents

Abstract This chapter reviews gender differences in suicide, commonly known as the gender paradox in suicide. While men are more likely to complete suicide, suicide attempts are more frequent in women. Although there are exceptions, this paradox occurs in most countries over the world, and it is partially explained by the preference of men for more lethal methods. Nevertheless, there are differences in the known risk factors for suicide between men and women, and this chapter summarizes the more relevant findings for the gender paradox. Apart from previous attempts, which still is the strongest predictor of death by suicide, with a higher rate in males than in females, we will emphasize in the role of male depression. It is

M. L. Barrigón (✉) and F. Cegla-Schvartzman
Hospital Universitario Fundación Jiménez Díaz, Madrid, Spain
e-mail: luisa.barrigon@fjd.es

Curr Topics Behav Neurosci
https://doi.org/10.1007/7854_2020_165

commonly recognized that over 90% of people who die by suicide had a psychiatric diagnosis, mostly depression, and male depression seems to be a distinct clinical phenotype challenging to recognize, which might contribute to the gender paradox. Finally, in light of all the information reviewed, some recommendations on prevention of suicide from a gender perspective in the clinical setting will be made.

Keywords Gender · Gender differences · Prevention · Risk factors · Sex · Suicide · Suicide attempted

1 Introduction

Suicide is a complex multifactorial phenomenon; the precise process underlying suicide involves many issues still unresolved. Suicide rates are different according to region, sex, age, time, ethnic origin, and, presumably, death registration methods (Hawton and van Heeringen 2009). One of the classic risk factors for suicide is male sex/gender, with men being more likely to complete suicide and women to try it; this is known as the gender paradox in suicide (Canetto and Sakinofsky 1998). Apart from China, Bangladesh, and Lesotho, where there are more deaths by suicide in women than in men, this is a demonstrated fact worldwide (Naghavi 2019; WHO, Suicide Data 2019).

Sex differences in suicidal behavior have been extensively studied, but most of the research in this topic is made assuming equivalence between sex and gender, and, furthermore, only in most recent times, research has also focused on suicidal behavior in a non-binary gender (Fox et al. 2018). Sex and gender are constructs that should be differentiated. Essentially, sex refers to biology while gender involves cultural constructs (Clayton and Tannenbaum 2016), and differences in suicidal behavior are probably more related to cultural issues (i.e., gender) than only to a biological basis (i.e., sex). In suicidology, both constructs should be taken into account, but by now it is difficult to perform a specific search on sex or gender exclusively; therefore, this chapter will use both terms interchangeably, with a preference for the term "gender," given the role of cultural factors in suicide phenomena.

In this chapter, we are going to review gender differences in suicide rates across different countries of the world. Then we will describe and reflect on gender differences in suicide methods. Next, suicide risk factors will be study focusing on the effect of male or female gender. Finally, different preventive approaches will be suggested in light of gender differences previously reviewed.

Across different sections of the chapter, it will be observed how the role of traditional masculinity is frequently involved in gender differences in suicide, and how men are a high-risk population, and specifical and tailored interventions are needed for them.

2 Terminology

2.1 Suicidal Behavior

In this chapter, we use the terminology for suicide based on the definition given by O'Carroll et al. in 1996 (O'Carroll et al. 1996) and later redefined by Silverman et al. in 2007 (Silverman et al. 2007). Thus, here we use the term *suicidal behavior* for denoting any type of suicidality, that is, suicidal ideation, suicidal plans, non-fatal suicide attempts, and deaths by suicide. Subsequently, we will use the term *suicidal ideation* for "unelaborated thoughts related to the wish and/or intention of taking one's life," *suicidal plan* for "an elaborated and structured suicidal ideation with decisions made as how to perform the suicide attempt," and *suicide attempts* for "any act of self-harm performed with the intention of taking one's life"; suicide attempts could lead to *non-fatal suicide attempts* or *death by suicide*. Hence, in this chapter, we will mostly refer to *death by suicide* or just *suicide.* Non-suicidal self-injuries (NSSI), a descriptive term employed in the DSM 5, will not be addressed in this chapter. Furthermore, extended and assisted suicide are not part of this chapter.

2.2 Sex and Gender

Sex and gender are terms frequently used interchangeably in ordinary speech. Indeed, in some languages there are not two different words for both constructs. Although in scientific terms, sex and gender are not strictly exchangeable, both terms are non-exclusive, but are related to each other and influence health in different ways (Clayton and Tannenbaum 2016). Primarily, while sex refers to biology, the term gender includes psychosocial factors (Clayton and Tannenbaum 2016).

Sex refers to the biological characteristics that define humans as female or male, which is determined by the genetic information of chromosomes, and includes cellular and molecular differences (Dunn et al. 2016). The World Health Organization (WHO) states that "sex refers to the biological and physiological characteristics that define men and women" and "'male' and 'female' are sex categories" (WHO, Defining Sexual Health 2019). Male or female sexual differentiation is based in karyotype at birth, 46XX for female sex and 46XY for males, and is physiologically characterized by the gonads (ovary or testes), sex hormones (testosterone and estrogen), external genitalia (e.g., penis or vulva), and internal reproductive organs (e.g., uterus or prostate gland) (Clayton and Tannenbaum 2016). People with mixed sex factors are intersex.

On the other hand, **gender** refers to the socially constructed characteristics of women and men and comprises the social, environmental, cultural, and behavioral factors that influence a person's identity of being a man or a woman (Clayton and Tannenbaum 2016). In the sphere of gender, several aspects must be distinguished: gender assignment, gender roles, and gender identity. *Gender assignment* is how an

infant is classified at birth, as either male or female based on external genitalia (WHO, Defining Sexual Health 2019). *Gender roles or gender norms* are unspoken rules in the family, workplace, institutions, or global culture that influence individual attitudes and behaviors (Schiebinger and Stefanick 2016). Finally, *gender identity* refers to how individuals and groups perceive and present themselves (Clayton and Tannenbaum 2016), but rather than a binary concept, there are gender identity gradations from masculinity to femininity (Fausto-Sterling 2008). When a mismatch between gender identity and gender assigned exists, we refer to "transgender," and so the term transgender includes people whose gender identity is the opposite of their assigned sex (trans men and trans women), but also includes people who feel not exclusively masculine or feminine (genderqueer, non-binary, bigender, pangender, genderfluid, or agender) (Fausto-Sterling 2008). Thus, gender identity is not an entirely fixed characteristic, and many transgender people move fluidly between identities over time, often without any specific labels (Haas et al. 2011).

Although, as it has been previously exposed, gender and sex are not equivalent, we should point out that in this chapter both terms will be used indistinctly since a specific search for each term is complicated due to the fact that previous scientific research has not generally made the distinction (Clayton and Tannenbaum 2016). A reflection on this should be made, and currently, many journals encourage authors to transparently report sex, gender, or even both in their works. Generally in research sex/gender are visually assigned to research participants without specifically asking, and even more, there are no validated tools for assessing gender, and an approach in which participants were asked first about sex assigned at birth and then about gender identity has been proposed (Clayton and Tannenbaum 2016).

Furthermore, transgender condition impacts death by suicide and suicide behavior, and it has been extensively studied, especially in recent years (Fox et al. 2018; Narang et al. 2018). Although highly interesting, this topic is beyond the scope of our review and will not be covered in this chapter.

3 Worldwide Suicide Rates by Gender

The World Health Organization (WHO) provides the most exhaustive and unbiased data on suicide rates from its member states and periodically updates them. Currently, the last available suicide data are from 2016 (WHO, Suicide Data 2019). According to WHO data, in 2016 the global male/female ratio of age-standardized suicide rates was 1.8, meaning that worldwide, men complete suicide almost twice more often than women (WHO, Suicide Rates (per 100 000 population) 2019). Interestingly, this ratio is particularly high in Europe (around 4:1) and in high-income countries but lower in low- and middle-income countries (around 1.6:1) (Saxena et al. 2014). Asian countries typically show much lower male/female ratios (Chen et al. 2012). Furthermore, comparing the information from the WHO countries, the male/female ratio ranged from 0.8 in Bangladesh and China to 12.2 in St. Vincent and the Grenadines (Bachmann 2018).

This is graphically shown in the map developed periodically by WHO, in which the lighter-colored countries represent those in which more women than men die by suicide and, on the contrary, the darker-colored ones represent those countries in which more men than women die by suicide (Fig. 1).

Similar figures are thrown by the Global Burden of Disease (GBD) Study (Naghavi 2019). According to GBD, male suicide age-standardized rate was higher (15.6 deaths/100,000, 95% uncertainty interval 13.7 to 17.2) than female rate (7.0 deaths/100,000, 95% uncertainty interval 6.5 to 7.4), However, the rate of decrease from 1990 to 2016 was lower for male (23.8%, 95% uncertainty interval 15.6% to 32.7%) than for female (49.0%, 42.6% to 54.6%). Figure 2 shows regional trends of age-standardized suicide rates for women and men.

Suicide rates vary in different countries throughout the world. Details of rates by country, according to WHO data (WHO, Suicide Rates (per 100 000 population) 2019), are shown in Table 1.

In this table, it could be observed how, in most countries, male suicide rates are higher than female, with Ukraine, Lithuania, or Russia among the top. Only in a limited number of countries, most from East Asia, the opposite happens. This is widely known for China (Simon et al. 2013), Bangladesh (Sharmin Salam et al. 2017), and Pakistan (Shekhani et al. 2018), but also consistently observed in African countries such as Morocco or Lesotho. Concerning Morocco, as in many other Arabic countries where suicide is a taboo act, there is a notable lack of national suicide rates, and studies on the topic are scarce (Bjegovic-Mikanovic et al. 2019). Finally, for Lesotho no specific studies about suicides have been found. Altogether, and in a simplistic approach, the highest male/female suicide rates are found in Eastern European countries and the lowest in the WHO Southeast Asia Region.

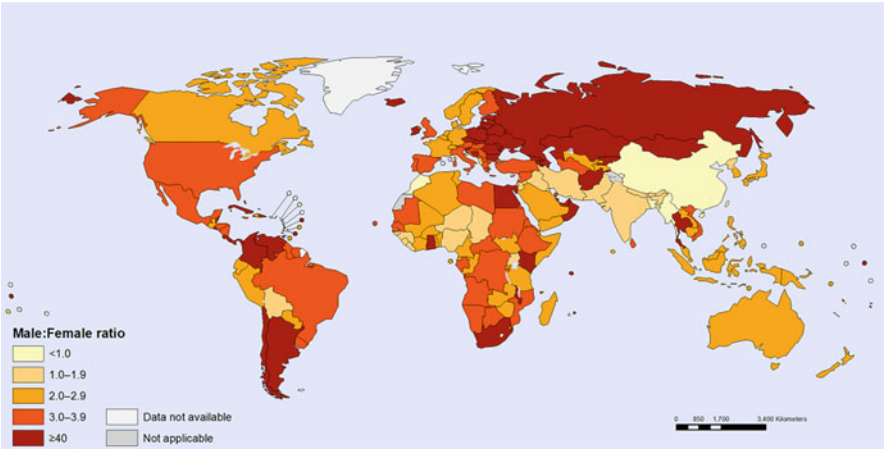


Fig. 1 2016 map of male/female ratio of age-standardized suicide rates from 2016. Picture obtained from WHO Global Health Observatory data repository

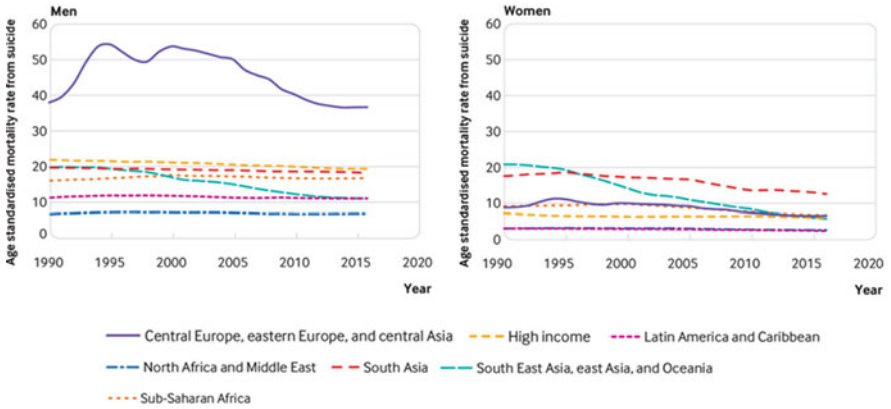


Fig. 2 Global Burden of Disease regions' age-standardized suicide rates for women and men (1990 to 2016). Modification of figure from the Global Burden of Disease Study 2016 (British Medical Journal, 2019; 364: l94)

Although there are many unanswered questions regarding these differences in the male/female suicide ratio across countries, probably cultural factors must be taken into account to understand them (Canetto 2008). In this sense, the review of Ahmed et al. highlighted how, in the United Kingdom, rates of self-harm among South Asian women are much higher than among their White counterparts (Ahmed et al. 2007).

Finally, it is worth to mention how the research investment in suicide does not correspond with the worldwide distribution of suicide, and we should point out that more research should be developed to better understand the male-female gap in suicide (Lopez-Castroman et al. 2015).

4 Gender Differences in Suicide Methods

Returning to the concept of gender paradox in suicide, men completed suicide up to three times more than women, while for suicide attempts, an inverse ratio is found (Bachmann 2018), and this difference is partially explained by the preference of men for more lethal methods (Ajdacic-Gross et al. 2008; Saxena et al. 2014). In the latest overview of suicide methods, in 2008, authors differentiated hanging, drowning, falls, pesticide poisoning, other poisoning, firearms, and others (Ajdacic-Gross et al. 2008). Globally, for both genders together, the most commonly used methods to complete suicide are hanging, self-poisoning with pesticides, and use of firearms (Saxena et al. 2014). Methods differ between world regions and between males and females. Next, is summarized the latest data on methods for males and females according to different world regions:

Table 1 Year 2016 age-standardized suicide rates (per 100,000 population)

Country	Both sexes	Male	Female
Afghanistan	6.4	10.6	2.1
Albania	5.6	7.0	4.3
Algeria	3.3	4.9	1.8
Angola	8.9	14.0	4.6
Antigua and Barbuda	0.5	0.0	0.9
Argentina	9.1	15.0	3.5
Armenia	5.7	10.1	2.0
Australia	11.7	17.4	6.0
Austria	11.4	17.5	5.7
Azerbaijan	2.6	4.3	1.0
Bahamas	1.6	2.8	0.5
Bahrain	5.7	7.9	2.1
Bangladesh	6.1	5.5	6.7
Barbados	0.4	0.8	0.3
Belarus	21.4	39.3	6.2
Belgium	15.7	22.2	9.4
Belize	5.9	9.9	2.0
Benin	15.7	22.6	9.6
Bhutan	11.6	13.8	8.9
Bolivia	12.9	16.9	8.9
Bosnia and Herzegovina	6.4	10.6	2.5
Botswana	11.5	18.3	5.7
Brazil	6.1	9.7	2.8
Brunei Darussalam	4.5	6.2	2.8
Bulgaria	7.9	13.1	3.2
Burkina Faso	14.8	22.4	9.1
Burundi	15.0	23.1	7.7
Cabo Verde	15.1	24.1	7.7
Cambodia	5.9	9.0	3.2
Cameroon	19.5	26.9	12.5
Canada	10.4	15.1	5.8
Central African Republic	11.6	18.0	6.0
Chad	15.5	17.1	13.8
Chile	9.7	16.0	3.8
China	8.0	7.9	8.3
Colombia	7.0	11.5	2.8
Comoros	11.1	17.6	5.4
Congo	9.3	13.9	5.0
Costa Rica	7.5	12.8	2.3
Cote d'Ivoire	23.0	32.0	13.0
Croatia	11.5	18.8	5.1
Cuba	10.1	16.4	4.1

(continued)

Table 1 (continued)

Country	Both sexes	Male	Female
Cyprus	4.5	7.2	1.9
Czechia	10.5	17.2	4.2
South Korea	10.6	14.8	8.0
Congo	9.7	15.0	4.9
Denmark	9.2	13.2	5.2
Djibouti	8.5	11.9	5.3
Dominican Republic	10.5	17.9	3.2
Ecuador	7.2	10.7	3.8
Egypt	4.4	7.2	1.7
El Salvador	13.5	24.8	4.3
Equatorial Guinea	22.0	31.3	10.8
Eritrea	13.8	22.4	6.1
Estonia	14.4	25.6	4.4
Eswatini	16.7	25.4	9.6
Ethiopia	11.4	18.7	4.7
Fiji	5.5	8.8	2.5
Finland	13.8	20.8	6.8
France	12.1	17.9	6.5
Gabon	9.6	15.0	4.3
Gambia	10.0	12.8	7.3
Georgia	6.7	12.3	1.9
Germany	9.1	13.6	4.8
Ghana	8.7	15.8	2.9
Greece	3.8	6.1	1.5
Grenada	1.7	2.1	1.0
Guatemala	2.9	4.4	1.7
Guinea	10.5	12.7	8.4
Guinea-Bissau	7.4	8.9	6.1
Guyana	30.2	46.6	14.2
Haiti	12.2	18.3	6.4
Honduras	3.4	5.3	1.7
Hungary	13.6	22.2	6.2
Iceland	13.3	21.7	4.7
India	16.5	18.5	14.5
Indonesia	3.7	5.2	2.2
Iran	4.0	4.9	3.1
Iraq	4.1	4.7	3.4
Ireland	10.9	17.6	4.2
Israel	5.2	8.2	2.4
Italy	5.5	8.4	2.6
Jamaica	2.0	3.2	0.9
Japan	14.3	20.5	8.1

(continued)

Table 1 (continued)

Country	Both sexes	Male	Female
Jordan	3.7	4.7	2.7
Kazakhstan	22.8	40.1	7.7
Kenya	5.6	9.7	2.1
Kiribati	15.2	25.9	5.4
Kuwait	2.2	2.5	1.7
Kyrgyzstan	9.1	14.8	3.7
Lao	9.3	12.9	6.1
Latvia	17.2	31.0	5.1
Lebanon	3.2	4.2	2.2
Lesotho	28.9	22.7	32.6
Liberia	13.4	13.8	13.0
Libya	5.5	8.7	2.3
Lithuania	25.7	47.5	6.7
Luxembourg	10.4	15.0	5.8
Madagascar	6.9	10.5	3.6
Malawi	7.8	13.7	3.2
Malaysia	6.2	8.7	3.6
Maldives	2.7	3.6	1.6
Mali	8.9	13.5	4.7
Malta	6.5	10.3	2.8
Mauritania	7.5	12.1	3.6
Mauritius	7.3	12.5	2.2
Mexico	5.2	8.2	2.3
Micronesia	11.3	16.2	6.2
Mongolia	13.3	23.3	3.8
Montenegro	7.9	12.6	3.6
Morocco	3.1	2.5	3.6
Mozambique	8.4	14.0	4.1
Myanmar	8.1	6.3	9.8
Namibia	11.5	19.4	4.9
Nepal	9.6	11.4	8.0
Netherlands	9.6	12.9	6.4
New Zealand	11.6	17.3	6.2
Nicaragua	11.9	19.2	5.0
Niger	9.0	11.5	6.7
Nigeria	17.3	17.5	17.1
Norway	10.1	13.6	6.5
Oman	3.5	4.8	0.9
Pakistan	3.1	3.0	3.1
Panama	4.4	7.6	1.2
Papua New Guinea	7.0	10.2	3.8
Paraguay	9.3	12.3	6.2

(continued)

Table 1 (continued)

Country	Both sexes	Male	Female
Peru	5.1	7.6	2.7
Philippines	3.7	5.2	2.3
Poland	13.4	23.9	3.4
Portugal	8.6	14.3	3.8
Qatar	5.8	7.3	1.1
North Korea	20.2	29.6	11.6
Moldova	13.4	24.1	3.8
Romania	8.0	13.9	2.4
Russia	26.5	48.3	7.5
Rwanda	11.0	16.9	0.0
Saint Lucia	7.3	12.7	2.1
Saint Vincent	2.4	3.9	0.9
Samoa	5.4	8.7	2.2
Sao Tome and Principe	3.1	4.2	2.1
Saudi Arabia	3.4	4.6	1.7
Senegal	11.8	20.3	5.2
Serbia	10.9	17.3	5.2
Seychelles	8.3	15.0	2.1
Sierra Leone	16.1	18.2	14.2
Singapore	7.9	11.1	4.9
Slovakia	10.1	18.4	2.6
Slovenia	13.3	22.4	4.5
Solomon Islands	5.9	8.5	3.2
Somalia	8.3	11.5	5.4
South Africa	12.8	21.7	5.1
South Sudan	6.1	8.3	4.1
Spain	6.1	9.3	3.1
Sri Lanka	14.2	23.3	6.2
Sudan	9.5	14.5	4.6
Suriname	23.2	36.1	10.9
Sweden	11.7	15.8	7.4
Switzerland	11.3	15.8	6.9
Syria	2.4	3.8	1.1
Tajikistan	3.3	5.0	1.7
Thailand	12.9	21.4	4.8
Macedonia	6.2	9.7	3.0
Timor-Leste	6.4	9.0	3.7
Togo	16.6	22.7	10.9
Tonga	4.0	5.2	2.9
Trinidad and Tobago	12.9	21.9	4.3
Tunisia	3.2	4.4	2.2
Turkey	7.2	11.3	3.2

(continued)

Table 1 (continued)

Country	Both sexes	Male	Female
Turkmenistan	7.2	11.0	3.7
Uganda	20.0	21.2	18.7
Ukraine	18.5	34.5	4.7
United Arab Emirates	2.7	3.5	0.8
United Kingdom	7.6	11.9	3.5
Tanzania	9.6	14.3	5.4
USA	13.7	21.1	6.4
Uruguay	16.5	26.8	7.1
Uzbekistan	7.4	10.3	4.6
Vanuatu	5.4	8.1	2.7
Venezuela	3.8	6.6	1.2
Vietnam	7.0	10.8	3.4
Yemen	9.8	13.4	6.2
Zambia	11.3	17.5	6.2
Zimbabwe	19.1	29.1	11.1

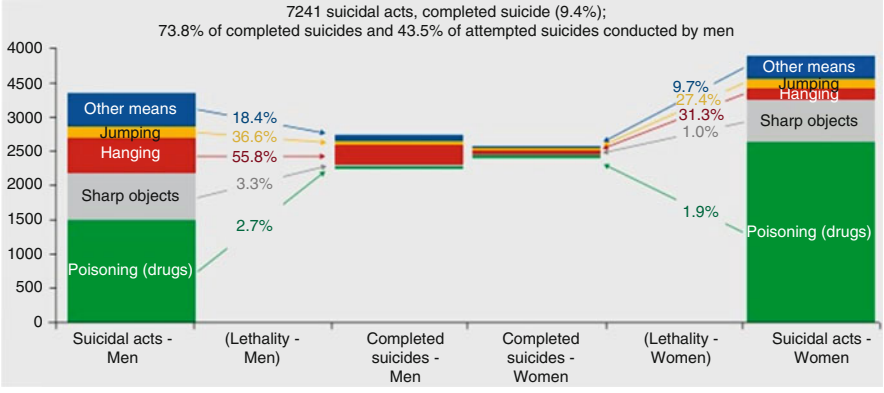
1. Africa: Suicide research in this continent is limited by a lack of systematic data collection; available data are from 60% of African population, which represents less than one third of African countries. The more common methods for suicide are hanging and poisoning, although across studies rates substantially varied (hanging 8–70%; poisoning 8–83%). Firearms are also a common method in some countries (range 0–32%). For both, suicides and suicide attempts, women have higher poisoning rates than males, whereas men tend to use violent methods such as hanging and firearms (Mars et al. 2014).
2. The Americas: Here, a distinction should be made between the United States and the rest of the countries, where significant differences are found in the last group. In the United States, suicides broadly occur by firearms in both genders (61% in males and 36% in females); furthermore poisoning represents 31% of suicide in women. In the other American countries, both men and women completed suicide by poisoning with pesticides (from 0.4% in Canada to 86% in El Salvador for men; from 1% in Canada to 95% in El Salvador for women) and hanging (from 8% in El Salvador to 77% in Chile for men; from 5% in Nicaragua to 63% in Chile for women) (Ajdacic-Gross et al. 2008; Fox et al. 2018).
3. Asia: globally, hanging and poisoning are the prevalent methods. In Hong Kong, people choose hanging (23% in men versus 48% in women) and falls (43% in men versus 23% in women). In the rest of Asian countries, females commit suicide by hanging (from 26% in South Korea to 60% in Japan) or by pesticide poisoning (from 4% in Japan to 43% in South Korea) (Jordans et al. 2014). China deserves a special mention, since more women die by suicide and poisoning with pesticides is the primary method, which reflects the role of women in rural China and the accessibility to pesticides here (Fox et al. 2018).

4. Europe: Globally, most common method, in both males and females, is hanging, except in Swiss males, were is firearm use. Men use firearm to commit suicide in second place in Finland, Norway, France, Austria, and Croatia (21–27%), while women poisoning or fall themselves. Slightly differences were found between countries (Bachmann 2018).
5. Australia and New Zealand: Hanging predominates in both male (45%) and females (36%), followed by firearms (12% in men versus 11% in women) and poisoning representing the third method of suicide (27% in men versus 20% in women) (Australian Institute of Health and Welfare 2014).

While globally these data point out how women who die by suicide choose methods with the same lethality than men, the same is not found for suicide attempts, and it has been shown how females survive suicide attempts more often than males because they use less lethal means (Cibis et al. 2012). Various studies developed in Europe illustrate these facts and explain the gender paradox by males choosing more lethal suicide methods and, in a minor extent, by a higher lethality of men's suicidal acts, even using the same method than women (Freeman et al. 2017)

This is graphically shown in Fig. 3 from the work of Hegerl with data from OSPI-Europe project ("Optimising Suicide Prevention Programmes and their Implementation in Europe") (Hegerl et al. 2009), where it is shown how women preferably choose drug overdose for attempting suicide with a less lethal outcome compared with men using poisoning with drugs as well. The same applies to all other methods, emphasizing that hanging is the most effective method in both men and women, with more lethal results in men (Hegerl 2016).

In Europe, more than 95% of people attempting suicide by poisoning survive, representing a low lethality of intoxications (Mergl et al. 2015). Nevertheless, in most low- and middle-income countries, poisoning is made with pesticides and other substances not available in Europe and more lethal than substances used in "European poisonings." So, many of the suicidal poisonings in these countries result



Open access graphic (Hegerl 2016)

Fig. 3 Differences in suicide methods in men and women

in a lethal outcome and tend to equal male/female suicide rates or even exceed them, as is in the case of China, where in rural areas young women have easy access to pesticides.

5 Gender Differences in Suicide Risk and Protective Factors

This section reviews the main suicide risk factors from a gender perspective, focusing on facts and findings that try to provide evidence for clarifying the gender paradox in suicide. Thus, we consider the known risk factors such as sociodemographic factors, sexual orientation, religiosity, suicide family history, previous suicide attempts, mental disorders, medical conditions, childhood trauma and stressful life events, help-seeking and coping strategies, and biological risk factors.

Here, it is essential to point out that a synergistic relationship is usually found between risk factors. Suicide is never the consequence of a single cause, and complex explanatory models have been proposed to better understand the path toward suicide behaviors (Turecki and Brent 2016; van Heeringen 2012). The stress-diathesis model proposes that direct or proximal stressors interact with distal risk factors (neurobiological and psychological susceptibilities) to predict suicidal behaviors (van Heeringen 2012). Thus, the interpretation of each single risk factor listed in this chapter should be cautiously made.

5.1 Demographics Risk Factors

Age is a relevant moderator factor for gender differences for suicide. Although no gender differences in suicide rates are found under the age of 14 (Fox et al. 2018), there is a consistent tendency for suicide rates to increase with age (Bertolote and Fleischmann 2002). Suicide is the second leading cause of death among 15–19-year-old females worldwide (Saxena et al. 2014) and the first one in Southeast Asia among young females aged 15–29 (Jordans et al. 2014). In most countries, suicide risk is highest in older males, and in younger females, the risk for suicide attempt is highest (Naghavi 2019; Saxena et al. 2014). Overall, female suicide rates are relatively stable with increasing age, whereas for males suicide rates increase with age, tend to plateau in midlife, and reach the highest point with men aged over 75. This final late-life peak is most evident for certain countries such as the United States, France, and Germany (Kiely et al. 2019).

Being unemployed, retired, and single were all significant risk factors for suicide in men, with no effect in females, although it should be acknowledged as a limitation that most studies on this topic come from Europe (Qin et al. 2000; Tóth et al. 2014).

In women, having a young child is identified as a protective factor (Qin et al. 2003). These differences could be explained by gender differences in burdensomeness perception according to Interpersonal-Psychological Theory of Suicidal Behavior (Donker et al. 2014).

Regarding employment, it should be noted how certain professions are more related to suicide, and a higher risk is found in physicians in most countries, more in female than in male doctors (Schernhammer and Colditz 2004). Also, female nurses have a high risk (Agerbo et al. 2002). In these professional groups, a crucial factor involved in the high rates of suicide is access to methods (Agerbo et al. 2002).

5.2 Sexual Orientation

Minority sexual orientations (i.e., being gay, lesbian, or bisexual) have been linked to suicidal behavior. Some authors propose that "minority stress theory" may explain this relationship, as homosexual or bisexual people are frequently exposed to external stressors (more significant stigma, discrimination, or victimization) or even internal stressors such as internalized homophobia that may predispose to suicidal behavior (Miranda-Mendizábal et al. 2017). However, as death records do not routinely include sexual orientation, there are no accurate rates of completed suicide in people with minority sexual orientation (Haas et al. 2011). Some researchers have approached this issue through psychological autopsies studies and have generally concluded that minority sexual orientations are not over-represented among suicide victims; nevertheless, these studies have limitations, especially small samples, and results should be cautiously interpreted (Haas et al. 2011).

It should be noted that suicide behavior in minority sexual orientations is more prevalent in young people (Haas et al. 2011). A meta-analysis of longitudinal studies in youths found that sexual orientation is significantly associated with suicide attempts, but the relationship is not clear for completed suicide, as until then it has been explored in only one longitudinal study. Regarding gender differences, sexual orientation was found to be an independent risk factor for suicide attempts among males, more than among females (Miranda-Mendizábal et al. 2017).

5.3 Religiosity

Beliefs and personal values strongly influence a possible decision to commit suicide. Higher levels of religiosity across the main religions (Christianity, Hinduism, Islam, and Judaism) are historically related to decreased suicide risk (Gearing and Alonzo 2018). Many factors have been postulated to be involved in this protective effect: religious beliefs, involvement in public religious practices by church attendance, moral objections to suicide, lower aggression level among religious individuals, spirituality, or less substance abuse (Kralovec et al. 2018).

To the best of our knowledge, no specific studies taking into account the gender role on the influence of religiosity in death by suicide have been developed. Regarding suicide thoughts and suicidal behavior, different studies have shown how religion in women seems to be a stronger protective effect than in men, either in the general population (Neeleman et al. 1997; Neeleman and Lewis 1999; Rasic et al. 2011), in clinical samples (Kralovec et al. 2018), or in special populations such as high-risk pregnant women (Benute et al. 2011). Only in one study, developed in college students, no significant interactions between gender, religiosity, and suicide ideation were found (Taliaferro et al. 2009).

5.4 Family History

It is well known that family history of suicide increases suicide risk independent of family psychiatric history, and this seems to be stronger in women than in men (Qin et al. 2003). Similarly, family transmission of suicide risk is especially important when suicide happens on the maternal side (Agerbo et al. 2002). Nevertheless, to the best of our knowledge, no systematic research has been developed on this topic.

5.5 Previous Suicide Attempts

The most robust predictor for complete suicide are previous suicide attempts. The effect with 38% of women who completed suicide had a previous suicidal behavior in men, the figure rises to 62% (Ayuso-Mateos et al. 2012).

5.6 Mental Disorders

The majority of deaths by suicide are related to underlying mental diseases, with depression on the top (Bertolote et al. 2004; Bertolote and Fleischmann 2002; Too et al. 2019). It is commonly recognized that over 90% of people who die by suicide had a psychiatric diagnosis and even higher figures (98%) are found in an extensive review of 15,629 cases (Bertolote and Fleischmann 2002). Among all diagnoses, mood disorders were found in 30.2% of suicides, followed by substance use disorders (17.6%), schizophrenia (14.1%), and personality disorders (13.0%) (Bertolote et al. 2004). Globally, females suffer more frequently mental disorders than males (Balta et al. 2019), and also gender differences are known for the more prevalent disorders involved in suicides. Here is again the gender paradox: women suffer more from mental disorders while more men die by suicide.

In psychological autopsies, it is shown that affective disorders prevail in suicide in both genders. Substance use and schizophrenia are more common in male

suicides, whereas in anorexia nervosa, most of patients who died by suicide are women (Hawton 2000).

Next, we summarize information regarding gender differences for the most frequent disorders underlying deaths by suicide: mood disorders (depression and bipolar disorders), substance use disorders, schizophrenia (and psychosis in general), and personality disorders. Also, it should be taken into account that comorbidity of mental disorders increases the suicide risk (Cavanagh et al. 2003).

Concerning **depression**, it is known that it doubles the risk of suicide in the 90 days after hospital discharge (Olfson et al. 2016). As women suffer from a major depressive disorder 2–3 more times than men (Alonso et al. 2000; Kessler et al. 1993), it might be expected more deaths by suicide in women than in men. Possible explanations for this paradox are the different expressions of depression in men and women and the interaction of depression with other risk factors such as alcohol use in men (Lenz et al. 2019). Although not recognized in classification systems, a "male depressive syndrome," widely supported by population studies and meta-analyses, has been proposed (Oliffe et al. 2019; Wålinder and Rutz 2001). This male depression would be a distinct clinical phenotype characterized by a range of externalizing symptoms not captured by diagnostic criteria and, consequently, underdiagnosed and undertreated (Genuchi 2015; Martin et al. 2013). Thus, depressive men are more likely than women to present irritability, anger, aggression, substance misuse, low impulse control, risk-taking, impulsivity, and over-involvement in work (Oliffe et al. 2019), and this depression appearance seems to be mainly influenced by men adjustment to masculine gender role norms (Genuchi and Valdez 2015). It should be noted how these "male traits" of depression are by themselves known suicide risk factors.

In **bipolar disorder**, a strong association with suicide has been found. In a large Danish register, the absolute risk of suicide in bipolar patients after their first hospitalization was around 8% for men and 5% for women (Nordentoft et al. 2011). The gender paradox of suicide is also present for bipolar disorders; however it might be less intense for bipolar disorder than for the general population (Beyer and Weisler 2016). The group of patients with higher suicide risk are young men in an early phase of the illness, especially those who have made a previous suicide attempt, those abusing alcohol, and those recently discharged from the hospital (Jamison 2000; Simpson and Jamison 1999). Among the risk factors specifically related to suicide in bipolar disorder, depressive polarity of the most recent mood episode, as well as depressive polarity of first episode, had the strongest association (Schaffer et al. 2015); this finding illustrates the gender paradox once again, as women tend to have a depressive polarity throughout the illness course.

In psychological autopsy studies, in 19% to 63% of suicides, there were found **substance use disorders (SUD)**, mostly **alcohol use disorders** (Schneider 2009), more commonly in male than in female suicides (Hawton 2000). However, only a limited number of observational studies have reported gender differences in SUD and suicide; therefore, in a recent meta-analysis on SUD and suicide, it was not possible to carry out a meta-analysis risk of suicide by gender (Poorolajal et al. 2016). Specifically for alcohol use, there is evidence from different studies on the

association of male gender, alcohol use, and suicide attempts (Boenisch et al. 2010). Acute alcohol use, or alcohol intoxication, deserves special mention, as it is related to suicide by itself (Bachmann 2018); according to a gender-stratified analysis (Kaplan et al. 2013), acute intoxication in deaths by suicide was more frequent in males than in females.

Comorbidity of SUD and other mental disorders seems to confer a heightened risk of suicide via impulsivity, hostility, and violence (Vijayakumar et al. 2011); all these are characteristically masculine traits (Lenz et al. 2019). Thus, although in the absence of evidence from meta-analysis (Poorolajal et al. 2016), the role of male gender should be taken into account in assessing the risk of suicide in men with mental disorders who also use drugs, especially alcohol.

In **schizophrenia,** male gender is traditionally considered a risk factor for suicide (Popovic et al. 2014). Therefore, the gender pattern of suicide in schizophrenia is similar to general population, and most studies have found higher suicide rates in men than in women (Hawton et al. 2005; Lester 2006), but differences between sex seem to be less marked than in general population (Carlborg et al. 2010) and there even are studies reporting no gender differences (Carlborg et al. 2008; Reutfors et al. 2009). The risk of suicide is highest within the first year after being diagnosed (Nordentoft et al. 2015), but in first-episode psychosis, the traditional gender pattern of suicide is not always found (Austad et al. 2015). Finally, in early-onset psychosis, which is psychosis starting before the age of 18, gender is not a consistent predictor of suicidality (Díaz-Caneja et al. 2015).

Personality disorders represent a high-risk group for suicide with 15% of inpatient and almost 12% of outpatient suicides (Bachmann 2018). Among personality disorders, in **borderline personality disorder (BPD)**, the association with suicide behavior is clear, even included as a diagnostic criterion (Vera-Varela et al. 2019). Nevertheless, in BPD, gender differences in suicidal behavior have been scarcely studied (Sher et al. 2019). In a recent meta-analysis of prospective studies, mean suicide rate ranged from 2% to 5%, but the effect of moderators, including gender, could not be studied due to the heterogeneity among studies (Álvarez-Tomás et al. 2019). Again, while most of BPD patients are women (Silberschmidt et al. 2015), almost 70% of BPD patients who completed suicide are men (Doyle et al. 2016); but contrary to general population, in BPD there are no gender differences in the proportion of suicide attempters or in lifetime number of suicide attempts (Sher et al. 2019). The second highest suicide risk group in personality disorders is **narcissistic personality disorder** (Bachmann 2018), but gender differences have not been studied in this subgroup.

5.7 Medical Conditions

The prevalence of suicide and suicide attempts is elevated not only in individuals with psychiatric illness but also in the context of physical health problems. Ultimately, any chronic disease may be associated with an elevated risk of suicide. An

essential issue in chronic physical illness is disability, which leads to an increase in suicidality. Studies show a variety of chronic diseases related to increased risk for suicide: chronic pain, heart disease, chronic obstructive pulmonary disease, stroke, cancer, congestive heart failure, and asthma (Bachmann 2018). Research also suggests that suffering from multiple physical health conditions confers an even greater risk for suicide (Juurlink et al. 2004).

In suicidality linked to cancer, a review has studied gender differences, concluding that also the gender paradox appears in this population. Thus, although there are exceptions, most studies found that suicide risk is higher in men than women (Robson et al. 2010). In other medical conditions, no systematic research on suicide gender differences has been found.

Nevertheless, an important issue to be taken into account when studying the relationship between somatic diseases and suicide is the role of comorbid psychiatric conditions. Many authors suggest that this comorbidity is what confers a higher risk for suicide in somatic diseases (Qin et al. 2014)

5.8 Childhood Trauma and Stressful Life Events

Suicide attempts and death by suicide are more frequent in people exposed to traumatic events in childhood compared with the general population, and this happens in both males and females (Zatti et al. 2017).

Concerning childhood trauma in a general sense, that is all kind of childhood trauma without distinctions. Some studies have found that suicidality is higher in women who have suffered childhood trauma than in men (Angst et al. 2014), but few works have separately study genders, so there is a lack of strong evidence (Zatti et al. 2017). In particular diagnoses, a recent review on the impact of gender and childhood abuse in psychosis found that women who suffered childhood abuse reported more suicide attempts compared to men (Comacchio et al. 2019).

The role of early sexual abuse on suicide and suicidal behavior has been extensively studied, and there is strong evidence about this relationship (Devries et al. 2014). Gender differences have been analyzed in at least two reviews. The first one is made with cross-sectional data, supporting previous knowledge of an increased odd of suicide in people (men and women) who have suffered childhood sexual abuse, and although sexual abuse was more frequent among females, the association between abuse and suicide attempts was higher in males (Rhodes et al. 2011). The second review is a meta-analysis of longitudinal studies (Devries et al. 2014) that found only two works which separately analyzed genders: in one of them, authors found higher risk of suicide attempts in males versus females (Brezo et al. 2008); the other revealed higher risk of death by suicides in females versus males (Cutajar et al. 2010). These findings are quite interesting, as in people who have suffered sexual abuses during childhood the gender paradox of suicide seems to be reversed.

While childhood trauma is a distal risk factor in explicative models of suicide, life stressors would be a proximal factor also playing a role in the suicide pathway.

Regarding life stressors, differences between genders are described; while men are more likely to experience different types of trauma, except for sexual and violent trauma, women tend to engage more in suicidal behaviors (Ásgeirsdóttir et al. 2018). Similarly, different types of stressors are more frequent according to gender; women tend to react to relational problems such as breakups and men to economic or work-related issues (Shaik et al. 2017). Here, traditional masculinity seems to play a critical role.

5.9 Coping Strategies and Help-Seeking

Men tend to respond to emotional stress with externalizing strategies like risk-taking, aggression, or substance use. Anger is also a negative emotion that men are culturally allowed to show. As previously exposed in this chapter, these coping strategies are related to traditional masculine traits, and, similarly, conformity to masculine norms is linked to a lower probability of help-seeking, as to be strong, resilient, and in control, also identified as male traits (Lenz et al. 2019; Seidler et al. 2016). Men often deny illness, suppress negative feelings, and refuse to admit depressive symptoms, waiting until late before seeking help (Oliffe and Phillips 2008). Thus, men are less likely than women to use healthcare services in general and mental healthcare services in particular; furthermore, men who look for help tend to delay service-seeking, to be reluctant to disclose health concerns, and worst to comply medical recommendations (Fox et al. 2018).

Help-seeking process involves, in addition to the initial act of seeking help, the patient's experience in consultation and subsequent treatment; and the effects of compliance with traditional male norms may also interfere with the therapy process, resulting in difficulties of attendance, compromise, or a non-stable therapeutic alliance (Seidler et al. 2016).

Nevertheless, contrary to the frequent assumption that men's engagement in help-seeking behaviors is rare, a recent review found that men do seek help if it is accessible, appropriate, and engaging (Seidler et al. 2016). This should be taken into account for designing resources tailored according patient gender.

Finally, it also should be noted that men tend to use emergency psychiatric services more than other healthcare facilities (Bachmann 2018). This situation turns emergency departments in critical spots for suicide treatment interventions, and when men with suicidal crisis attend to emergency departments, clinicians should make a special effort to initiate interventions in order to promote their commitment in a therapeutic plan.

5.10 Biological Risk Factors

A biological basis for suicide is known throughout brain post-mortem studies, genomic studies, and neuroimaging studies. Around 50% of suicide risk due to

diathesis is inherited, and this percentage might be higher in females compared to males (van Heeringen and Mann 2014). Despite a large number of studies on biological risk factors for suicide, the knowledge of biological mechanisms underpinning in suicide completion is limited, and the studies focusing on gender differences are scarce.

Genetic differences by gender have been reported in suicide in different samples: in a Portuguese sample, the 5-HTR6 gene 268 C/T SNP has a role in male suicide but not in females (Azenha et al. 2009); in a Japanese sample, men who died by suicide had a lower frequency of the minor allele of a single SNP in the NOS1 gene compared to controls and suicide in women (Cui et al. 2010). Also, a dysfunction in the serotoninergic system is probably the most consistent biological risk factor for suicide, and this is connected with aggression and violence, both considered male traits (Lenz et al. 2019).

Studies on the biology of suicide from a gender perspective often focus on the main biological difference between men and women: sexual hormones. From this perspective, an attractive explanatory model of suicide, "the androgen model of suicide completion" (Lenz et al. 2019), has been proposed. The authors of this model posit that taking into account that male gender is a specific risk factor for suicide, androgen effects might be implicated in the suicidal process and numerous studies are presented showing direct and indirect evidence that increased prenatal androgen levels and also increased androgen activity in adulthood are involved in death by suicide.

The fact that male traits such as aggression, violence, and impulsivity are related to suicide *s*upports the role of androgens in suicide (Lenz et al. 2019). Similarly, the finding of women attempting suicide more frequently during the follicular phase when there are higher testosterone levels also advocates this hypothesis (Baca-Garcia et al. 2010).

As far as biological factors are concerned, probably the most relevant fact to take into account is the role of the interaction of different distal suicide risk factors in the onset of epigenetic mechanisms. In this sense, the work of Turecki et al. is particularly enlightening when it states how sexual hormone activity and early-life stressful events interact and lead to a dysregulation of hypothalamic-pituitary-adrenal (HPA) axis which is known to be involved in suicide (Turecki 2014).

6 Toward a Tailored Prevention According to Gender

Suicide prevention programs include multilevel strategies to address population and individual suicide risk factors and generally include public awareness campaigns, training of community "gatekeepers," and educational initiatives for GPs (Saxena et al. 2014). In this section, we will focus on initiatives that take place in clinical settings, even though population and public health approaches are essential to suicide prevention.

Previous reviews of the effectiveness of prevention programs have recommended the development of tailor-made interventions for specific risk groups (Zalsman et al. 2016). As previously shown throughout this chapter, men are a particular risk group, but gender differences in response to preventive strategies have received little research attention, and specifical interventions focus in men are scarce (Struszczyk et al. 2019). The specificity of certain risk factors in men suggests that there is a need for specific interventions focusing on male factors. Some suggestions, based on previous research, are proposed below.

First of all, depression plays a crucial role in suicide behavior, and depression in men is poorly understood and, consequently, underdiagnosed and undertreated (Olfson et al. 2016). Results of Gotland study highlight how, after an educational program to enhance GP detection of depression, the overall rate of suicide decreased by 60%, but this change was related to female suicide reduction, whereas suicide males were not affected (Rutz et al. 1995). In many men, depression is manifested atypically, and their distress is undetected by the existing diagnostic tools (Seidler et al. 2016). All these findings reflect the necessity of specific tools for screening depression in men and changes in the training of GPs and mental health professionals, including a gender perspective. In response to the first requirement, some specific tools have been developed, such as the Gotland Male Depression Scale (Zierau et al. 2002) or the Masculine Depression Scale (Magovcevic and Addis 2008), the last one divided depression symptoms into internalizing and externalizing (e.g., aggression and irritability), with externalizing symptoms being more representative of depression in men.

The expression of traditional masculinity is closely related to the manifestation of depression in men (Wide et al. 2011), and depressive symptoms are contrary to male ideals, such as feelings of control, stoicism, strength, and success (Seidler et al. 2016). This contradiction usually causes men not to seek help and instead to have feelings of shame or weakness (Seidler et al. 2016). Here, psychotherapeutic and social approaches to redefine masculinity are useful. Reframing traditional male roles to a fluid and flexible masculinity according to contexts allows to cope better with depression and mental health problems (Seidler et al. 2016).

Furthermore, this masculinity, along with the difficulty to recognize depression in men, contributes to delay help-seeking (Seidler et al. 2016). Additionally, when men finally reach mental health services, therapies should be tailored according to most men's preferences. Studies have shown that men tend to prefer interventions based on problem-solving, short-term therapies, and group-based treatment options (Oliffe and Phillips 2008; Seidler et al. 2016). Gender differences in verbal abilities and the resistance of many men to share emotional problems may make talking therapies less attractive to some men (Hawton 2000). Also, sharing experiences with other suicide survivors has been shown to be helpful (Seidler et al. 2016).

Not only the therapeutic approach seems to be important for men but also the environment. Thus, many men demand less formal settings. In this sense, the use of interventions that promote social interaction and informal community-based support centers is highly valued by men. This is of particular interest in young men. Nevertheless, ultimately, and in order to avoid this stigmatization of mental health

facilities by men, incorporating mental health promotion strategies into the educational curriculum from a young age might be a solution (Seidler et al. 2016). Also, this strategy could help men to be more open and to recognize and express their feelings, helping to normalize the need for psychiatric care.

References

Agerbo E, Nordentoft M, Mortensen PB (2002) Familial, psychiatric, and socioeconomic risk factors for suicide in young people: nested case-control study. BMJ 325:74

Ahmed K, Mohan RA, Bhugra D (2007) Self-harm in South Asian women: a literature review informed approach to assessment and formulation. Am J Psychother 61:71–81

Ajdacic-Gross V, Weiss MG, Ring M, Hepp U, Bopp M, Gutzwiller F, Rössler W (2008) Methods of suicide: international suicide patterns derived from the WHO mortality database. Bull World Health Organ 86:726–732

Alonso J, Angermeyer MC, Bernert S, Bruffaerts R, Brugha TS, Bryson H, de Girolamo G, Graaf R, Demyttenaere K, Gasquet I, Haro JM, Katz SJ, Kessler RC, Kovess V, Lépine JP, Ormel J, Polidori G, Russo LJ, Vilagut G, Almansa J, Arbabzadeh-Bouchez S, Autonell J, Bernal M, Buist-Bouwman MA, Codony M, Domingo-Salvany A, Ferrer M, Joo SS, Martínez-Alonso M, Matschinger H, Mazzi F, Morgan Z, Morosini P, Palacín C, Romera B, Taub N, Vollebergh WAM, ESEMeD/MHEDEA (2000) Investigators, European Study of the Epidemiology of Mental Disorders (ESEMeD) project (2004) Prevalence of mental disorders in Europe: results from the European Study of the Epidemiology of Mental Disorders (ESEMeD) project. Acta Psychiatr Scand Suppl:21–27

Álvarez-Tomás I, Ruiz J, Guilera G, Bados A (2019) Long-term clinical and functional course of borderline personality disorder: a meta-analysis of prospective studies. Eur Psychiatry 56:75–83

Angst J, Hengartner MP, Rogers J, Schnyder U, Steinhausen H-C, Ajdacic-Gross V, Rössler W (2014) Suicidality in the prospective Zurich study: prevalence, risk factors and gender. Eur Arch Psychiatry Clin Neurosci 264:557–565

Ásgeirsdóttir HG, Valdimarsdóttir UA, Þorsteinsdóttir ÞK, Lund SH, Tomasson G, Nyberg U, Ásgeirsdóttir TL, Hauksdóttir A (2018) The association between different traumatic life events and suicidality. Eur J Psychotraumatol 9

Austad G, Joa I, Johannessen JO, Larsen TK (2015) Gender differences in suicidal behaviour in patients with first-episode psychosis. Early Interv Psychiatry 9:300–307

Australian Institute of Health and Welfare (2014) Suicide and hospitalised self-harm in Australia: trends and analysis. Canberra, ACT, Australian Institute of Health and Welfare

Ayuso-Mateos JL, Baca-García E, Bobes J, Giner J, Giner L, Pérez V, Sáiz PA, Ruiz JS (2012) Recommendations for the prevention and management of suicidal behaviour. Revista de Psiquiatría y Salud Mental (English Edition) 5:8–23

Azenha D, Alves M, Matos R, Santa JF, Silva B, Cordeiro C, Vieira DN, Ambrósio AM (2009) Male specific association between the 5-HTR6 gene 267C/T SNP and suicide in the Portuguese population. Neurosci Lett 466:128–130

Baca-Garcia E, Diaz-Sastre C, Ceverino A, Perez-Rodriguez MM, Navarro-Jimenez R, Lopez-Castroman J, Saiz-Ruiz J, de Leon J, Oquendo MA (2010) Suicide attempts among women during low estradiol/low progesterone states. J Psychiatr Res 44:209–214

Bachmann S (2018) Epidemiology of suicide and the psychiatric perspective. Int J Environ Res Public Health 15:1425

Balta G, Dalla C, Kokras N (2019) Women's psychiatry. Adv Exp Med Biol 1192:225–249

Benute GRG, Nomura RMY, Jorge VMF, Nonnenmacher D, Fráguas R Jr, de Lucia MCS, Zugaib M (2011) Risk of suicide in high risk pregnancy: an exploratory study. Rev Assoc Med Bras 57:583–587

Bertolote JM, Fleischmann A (2002) Suicide and psychiatric diagnosis: a worldwide perspective. World Psychiatry 1:181–185

Bertolote JM, Fleischmann A, De Leo D, Wasserman D (2004) Psychiatric diagnoses and suicide: revisiting the evidence. Crisis 25:147–155

Beyer JL, Weisler RH (2016) Suicide behaviors in bipolar disorder: a review and update for the clinician. Psychiatr Clin North Am 39:111–123

Bjegovic-Mikanovic V, Abousbie ZAS, Breckenkamp J, Wenzel H, Broniatowski R, Nelson C, Vukovic D, Laaser U (2019) A gap analysis of SDG 3 and MDG 4/5mortality health targets in the six Arabic countries of North Africa: Egypt, Libya, Tunisia, Algeria, Morocco, and Mauritania. Libyan J Med 14:1607698

Boenisch S, Bramesfeld A, Mergl R, Havers I, Althaus D, Lehfeld H, Niklewski G, Hegerl U (2010) The role of alcohol use disorder and alcohol consumption in suicide attempts--a secondary analysis of 1921 suicide attempts. Eur Psychiatry 25:414–420

Brezo J, Paris J, Vitaro F, Hébert M, Tremblay RE, Turecki G (2008) Predicting suicide attempts in young adults with histories of childhood abuse. Br J Psychiatry 193:134–139

Canetto SS (2008) Women and suicidal behavior: a cultural analysis. Am J Orthopsychiatry 78:259–266

Canetto SS, Sakinofsky I (1998) The gender paradox in suicide. Suicide Life Threat Behav 28:1–23

Carlborg A, Jokinen J, Jönsson EG, Nordstrom A-L, Nördstrom P (2008) Long-term suicide risk in schizophrenia spectrum psychoses: survival analysis by gender. Arch Suicide Res 12:347–351

Carlborg A, Winnerbäck K, Jönsson EG, Jokinen J, Nordström P (2010) Suicide in schizophrenia. Expert Rev Neurother 10:1153–1164

Cavanagh JTO, Carson AJ, Sharpe M, Lawrie SM (2003) Psychological autopsy studies of suicide: a systematic review. Psychol Med 33:395–405

Chen Y-Y, Chien-Chang Wu K, Yousuf S, Yip PSF (2012) Suicide in Asia: opportunities and challenges. Epidemiol Rev 34:129–144

Cibis A, Mergl R, Bramesfeld A, Althaus D, Niklewski G, Schmidtke A, Hegerl U (2012) Preference of lethal methods is not the only cause for higher suicide rates in males. J Affect Disord 136:9–16

Clayton JA, Tannenbaum C (2016) Reporting sex, gender, or both in clinical research? JAMA 316:1863

Comacchio C, Lasalvia A, Ruggeri M (2019) Current evidence of childhood traumatic experiences in psychosis – focus on gender differences. Psychiatry Res 281:112507

Cui H, Supriyanto I, Asano M, Ueno Y, Nagasaki Y, Nishiguchi N, Shirakawa O, Hishimoto A (2010) A common polymorphism in the 3'-UTR of the NOS1 gene was associated with completed suicides in Japanese male population. Prog Neuro-Psychopharmacol Biol Psychiatry 34:992–996

Cutajar MC, Mullen PE, Ogloff JRP, Thomas SD, Wells DL, Spataro J (2010) Suicide and fatal drug overdose in child sexual abuse victims: a historical cohort study. Med J Aust 192:184–187

Devries KM, Mak JYT, Child JC, Falder G, Bacchus LJ, Astbury J, Watts CH (2014) Childhood sexual abuse and suicidal behavior: a meta-analysis. Pediatrics 133:e1331–e1344

Díaz-Caneja CM, Pina-Camacho L, Rodríguez-Quiroga A, Fraguas D, Parellada M, Arango C (2015) Predictors of outcome in early-onset psychosis: a systematic review. NPJ Schizophr 1:14005

Donker T, Batterham PJ, Van Orden KA, Christensen H (2014) Gender-differences in risk factors for suicidal behaviour identified by perceived burdensomeness, thwarted belongingness and acquired capability: cross-sectional analysis from a longitudinal cohort study. BMC Psychol 2:20

Doyle M, While D, Mok PLH, Windfuhr K, Ashcroft DM, Kontopantelis E, Chew-Graham CA, Appleby L, Shaw J, Webb RT (2016) Suicide risk in primary care patients diagnosed with a personality disorder: a nested case control study. BMC Fam Pract 17:106

Dunn AG, Coiera E, Mandl KD, Bourgeois FT (2016) Conflict of interest disclosure in biomedical research: a review of current practices, biases, and the role of public registries in improving transparency. Res Integr Peer Rev 1:1

Fausto-Sterling A (2008) Sexing the body: gender politics and the construction of sexuality. Hachette, London

Fox KR, Millner AJ, Mukerji CE, Nock MK (2018) Examining the role of sex in self-injurious thoughts and behaviors. Clin Psychol Rev 66:3–11

Freeman A, Mergl R, Kohls E, Székely A, Gusmao R, Arensman E, Koburger N, Hegerl U, Rummel-Kluge C (2017) A cross-national study on gender differences in suicide intent. BMC Psychiatry 17

Gearing RE, Alonzo D (2018) Religion and suicide: new findings. J Relig Health 57:2478–2499

Genuchi M (2015) Anger and hostility as primary externalizing features of depression in college men. Int J Mens Health 14:113–128

Genuchi MC, Valdez JN (2015) The role of anger as a component of a masculine variation of depression. Psychol Men Masculinity 16:149–159

Haas AP, Eliason M, Mays VM, Mathy RM, Cochran SD, D'Augelli AR, Silverman MM, Fisher PW, Hughes T, Rosario M, Russell ST, Malley E, Reed J, Litts DA, Haller E, Sell RL, Remafedi G, Bradford J, Beautrais AL, Brown GK, Diamond GM, Friedman MS, Garofalo R, Turner MS, Hollibaugh A, Clayton PJ (2011) Suicide and suicide risk in lesbian, gay, bisexual, and transgender populations: review and recommendations. J Homosex 58:10–51

Hawton K (2000) Sex and suicide: gender differences in suicidal behaviour. Br J Psychiatry 177:484–485

Hawton K, van Heeringen K (2009) Suicide. Lancet 373:1372–1381

Hawton K, Sutton L, Haw C, Sinclair J, Deeks JJ (2005) Schizophrenia and suicide: systematic review of risk factors. Br J Psychiatry 187:9–20

Hegerl U (2016) Prevention of suicidal behavior. Dialogues Clin Neurosci 18:183–190

Hegerl U, Wittenburg L, Arensman E, Van Audenhove C, Coyne JC, McDaid D, van der Feltz-Cornelis CM, Gusmão R, Kopp M, Maxwell M, Meise U, Roskar S, Sarchiapone M, Schmidtke A, Värnik A, Bramesfeld A (2009) Optimizing suicide prevention programs and their implementation in Europe (OSPI Europe): an evidence-based multi-level approach. BMC Public Health 9:428

Jamison KR (2000) Suicide and bipolar disorder. J Clin Psychiatry 61:47–51

Jordans MJD, Kaufman A, Brenman NF, Adhikari RP, Luitel NP, Tol WA, Komproe I (2014) Suicide in South Asia: a scoping review. BMC Psychiatry 14:358

Juurlink DN, Herrmann N, Szalai JP, Kopp A, Redelmeier DA (2004) Medical illness and the risk of suicide in the elderly. Arch Intern Med 164:1179–1184

Kaplan MS, McFarland BH, Huguet N, Conner K, Caetano R, Giesbrecht N, Nolte KB (2013) Acute alcohol intoxication and suicide: a gender-stratified analysis of the national violent death reporting system. Inj Prev 19:38–43

Kessler RC, McGonagle KA, Swartz M, Blazer DG, Nelson CB (1993) Sex and depression in the national comorbidity survey. I: Lifetime prevalence, chronicity and recurrence. J Affect Disord 29:85–96

Kiely KM, Brady B, Byles J (2019) Gender, mental health and ageing. Maturitas 129:76–84

Kralovec K, Kunrath S, Fartacek C, Pichler E-M, Plöderl M (2018) The gender-specific associations between religion/spirituality and suicide risk in a sample of Austrian psychiatric inpatients. Suicide Life Threat Behav 48:281–293

Lenz B, Röther M, Bouna-Pyrrou P, Mühle C, Tektas OY, Kornhuber J (2019) The androgen model of suicide completion. Prog Neurobiol 172:84–103

Lester D (2006) Sex differences in completed suicide by schizophrenic patients: a meta-analysis. Suicide Life Threat Behav 36:50–56

Lopez-Castroman J, Blasco-Fontecilla H, Courtet P, Baca-Garcia E, Oquendo MA (2015) Are we studying the right populations to understand suicide? World Psychiatry 14:368–369

Magovcevic M, Addis ME (2008) The masculine depression scale: development and psychometric evaluation. Psychol Men Masculinity 9:117–132

Mars B, Burrows S, Hjelmeland H, Gunnell D (2014) Suicidal behaviour across the African continent: a review of the literature. BMC Public Health 14:606

Martin LA, Neighbors HW, Griffith DM (2013) The experience of symptoms of depression in men vs women: analysis of the national comorbidity survey replication. JAMA Psychiat 70:1100–1106

Mergl R, Koburger N, Heinrichs K, Székely A, Tóth MD, Coyne J, Quintão S, Arensman E, Coffey C, Maxwell M, Värnik A, van Audenhove C, McDaid D, Sarchiapone M, Schmidtke A, Genz A, Gusmão R, Hegerl U (2015) What are reasons for the large gender differences in the lethality of suicidal acts? An epidemiological analysis in four European countries. PLoS One 10: e0129062

Miranda-Mendizábal A, Castellví P, Parés-Badell O, Almenara J, Alonso I, Blasco MJ, Cebrià A, Gabilondo A, Gili M, Lagares C, Piqueras JA, Roca M, Rodríguez-Marín J, Rodríguez-Jiménez T, Soto-Sanz V, Vilagut G, Alonso J (2017) Sexual orientation and suicidal behaviour in adolescents and young adults: systematic review and meta-analysis. Br J Psychiatry 211:77–87

Naghavi M (2019) Global, regional, and national burden of suicide mortality 1990 to 2016: systematic analysis for the Global Burden of Disease Study 2016. BMJ 364:l94

Narang P, Sarai SK, Aldrin S, Lippmann S (2018) Suicide among transgender and gender-nonconforming people. Prim Care Companion CNS Disord 20

Neeleman J, Lewis G (1999) Suicide, religion, and socioeconomic conditions. An ecological study in 26 countries, 1990. J Epidemiol Community Health 53:204–210

Neeleman J, Halpern D, Leon D, Lewis G (1997) Tolerance of suicide, religion and suicide rates: an ecological and individual study in 19 Western countries. Psychol Med 27:1165–1171

Nordentoft M, Mortensen PB, Pedersen CB (2011) Absolute risk of suicide after first hospital contact in mental disorder. Arch Gen Psychiatry 68:1058–1064

Nordentoft M, Madsen T, Fedyszyn I (2015) Suicidal behavior and mortality in first-episode psychosis. J Nerv Ment Dis 203:387–392

O'Carroll PW, Berman AL, Maris RW, Moscicki EK, Tanney BL, Silverman MM (1996) Beyond the tower of babel: a nomenclature for suicidology. Suicide Life Threat Behav 26:237–252

Olfson M, Wall M, Wang S, Crystal S, Liu S-M, Gerhard T, Blanco C (2016) Short-term suicide risk after psychiatric hospital discharge. JAMA Psychiat 73:1119–1126

Oliffe JL, Phillips MJ (2008) Men, depression and masculinities: a review and recommendations. J Men's Health 5:194–202

Oliffe JL, Rossnagel E, Seidler ZE, Kealy D, Ogrodniczuk JS, Rice SM (2019) Men's depression and suicide. Curr Psychiatry Rep 21:103

Poorolajal J, Haghtalab T, Farhadi M, Darvishi N (2016) Substance use disorder and risk of suicidal ideation, suicide attempt and suicide death: a meta-analysis. J Public Health 38:e282–e291

Popovic D, Benabarre A, Crespo JM, Goikolea JM, González-Pinto A, Gutiérrez-Rojas L, Montes JM, Vieta E (2014) Risk factors for suicide in schizophrenia: systematic review and clinical recommendations. Acta Psychiatr Scand 130:418–426

Qin P, Agerbo E, Westergård-Nielsen N, Eriksson T, Mortensen PB (2000) Gender differences in risk factors for suicide in Denmark. Br J Psychiatry 177:546–550

Qin P, Agerbo E, Mortensen PB (2003) Suicide risk in relation to socioeconomic, demographic, psychiatric, and familial factors: a national register-based study of all suicides in Denmark, 1981-1997. Am J Psychiatry 160:765–772

Qin P, Hawton K, Mortensen PB, Webb R (2014) Combined effects of physical illness and comorbid psychiatric disorder on risk of suicide in a national population study. Br J Psychiatry 204:430–435

Rasic D, Kisely S, Langille DB (2011) Protective associations of importance of religion and frequency of service attendance with depression risk, suicidal behaviours and substance use in adolescents in Nova Scotia, Canada. J Affect Disord 132:389–395

Reutfors J, Brandt L, Jönsson EG, Ekbom A, Sparén P, Osby U (2009) Risk factors for suicide in schizophrenia: findings from a Swedish population-based case-control study. Schizophr Res 108:231–237

Rhodes AE, Boyle MH, Tonmyr L, Wekerle C, Goodman D, Leslie B, Mironova P, Bethell J, Manion I (2011) Sex differences in childhood sexual abuse and suicide-related behaviors. Suicide Life Threat Behav 41:235–254

Robson A, Scrutton F, Wilkinson L, MacLeod F (2010) The risk of suicide in cancer patients: a review of the literature. Psycho-Oncology 19:1250–1258

Rutz W, von Knorring L, Pihlgren H, Rihmer Z, Wålinder J (1995) Prevention of male suicides: lessons from Gotland study. Lancet 345:524

Saxena S, Krug EG, Chestnov O, World Health Organization (2014) Preventing suicide: a global imperative. World Health Organization, Geneva

Schaffer A, Isometsä ET, Azorin J-M, Cassidy F, Goldstein T, Rihmer Z, Sinyor M, Tondo L, Moreno DH, Turecki G, Reis C, Kessing LV, Ha K, Weizman A, Beautrais A, Chou Y-H, Diazgranados N, Levitt AJ, Zarate CA, Yatham L (2015) A review of factors associated with greater likelihood of suicide attempts and suicide deaths in bipolar disorder: part II of a report of the international society for bipolar disorders task force on suicide in bipolar disorder. Aust N Z J Psychiatry 49:1006–1020

Schernhammer ES, Colditz GA (2004) Suicide rates among physicians: a quantitative and gender assessment (meta-analysis). Am J Psychiatry 161:2295–2302

Schiebinger L, Stefanick ML (2016) Gender matters in biological research and medical practice. J Am Coll Cardiol 67:136–138

Schneider B (2009) Substance use disorders and risk for completed suicide. Arch Suicide Res 13:303–316

Seidler ZE, Dawes AJ, Rice SM, Oliffe JL, Dhillon HM (2016) The role of masculinity in men's help-seeking for depression: a systematic review. Clin Psychol Rev 49:106–118

Shaik S, Rajkumar RP, Menon V, Sarkar S (2017) Gender, life events, and depression: an exploratory study. Indian J Psychol Med 39:330–335

Sharmin Salam S, Alonge O, Islam MI, Hoque DME, Wadhwaniya S, Ul Baset MK, Mashreky SR, El Arifeen S (2017) The Burden of suicide in rural Bangladesh: magnitude and risk factors. Int J Environ Res Public Health 14

Shekhani SS, Perveen S, Hashmi D-E-S, Akbar K, Bachani S, Khan MM (2018) Suicide and deliberate self-harm in Pakistan: a scoping review. BMC Psychiatry 18:44

Sher L, Rutter SB, New AS, Siever LJ, Hazlett EA (2019) Gender differences and similarities in aggression, suicidal behaviour, and psychiatric comorbidity in borderline personality disorder. Acta Psychiatr Scand 139:145–153

Silberschmidt A, Lee S, Zanarini M, Schulz SC (2015) Gender differences in borderline personality disorder: results from a multinational, clinical trial sample. J Personal Disord 29:828–838

Silverman MM, Berman AL, Sanddal ND, O'carroll PW, Joiner TE (2007) Rebuilding the tower of Babel: a revised nomenclature for the study of suicide and suicidal behaviors. Part 2: suicide-related ideations, communications, and behaviors. Suicide Life Threat Behav 37:264–277

Simon M, Chang E-S, Zeng P, Dong X (2013) Prevalence of suicidal ideation, attempts, and completed suicide rate in Chinese aging populations: a systematic review. Arch Gerontol Geriatr 57:250–256

Simpson SG, Jamison KR (1999) The risk of suicide in patients with bipolar disorders. J Clin Psychiatry 60:53–56

Struszczyk S, Galdas PM, Tiffin PA (2019) Men and suicide prevention: a scoping review. J Ment Health 28:80–88

Taliaferro LA, Rienzo BA, Pigg RM, Miller MD, Dodd VJ (2009) Spiritual well-being and suicidal ideation among college students. J Am Coll Heal 58:83–90

Too LS, Spittal MJ, Bugeja L, Reifels L, Butterworth P, Pirkis J (2019) The association between mental disorders and suicide: a systematic review and meta-analysis of record linkage studies. J Affect Disord 259:302–313

Tóth MD, Ádám S, Birkás E, Székely A, Stauder A, Purebl G (2014) Gender differences in deliberate self-poisoning in Hungary: analyzing the effect of precipitating factors and their relation to depression. Crisis 35:145–153

Turecki G (2014) The molecular bases of the suicidal brain. Nat Rev Neurosci 15:802–816
Turecki G, Brent DA (2016) Suicide and suicidal behaviour. The Lancet 387:1227–1239
van Heeringen K (2012) Stress–diathesis model of suicidal behavior. In: Dwivedi Y (ed) The neurobiological basis of suicide, frontiers in neuroscience. CRC Press/Taylor & Francis, Boca Raton
van Heeringen K, Mann JJ (2014) The neurobiology of suicide. Lancet Psychiatry 1:63–72
Vera-Varela C, Giner L, Baca-García E, Barrigón ML (2019) Trastorno límite de personalidad: el reto del manejo de las crisis. Psiquiatr Biol 26:1–6
Vijayakumar L, Kumar MS, Vijayakumar V (2011) Substance use and suicide. Curr Opin Psychiatry 24:197–202
Wålinder J, Rutz W (2001) Male depression and suicide. Int Clin Psychopharmacol 16:S21
WHO, Defining Sexual Health (2019). WHO. http://www.who.int/reproductivehealth/topics/sexual_health/sh_definitions/en/ Accessed 27 Nov 2019
WHO, Suicide Data (2019) WHO. http://www.who.int/mental_health/prevention/suicide/suicideprevent/en/ Accessed 10 Nov 2019
WHO, Suicide Rates (per 100 000 population) (2019) WHO. http://www.who.int/gho/mental_health/suicide_rates_male_female/en/ Accessed 10 Nov 2019
Wide J, Mok H, McKenna M, Ogrodniczuk JS (2011) Effect of gender socialization on the presentation of depression among men: a pilot study. Can Fam Phys 57:e74–e78
Zalsman G, Hawton K, Wasserman D, van Heeringen K, Arensman E, Sarchiapone M, Carli V, Höschl C, Barzilay R, Balazs J, Purebl G, Kahn JP, Sáiz PA, Lipsicas CB, Bobes J, Cozman D, Hegerl U, Zohar J (2016) Suicide prevention strategies revisited: 10-year systematic review. Lancet Psychiatry 3:646–659
Zatti C, Rosa V, Barros A, Valdivia L, Calegaro VC, Freitas LH, Ceresér KMM, da Rocha NS, Bastos AG, Schuch FB (2017) Childhood trauma and suicide attempt: a meta-analysis of longitudinal studies from the last decade. Psychiatry Res 256:353–358
Zierau F, Bille A, Rutz W, Bech P (2002) The Gotland male depression scale: a validity study in patients with alcohol use disorder. Nord J Psychiatry 56:265–271