



UNIVERSITA' DEGLI STUDI DI PAVIA
DIPARTIMENTO DI SCIENZE DEL SISTEMA NERVOSO E DEL
COMPORTAMENTO
CORSO DI LAUREA MAGISTRALE IN PSICOLOGIA

**L'AFFIDABILITÀ DIAGNOSTICA NEI DISTURBI
DELL'UMORE: UNA REVISIONE SISTEMATICA
CON METANALISI**

:

RELATORE:
Prof. Davide Liccione

CORRELATRICE:
Dr.ssa Carla Laura Di Forti

Tesi di Laurea di
Filippo Gava
Matricola n. 522911

Anno Accademico 2023/2024

Sommario

<i>ABSTRACT</i>	4
<i>RIASSUNTO</i>	4
<i>INTRODUZIONE</i>	6
1. DISTURBI DELL'UMORE	6
1.1. PROSPETTIVA STORICA SUI DISTURBI DELL'UMORE	7
1.2. CLASSIFICAZIONE DEI DISTURBI DELL'UMORE	11
1.3. LE RECENTI REVISIONI DEI DISTURBI DELL'UMORE: DSM-5-TR E ICD-11	16
1.3.1. DISTURBI DEPRESSIVI	16
1.3.2. DISTURBI BIPOLARI	18
1.4. LIMITI DEI SISTEMI NOSOGRAFICI: CATEGORIZZARE VS INDIVIDUALIZZARE	19
2. L'APPROCCIO DIMENSIONALE AI DISTURBI DELL'UMORE: LE CONDIZIONI SPETTRALI	21
1. TEORIZZAZIONE DELLO SPETTRO BIPOLARE	24
1.1.1. I TEMPERAMENTI COME PRODROMI: DA KRAEPELIN A TELLENBACH	25
1.1.2. SOFT BIPOLAR SPECTRUM: CONTINUUM TRA I DUE POLI MANIACALE E DEPRESSIVO	28
2.2. TEORIZZAZIONE DELLO SPETTRO DELL'UMORE	31
2.3. CRITICITÀ E PROSPETTIVE FUTURE	32
3. DIAGNOSI DEI DISTURBI DELL'UMORE: SFIDE E LIMITI	34
3.1. METODOLOGIE DIAGNOSTICHE	34
3.2. AFFIDABILITÀ E VALIDITÀ DELLA DIAGNOSI	37
3.2.1. I <i>FIELD TRIALS</i> SULLE SINDROMI AFFETTIVE	39
3.2.3. <i>BIAS</i> COGNITIVI DEL CLINICO E DEL PAZIENTE	41
<i>CAPITOLO 1. LO STUDIO</i>	45
1.1. OBIETTIVI	45
1.2. MATERIALI E METODI	47
1.3. RISULTATI	50
<i>CAPITOLO 2. DISCUSSIONE</i>	59
<i>CONCLUSIONE</i>	64
<i>BIBLIOGRAFIA</i>	66
<i>SITOGRAFIA</i>	81

ABSTRACT

The literature suggests a significant delay between the onset of bipolar disorder and accurate diagnosis. Bipolar disorder can be difficult to recognize and misdiagnosis is common, often influenced by clinician and patient bias. The results of the DSM-5 field trial (Regier *et al.* 2013) on the diagnosis of major depressive disorder were unsatisfactory. Research into the characteristics of the diagnostic process for mood disorders is therefore crucial for improvement.

After an introduction to the diagnostic process in affective disorders, the paper presents a systematic review with meta-analysis of inter-rater reliability in mood disorders, conducted according to the PRISMA guidelines. A total of 1923 articles were screened, and 38 were included in the study. Data analysis showed a substantial agreement, with $K = 0.74$ (CI 0.71-0.78). Contrary to expectations, the mean kappa for bipolar disorder was 0.76. The funnel plot shows the presence of publication bias. The heterogeneity between trials was very high, with a Higgins index of 95%. Further research is needed to reduce the heterogeneity and the publication bias.

RIASSUNTO

La letteratura indica un significativo ritardo tra l'insorgenza del disturbo bipolare e una diagnosi accurata. Il disturbo bipolare può essere difficile da riconoscere e la diagnosi errata è molto comune, spesso influenzata dai *bias* cognitivi del clinico e del paziente. I risultati del *field trial* sul DSM-5 (Regier *et al.* 2013) per la diagnosi di disturbo depressivo maggiore sono stati insoddisfacenti. Le attuali difficoltà nella diagnosi dei disturbi dell'umore richiedono dunque un impegno maggiore nella ricerca.

Dopo aver introdotto il processo diagnostico nei disturbi affettivi, l'elaborato presenta una revisione sistematica con metanalisi sull'accordo tra valutatori dei disturbi dell'umore, condotta secondo le linee guida PRISMA. Sono stati visionati 1923 abstract di articoli e 38 sono stati gli studi inclusi. Dai risultati emerge un accordo sostanziale con $K = 0,74$ (CI 0,71-0,78). Contrariamente alle aspettative, la media di kappa per il disturbo bipolare è stata di 0,76. Il grafico a imbuto mostra la presenza di *bias* di pubblicazione. L'eterogeneità tra gli

studi è risultata essere molto elevata ($I^2= 95\%$). Ulteriori studi sono necessari al fine di ridurre l'eterogeneità e il *bias* di pubblicazione.

INTRODUZIONE

1. DISTURBI DELL'UMORE

La comprensione dei disturbi dell'umore, come il disturbo depressivo maggiore e il disturbo bipolare, si è certamente evoluta nel corso del tempo. Tuttavia, nonostante oggi disponiamo di una varietà di trattamenti psico-farmacologici nonché di strumenti diagnostici sempre più affidabili, in relazione ai disturbi dell'umore ci troviamo di fronte a dati inquietanti. L'Institute for Health Metrics and Evaluation (IHME) stima che nel 2021 la prevalenza del disturbo depressivo è stata del 4,1% della popolazione (332 milioni di persone)¹. Si calcola che in Italia superino i 2,8 milioni (5,4% delle persone di 15 anni e più) coloro che hanno sofferto di depressione nel corso del 2015². Una metanalisi rileva che le donne in gravidanza e le puerpere presentano un rischio significativamente più elevato di depressione, con una prevalenza di circa il 10% nei paesi a reddito elevato; percentuale che aumenta nei paesi a basso/medio reddito (Woody *et al.* 2017). Sempre secondo l'IHME nel 2021 le persone con disturbo bipolare diagnosticato erano 38,5 milioni. La metanalisi di Dagani *et al.* (2017) ha identificato come la durata di malattia non trattata (DUI), per il disturbo bipolare, si attestasse intorno ai 6 anni, peggiorando significativamente la prognosi. Ciò accade anche a causa della difficoltà di approdare ad una diagnosi corretta. Il riconoscimento di questo disturbo, noto per l'alto rischio suicidario, è difficile e molto frequente è l'errata diagnosi (Stensland *et al.* 2010, Bruchmüller & Meyer 2009). A causa della presenza degli episodi depressivi, esso viene spesso diagnosticato come tale solo con il sopraggiungere di un episodio maniacale nonostante la presenza di precedenti episodi ipomaniacali non riconosciuti (Hirschfeld *et al.* 2004). Risulta necessario, pertanto, impiegare sempre maggiori sforzi di ricerca sui disturbi dell'umore per migliorarne la diagnosi; il miglioramento della diagnosi costituisce infatti il primo passo essenziale per il corretto trattamento.

Nell'introduzione di questo elaborato si darà una visione generale dei disturbi dell'umore per quanto concerne il processo diagnostico. Anzitutto verrà presentata una breve storia del disturbo depressivo e del disturbo bipolare concentrandosi sull'evoluzione della

¹ Cfr. Institute of Health Metrics and Evaluation. Global Health Data Exchange (GHDx): <https://vizhub.healthdata.org/gbd-results/>

² Cfr. il comunicato stampa dell'ISTAT: <https://www.istat.it/comunicato-stampa/la-salute-mentale-nelle-varie-fasi-della-vita/>

classificazione nosografica degli stessi; verranno poi esposti i criteri diagnostici delle recenti versioni dell'ICD e del DSM. Nel secondo paragrafo si prenderanno in considerazione i modelli dimensionali spettrali prevalenti dei disturbi affettivi, partendo dai pregressi storici degli stessi e soffermandosi sul ruolo del temperamento; nel terzo paragrafo, verranno presentati gli strumenti diagnostici più validati per la diagnosi dei disturbi dell'umore, verranno esposti i risultati di alcuni *field trials* condotti sulle recenti versioni dei manuali diagnostici e verranno infine evidenziati i principali *bias* in cui può incorrere il processo diagnostico. A partire dal capitolo 1 si andrà a presentare la metanalisi da noi condotta sull'accordo tra valutatori (*inter-rater reliability*) nei disturbi dell'umore.

1.1. PROSPETTIVA STORICA SUI DISTURBI DELL'UMORE

“A Oxford Dictionary of Psychology” alla voce *mood*, riporta la seguente definizione di umore: “a temporary but relatively sustained and pervasive affective state, often contrasted in psychology and psychiatry with a more specific and short-term emotion” (Colman 2015). Solo parzialmente dipendente dalle circostanze esterne, l'umore influenza profondamente la percezione di sé stessi, dell'ambiente circostante e il modo in cui una persona si relaziona con gli altri. Le sue fluttuazioni – entro una certa norma - sono fisiologiche perché servono a rispondere alle richieste dell'ambiente. Queste fluttuazioni modulano il senso di iniziativa e le risposte della persona, permettendole di adottare diversi comportamenti in base ai cambiamenti delle circostanze ambientali. Negli stati francamente patologici queste fluttuazioni diventano troppo ampie e indipendenti per durata ed intensità dagli eventuali stimoli esterni. In questo contesto si configurano i disturbi dell'umore: l'umore diventa abnormemente depresso o elevato (Di Paolo 2008).

Con Ippocrate, l'interesse per l'osservazione dello stile comportamentale ha portato alla definizione della cosiddetta “teoria umorale” pressoché dominante fino al XVII secolo. Ippocrate e la sua scuola distinguono quattro temperamenti a cui corrispondono quattro sostanze endogene, gli umori: sangue, flegma, bile gialla e bile nera. L'idea soggiacente a questa teoria, sviluppata dopo Ippocrate da Galeno, era la seguente: nell'equilibrio dei quattro fluidi fondamentali consiste la salute, e dal variare del loro rapporto va determinandosi il temperamento (sanguigno, flemmatico, collerico, malinconico), mentre

dal prevalere eccessivo di uno sugli altri tre hanno origine gli stati morbosi³. Dal grado di prevalenza di uno degli umori sugli altri si distingue la semplice “disposizione personale” (temperamento) dallo stato patologico (Harnic *et al.* 2011). Secondo la scuola ippocratica l’eccesso di bile nera determinava la depressione - da cui il termine melanconia: *mélas* “nero” e *cholé* “bile” -, mentre la mania era determinata da una esagerata presenza di bile gialla.

La prima identificazione del disturbo bipolare viene solitamente ricondotta ad Areteo di Cappadocia (I secolo d. C), il quale ipotizzò l’esistenza di uno stretto legame tra depressione ed eccitamento euforico, considerandole due facce di una medesima condizione. Marneros e Angst (2002) sintetizzano le posizioni di Areteo in questo modo:

(a) Melanconia e mania hanno la stessa eziologia, ovvero sono disturbi della funzione del cervello e di alcuni altri organi.

(b) La mania è un peggioramento della melanconia.

(c) La mania è la controparte fenomenologica della melanconia.

Se durante il medioevo i disturbi mentali escono dall’ambito della medicina - venendo attribuiti a “cause demoniache” -, è solo tra il XVII e il XVIII secolo che l’osservazione della alternanza tra mania e melancolia da parte di Thomas Willis, Giovanni Battista Morgagni, Antoine Charles Lorry e Vincenzo Chiarugi, gettò le basi per la comprensione moderna del disturbo bipolare, superando l’attribuzione della malattia a cause soprannaturali⁴.

Kendler (2020a) in una revisione della letteratura basata su 28 testi pubblicati tra il 1780 e il 1900, opera una distinzione in tre fasi del modo di concepire la mania in quel periodo. Dal 1780 fino agli anni '20 dell'Ottocento, la mania era vista come un disturbo del ragionamento/giudizio che si manifestava come follia totale e/o uno stato di furia indifferenziata. Nei successivi 30 anni, il consenso cambiò e la mania fu intesa principalmente come un disturbo dell'umore elevato. Intorno al 1860, l’accordo attorno a questa visione si rompe: per il resto del XIX secolo, il modello della mania basato sull'umore ha concorso con la visione secondo cui la mania derivasse principalmente da processi mentali accelerati (Mendel) e dall’eccitazione (o iperattività) psicomotoria (Griesinger). Ci è sembrato importante fare un breve riferimento a questo dibattito poiché questo continuerà a esercitare la sua influenza sottotraccia. Come lo stesso Kendler (2020a) sottolinea, tra DSM-IV e DSM-5 è stata effettuata una modifica di rilievo: se nel DSM-IV era richiesto nel

³ <https://www.treccani.it/vocabolario/umore/?search=um%C3%B3re%2F>

⁴ Cfr. La pagina curata da Antonio Tundo, “Breve storia del disturbo bipolare”: <https://www.istitutodipsicopatologia.it/breve-storia-del-disturbo-bipolare/>

criterio A “umore elevato espanso o irritabile” ora è richiesto anche “un aumento anomalo e persistente dell’attività finalizzata o dell’energia” (APA 2014); questa modifica richiama il dibattito della seconda metà del XIX secolo, in particolare la posizione di Griesinger.

La storia della mania condivide alcune importanti somiglianze e differenze con la storia della melanconia nello stesso periodo. Lo stesso Kendler (2020b) pubblica una storia della melanconia tra il 1780 e il 1880. Anche in questo caso Kendler distingue tre periodi. Tra la fine del XVIII e l'inizio del XIX secolo, il concetto di melanconia era strettamente legato a visioni precedenti che la consideravano fundamentalmente un disturbo dell'intelletto - una follia parziale - spesso, ma non sempre, accompagnata da tristezza. Il primo allontanamento da questo paradigma si deve a Guislain, che scrisse poco dopo la metà del XIX secolo. Egli definì la melanconia elementare come un disturbo dell'umore e si concentrò sulla categoria trascurata ma significativa della melanconia senza deliri: i pazienti che ne presentavano i sintomi non manifestavano alcuna anomalia dell'intelletto o del giudizio. Nella fase finale, che va dagli anni '60 agli anni '80 del XIX secolo, prevalse l’idea che la melanconia fosse un disturbo primario dell'umore e che i deliri emergessero dall'umore anormale (Kendler 2020b).

Se abbiamo trattato delle concezioni nosografiche ottocentesche della mania e della depressione intesi come disturbi unipolari, altro discorso vale per il disturbo bipolare. Nel 1854, gli psichiatri francesi Jules Baillarger e Jean-Pierre Falret, in modo indipendente l'uno dall'altro, dichiararono di aver scoperto un nuovo disturbo mentale - il disturbo bipolare - a cui attribuirono nomi diversi: “Follia circolare” (Falret) “Follia a doppia forma” (Baillarger), e sarà infatti solo nella seconda metà del Novecento che il “disturbo bipolare” assumerà definitivamente questo nome (Marneros & Angst 2002).

A cavallo il XIX e il XX secolo, il contributo più rilevante è stato dato da quello che viene considerato il padre della nosografia psichiatrica moderna: Emil Kraepelin. È allo psichiatra tedesco che bisogna dare il merito di aver separato i disturbi mentali gravi (psicosi) endogeni in due grandi gruppi: la psicosi maniaco-depressiva e la *dementia praecox*, che con Eugen Bleuler assumerà poi il nome di “schizofrenia”. Kraepelin fissa come criteri differenziali di questi due grandi gruppi psicopatologici l’età di insorgenza, la familiarità, il decorso e l’esito. Rispetto all’esito, secondo Kraepelin, la psicosi maniaco-depressiva aveva una prognosi migliore rispetto all’infausto esito della “demenza precoce”. Il gruppo delle psicosi maniaco depressive aggrega – sulla base dell’ipotesi della medesima alterazione organismica - psicopatologie prima considerate indipendenti, in particolare la mania e la

depressione e la “folie circulaire” di Falret, eliminando così la distinzione fra disturbo unipolare e bipolare (Marneros & Angst 2002). Secondo Kraepelin (1905), i disturbi dell'umore sono parte di un continuum in cui la depressione 'pura' e la mania 'pura' rappresentano gli estremi di uno spettro definito da tre dimensioni psicopatologiche principali: affettività, pensiero e volizione-psicomotricità. Questo modello suggerisce che le tre dimensioni possono variare in modo patologico sia in senso espansivo sia in senso depressivo, permettendo una varietà di disturbi affettivi. Mania e depressione come polarità pure si sviluppano quando le tre dimensioni si muovono assieme nella medesima direzione, mentre le altre condizioni sono considerate da Kraepelin come ‘stati misti’ (Harnic *et al.* 2011).

L’opera di classificazione di Kraepelin è continuata con il suo allievo Wilhelm Weygard, il quale ha descritto dettagliatamente gli stati misti (*Mischzustände*) della psicosi maniaco-depressiva. Secondo Weygard, dalla combinazione dei tre ambiti funzionali della psiche, già descritti dal suo maestro, quali umore, psicomotricità e pensiero, possono derivare sei principali forme di stati misti; tuttavia, le forme descrivibili sono moltissime, trattandosi di un continuum psicopatologico. Agli 'stati misti', intesi come copresenza simultanea di sintomi depressivi e maniacali, Weygard ha affiancato i 'quadri misti' (*Mischbilder*), dove i sintomi di polarità opposta si alternano rapidamente, una condizione simile a quella oggi chiamata “a cicli rapidi” (Harnic *et al.* 2011).

La psicosi maniaco-depressiva introdotta da Kraepelin rimase dominante fino alla metà del XX secolo; ancora presente nel DSM-II (1968), viene definitivamente abolita solo nel DSM-III (1980). A questa svolta nosografica contribuirono prima i lavori di Karl Kleist e Karl Leonhard e, un decennio dopo, i lavori di Jules Angst e di Carlo Perris.

Kleist e Leonhard proposero di rompere in due categorie nosografiche distinte la psicosi maniaco-depressiva: nascono così i “disturbi bipolari”, in cui si alternano e si compenetrano mania e depressione e i “disturbi unipolari”, cioè la “mania pura” e la “melanconia pura”. Tuttavia, la distinzione da loro introdotta rimase quasi sconosciuta alla psichiatria internazionale per diversi anni (Marneros & Angst 2002).

La separazione delle forme unipolari da quelle bipolari fu poi confermata su base genetica da due importanti pubblicazioni indipendenti di Angst (1966) e Perris (1966), ricevendo così il consenso della nosografia internazionale. Entrambi gli studi furono svolti su due gruppi di soggetti diversamente diagnosticati: un gruppo con “psicosi ciclica” - cioè la psicosi maniaco depressiva che verrà ridenominata “disturbo bipolare” - e un gruppo con

“depressione ricorrente”; i due psichiatri osservarono che i famigliari del gruppo con psicosi ciclica presentavano più frequentemente un disturbo mentale rispetto ai famigliari dei soggetti con depressione ricorrente. È stata inoltre individuata una correlazione significativa tra il tipo di disturbo dei pazienti e il tipo di disturbo più prevalente fra i parenti: i parenti del gruppo con psicosi ciclica tendevano a presentare psicosi ciclica, tendenza rilevata anche per il la depressione ricorrente tra i parenti dei soggetti con depressione ricorrente (Angst & Perris 1972). I due psichiatri hanno inoltre mostrato come la "mania unipolare" fosse geneticamente correlata ai disturbi bipolari (Marneros & Angst 2002). Secondo Angst e Perris psichiatri, quindi, se è corretto avere una entità nosografica unipolare per la depressione, non è corretto averne una per la sola mania, perché questa tende a presentarsi con episodi depressivi, ponendo pertanto le basi alla distinzione tra disturbo bipolare di tipo I e di tipo II.

Altro sviluppo importante per gli attuali sistemi nosografici, presente tanto nel DSM-5-TR, quanto nell'ICD-11, è appunto la distinzione tra disturbo bipolare I e disturbo bipolare II introdotta da Fieve & Dunner (1975), approfondita da Dunner nel 1976 (Marneros & Angst 2002) e, infine, accolta nel DSM-IV (APA 1994). Se il disturbo bipolare I è caratterizzato da almeno un episodio maniacale o misto, il disturbo bipolare II è caratterizzato da almeno un episodio ipomaniacale e da un episodio depressivo maggiore (APA 1994). I pazienti affetti da disturbo bipolare di tipo II – da ciò deriva la costituzione dei due tipi - sono stati identificati da Fieve & Dunner (1975) sulla base del criterio dell'ospedalizzazione. Essi, infatti, presentavano gravi episodi depressivi maggiori, alternati a periodi ipomaniacali che non richiedevano l'ospedalizzazione, a differenza dei casi con mania (Perugi *et al.* 2003, Akiskal & Pinto 1999).

1.2. CLASSIFICAZIONE DEI DISTURBI DELL'UMORE

Le macrocategorie diagnostiche presenti nell'ICD-10 per i disturbi dell'umore che andremo a presentare in questo sottoparagrafo sono: episodio maniacale (F30), disturbo affettivo bipolare (F31) e disturbo depressivo ricorrente (F33).

Fra le suddivisioni dell'episodio maniacale (F30) vi sono l'ipomania (F30.0), e la mania senza sintomi psicotici (F30.1). L'OMS indica che per diagnosticare l'episodio maniacale o ipomaniacale questo deve presentarsi come isolato, in assenza di una storia

pregressa di disturbi dell'umore, altrimenti viene indicato il disturbo affettivo bipolare come diagnosi più appropriata. L'ipomania (F30.0) è descritta come un disturbo caratterizzato da

persistente lieve esaltazione dell'umore, aumento di energia e attività e, generalmente, marcato senso di benessere e di efficienza fisica e mentale. Sono spesso presenti aumento della socievolezza, loquacità, invadenza, aumento dell'energia sessuale e riduzione del bisogno di dormire, ma non in misura tale da comportare marcata compromissione lavorativa o rifiuto sociale (Frattura *et al.* 2022, p. 250)

La differenza essenziale tra ipomania e mania sta nella differente intensità dell'umore espanso e dell'attività (si parla infatti di *attività* nell'ipomania e di *iperattività* nella Mania); se inoltre l'ipomania non comporta una marcata compromissione lavorativa o rifiuto sociale nella mania vi è un "comportamento incauto, violento e inappropriato alle circostanze" (Frattura *et al.* 2022, p. 250) tale da compromettere l'ambito lavorativo e/o sociale.

Per quanto riguarda la macrocategoria del disturbo affettivo bipolare, l'OMS indica che esso è caratterizzato da almeno un episodio maniaco o ipomaniaco in aggiunta ad almeno un episodio depressivo oppure da almeno due episodi di mania o ipomania (Frattura *et al.* 2022 p. 251). In sostanza, servono due episodi di alterazione patologica dell'umore (ipomaniacali, maniacali, depressivi o misti), basta che non siano due episodi depressivi; come nel disturbo bipolare I del DSM-5, pertanto non è necessario un episodio depressivo per diagnosticare nell'ICD-10 il disturbo affettivo bipolare, giacché "episodi ripetuti esclusivamente di ipomania o mania sono classificati come bipolari" (Frattura *et al.* p. 251). Sulla revisione della macrocategoria diagnostica dei disturbi bipolari operata dall'ICD-11 rimando al paragrafo 1.3 della presente introduzione.

La descrizione dell'episodio depressivo (F32) dell'ICD-10 è sostanzialmente uguale a quella del DSM-5, pertanto rimandiamo alle pagine successive, in cui viene presentato il disturbo secondo i criteri del DSM-5.

Per quanto riguarda il disturbo depressivo ricorrente (F33), all'interno dell'ICD-10 viene descritto come un disturbo caratterizzato da ripetuti episodi di depressione senza anamnesi positiva di mania, ma in cui

ci possono comunque essere brevi episodi di lieve esaltazione del tono dell'umore e iperattività (ipomania) immediatamente dopo un episodio depressivo, talvolta precipitati dal trattamento antidepressivo (Frattura *et al.* p. 252).

Il disturbo depressivo ricorrente (F33) si distingue, pertanto, dal disturbo affettivo bipolare (F31) per ricorrenti episodi depressivi, sebbene il disturbo depressivo nella sua forma ricorrente non escluda la possibilità di lievi episodi ipomaniacali sottosoglia.

Differentemente dall'ICD-10, che riunisce in un unico gruppo i disturbi dell'umore, il DSM-5 (APA, 2014) separa i disturbi affettivi in due gruppi: “disturbo bipolare e disturbi correlati” e “disturbi depressivi”.

In questo paragrafo andremo a illustrare i criteri diagnostici per i seguenti disturbi che sono i più diagnosticati: disturbo bipolare I; disturbo bipolare II; disturbo depressivo maggiore. A questo fine, riportiamo i criteri diagnostici per l'episodio maniacale, ipomaniacale e depressivo.

L'episodio maniacale è definito dai seguenti criteri:

A. Un periodo definito di umore anormalmente e persistentemente elevato, espanso o irritabile e di aumento anomalo e persistente dell'attività finalizzata o dell'energia, della durata di almeno 1 settimana e presente per la maggior parte del giorno, quasi tutti i giorni (o di qualsiasi durata, se è necessaria l'ospedalizzazione).

B. Durante il periodo di alterazione dell'umore e di aumento di energia o attività, tre (o più) dei seguenti sintomi (quattro, se l'umore è solo irritabile) sono presenti a un livello significativo e rappresentano un cambiamento evidente rispetto al comportamento abituale:

1. Autostima ipertrofica o grandiosità.

2. Diminuito bisogno di sonno (per es., sentirsi riposati/e dopo solo 3 ore di sonno).

3. Maggiore loquacità del solito o spinta continua a parlare.

4. Fuga delle idee o esperienza soggettiva che i pensieri si succedano rapidamente.

5. Distraibilità (cioè attenzione troppo facilmente deviata da stimoli esterni non imo non pertinenti), riferita o osservata.

6. Aumento dell'attività finalizzata (sociale, lavorativa, scolastica o sessuale) o agitazione psicomotoria (cioè attività immotivata non finalizzata).

7. Eccessivo coinvolgimento in attività che hanno un alto potenziale di conseguenze dannose (per es., acquisti incontrollati, comportamenti sessuali sconvenienti o investimenti finanziari avventati).

C. L'alterazione dell'umore è sufficientemente grave da causare una marcata compromissione del funzionamento sociale o lavorativo o da richiedere l'ospedalizzazione per prevenire danni a sé o agli altri, oppure sono presenti manifestazioni psicotiche.

D. L'episodio non è attribuibile agli effetti fisiologici di una sostanza (per es., una sostanza di abuso, un farmaco, un altro trattamento) o a un'altra condizione medica (APA 2014, p. 144).

Rispetto al Criterio A del DSM-IV, come abbiamo già menzionato nel paragrafo precedente, il DSM-5 introduce tanto per l'episodio maniacale quanto per quello ipomaniacale “l'aumento anomalo e persistente dell'attività finalizzata o dell'energia”, assente nel DSM-IV. Nel DSM-5, i criteri per l'episodio ipomaniacale sono molto simili a quelli dell'episodio maniacale sopra riportati, con le seguenti differenze: rispetto al criterio A, la durata dell'umore espanso è inferiore, vengono cioè richiesti almeno 4 giorni al posto di 7; inoltre rispetto al criterio C dell'episodio maniacale, l'episodio ipomaniacale al criterio E specifica che “non è sufficientemente grave da causare una marcata compromissione del funzionamento sociale o lavorativo o da richiedere l'ospedalizzazione” (APA, 2014, p. 145)

– come per l'ICD-10 - anche se l'alterazione dell'umore risulta osservabile dagli altri (criterio D).

I criteri dell'episodio depressivo sono i seguenti:

A. Cinque (o più) dei seguenti sintomi sono stati contemporaneamente presenti durante un periodo di 2 settimane e rappresentano un cambiamento rispetto al precedente livello di funzionamento; almeno uno dei sintomi è 1) umore depresso o 2) perdita di interesse o piacere [...]

1. Umore depresso per la maggior parte del giorno, quasi tutti i giorni, come riportato dall'individuo (per es., si sente triste, vuoto/a, disperato/a) o come osservato da altri (per es., appare lamentoso/a). (Nota: Nei bambini e negli adolescenti l'umore può essere irritabile.)

2. Marcata diminuzione di interesse o piacere per tutte, o quasi tutte, le attività per la maggior parte del giorno, quasi tutti i giorni (come indicato dal resoconto soggettivo o dall'osservazione).

3. Significativa perdita di peso, non dovuta a dieta, o aumento di peso (per es., un cambiamento superiore al 5% del peso corporeo in un mese) oppure diminuzione o aumento dell'appetito quasi tutti i giorni. (Nota: Nei bambini, considerare l'incapacità di raggiungere i normali livelli ponderali.)

4. Insonnia o ipersonnia quasi tutti i giorni.

5. Agitazione o rallentamento psicomotori quasi tutti i giorni (osservabile dagli altri; non semplicemente sentimenti soggettivi di essere irrequieto/a o rallentato/a).

6. Faticabilità o mancanza di energia quasi tutti i giorni.

7. Sentimenti di autosvalutazione o di colpa eccessivi o inappropriati (che possono essere deliranti), quasi tutti i giorni (non semplicemente autoaccusa o sentimenti di colpa per il fatto di essere ammalato/a).

8. Ridotta capacità di pensare o di concentrarsi, o indecisione, quasi tutti i giorni (come impressione soggettiva o osservata da altri).

9. Pensieri ricorrenti di morte (non solo paura di morire), ricorrente ideazione suicidaria senza un piano specifico o un tentativo di suicidio o un piano specifico per commettere suicidio.

B. I sintomi causano disagio clinicamente significativo o compromissione del funzionamento in ambito sociale, lavorativo o in altre aree importanti.

C. L'episodio non è attribuibile agli effetti fisiologici di una sostanza o a un'altra condizione medica (APA, 2014, pp. 145s).

La diagnosi di disturbo depressivo maggiore nel DSM-5, oltre a richiedere la soddisfazione dei criteri per un episodio depressivo necessita il rispetto dei seguenti criteri:

D. Il verificarsi dell'episodio depressivo maggiore non è meglio spiegato dal disturbo schizoaffettivo, dalla schizofrenia, dal disturbo schizofreniforme, dal disturbo delirante o dal disturbo dello spettro della schizofrenia e altri disturbi psicotici con altra specificazione o senza specificazione.

E. Non vi è mai stato un episodio maniacale o ipomaniacale (APA, 2014, p. 187).

Per soddisfare la diagnosi di disturbo Bipolare I, è richiesto almeno un episodio maniacale, nonostante venga evidenziato che spesso l'episodio maniacale è preceduto o seguito da un episodio depressivo o ipomaniacale.

Per la diagnosi di disturbo Bipolare II, il DSM-5, tra gli altri, richiede di soddisfare i seguenti criteri:

- A. sono soddisfatti i criteri per almeno un episodio maniacale [*sic*: ipomaniacale non maniacale] (Criteri A-F sotto “*Episodio ipomaniacale*”, sopra) e per almeno un episodio depressivo maggiore (Criteri A-C sotto “*Episodio depressivo maggiore*”, sopra).
- B. Non vi è mai stato un episodio maniacale (APA, 2014, p. 156, corsivo mio).

L’entità nosografica di Bipolar disease “Mixed episode” del DSM-IV-TR (APA 2000, p. 362) che prevedeva che nello stesso periodo di una settimana fossero soddisfatti quasi tutti i giorni i criteri di episodio maniacale e di episodio depressivo, nel DSM-5 è stata sostituita dallo specificatore “con caratteristiche miste” (Kaltenboek *et al.* 2016). Questo specificatore può essere associato alla diagnosi di episodio maniacale (o ipomaniacale) se vengono soddisfatti almeno 3 dei criteri A e B dell’episodio depressivo⁵, simultaneamente all’episodio maniacale o ipomaniacale. Lo specificatore “con caratteristiche miste” può essere inoltre associato ad una diagnosi di episodio depressivo se vengono soddisfatti almeno 3 dei criteri A e B dell’episodio maniacale/ipomaniacale simultaneamente all’episodio depressivo. In questo caso le caratteristiche miste associate a un episodio depressivo maggiore “rappresentano un significativo fattore di rischio per lo sviluppo di disturbo bipolare I o bipolare II” (APA 2014), è dunque necessario prestare particolare attenzione a queste caratteristiche.

Lo specificatore “con caratteristiche miste” viene distinto nel DSM-5 dallo specificatore “a cicli rapidi”; la caratteristica essenziale del disturbo bipolare con cicli rapidi è il verificarsi di quattro o più episodi dell’umore durante i 12 mesi precedenti (APA 2014, p. 174).

Intendiamo ora mettere in luce le principali differenze fra DSM-5 e i ICD-10. Una di queste, come viene sottolineato da Kaltenboek *et al.* (2016), sta nel fatto che l’ICD-10 permette di diagnosticare singoli episodi maniacali, ipomaniacali e depressivi a differenza del DSM-5, in cui non è possibile diagnosticare gli episodi singoli come tali; inoltre, nell’ICD-10 servono almeno due episodi di alterazione dell’umore (tra cui uno maniacale) per diagnosticare il “disturbo affettivo bipolare, episodio maniacale in atto” (F31.1) mentre nel DSM-5 basta un episodio maniacale. Altra importante differenza sta nel modo di

⁵ Va fatto notare che in realtà in questo caso le caratteristiche depressive dei criteri per diagnosticare l’episodio misto in presenza di un episodio maniacale/ipomaniacale divergono leggermente rispetto a quelli A e B dell’episodio depressivo. Sullo specificatore “con caratteristiche miste” cfr. APA, 2014, pp. 173s; 212s.

suddividere il disturbo bipolare: l'ICD-10 non distingue esplicitamente tra disturbo bipolare I e II – tuttavia nell'ICD-10 il disturbo bipolare II può essere codificato come "Disturbi affettivi bipolari di altro tipo" - differenza con il DSM che, come vedremo, verrà meno nell'ICD-11.

1.3. LE RECENTI REVISIONI DEI DISTURBI DELL'UMORE: DSM-5-TR E ICD-11

Nonostante le ampie modifiche apportate al DSM-5, l'American Psychiatric Association (APA) ha ritenuto opportuno commissionarne una revisione, il DSM-5-TR (APA, 2022b). A questo scopo, sono stati invitati a partecipare molti degli esperti che avevano già collaborato alla stesura del DSM-5. In totale, 200 esperti hanno avuto il compito di revisionare il manuale, dando vita al DSM-5-TR, pubblicato nel 2022 (Bradley *et al.* 2023).

Nel 2007, il Dipartimento di Salute Mentale e Abuso di Sostanze dell'OMS ha nominato un gruppo consultivo internazionale per la revisione del capitolo dell'ICD-10 sui Disturbi mentali, comportamentali e del neurosviluppo. Nel 2019, la 72° Assemblea mondiale della Sanità ha votato per l'adozione dell'ICD-11, che è stato quindi adottato dagli Stati membri dell'OMS a partire dal 1° gennaio 2022.

Nei prossimi due paragrafi metteremo a confronto l'ICD-10 con l'ICD-11 e il DSM-5 con il DSM-5-TR prima in relazione ai disturbi depressivi, quindi in relazione ai disturbi bipolari.

1.3.1. DISTURBI DEPRESSIVI

Nel DSM-5-TR tra le diagnosi oggetto di una revisione vi è stato il “disturbo dell'umore senza specificazione” e il “disturbo depressivo con altra specificazione”.

Il “disturbo dell'umore non specificato” è una categoria nosografica residuale utilizzabile quando si presentano sintomi dell'umore che non soddisfano pienamente i criteri per nessuno dei disturbi nelle classi diagnostiche dei disturbi bipolari o depressivi e per i quali è difficile scegliere tra disturbo Bipolare e altri disturbi correlati senza specificazione

e disturbo depressivo senza specificazione (First *et al.* 2022). A causa dell'assenza di un raggruppamento per i disturbi dell'umore nel DSM-5 e nel DSM-5-TR - a differenza del DSM-IV - il disturbo dell'umore senza specificazione è presente sia nei capitoli sui disturbi depressivi che in quelli sui disturbi bipolari. Ciò allinea il DSM-5-TR con l'ICD-10-CM e con l'ICD-11⁶, che includono anch'essi questa categoria (APA 2022c).

Rispetto al “disturbo depressivo con altra specificazione”, una quarta specificazione è stata introdotta: l’episodio depressivo maggiore sovrapposto. Questa specificazione si applica quando un episodio depressivo maggiore si verifica contemporaneamente a un disturbo psicotico che non ha episodi dell'umore come parte dei suoi criteri diagnostici (APA 2022d).

Veniamo ora alle modifiche introdotte con l’ICD-11. Anzitutto il sotto-capitolo sui “Disturbi dell’umore” ha cambiato nome: l'alternativa "disturbi affettivi" è cioè stata omessa. Un cambiamento sostanziale tra ICD-10 e ICD-11 sta nel fatto che gli Episodi dell'umore, quindi anche l’episodio depressivo, non sono condizioni diagnosticabili indipendentemente dagli altri Disturbi dell’umore (Reed *et al.* 2019), ciò allinea l’ICD-11 al DSM-5 (e DSM-5-TR).

Inoltre, rispetto all’ICD-10 nell'ICD-11 gli Episodi depressivi nei disturbi depressivi e bipolari possono essere descritti in dettaglio utilizzando specificatori che indicano la presenza di sintomi specifici: il qualificatore delle caratteristiche melanconiche, lo specificatore dei sintomi d'ansia, il qualificatore degli attacchi di panico e lo specificatore “seasonal pattern”. Inoltre, gli episodi depressivi possono essere descritti in base alla gravità (lieve, moderata o grave), contrassegnando l’eventuale presenza di sintomi psicotici – non diversamente dall’ICD-10 – e in base allo stato di remissione parziale o completa (Gaebel *et al.* 2020, Krawczyk *et al.* 2020).

Nell’ICD-11 è stato introdotto il Premenstrual Dysphoric Disorder (PMDD), allineando pertanto l’ICD-11 al DSM-5. Il disturbo disforico premenstruale è caratterizzato da una varietà di sintomi umorali, somatici e/o cognitivi gravi che iniziano diversi giorni prima dell'inizio delle mestruazioni, cominciano a migliorare entro pochi giorni e diventano minimi o assenti entro circa una settimana dall'inizio del ciclo. Sebbene la sua collocazione principale nell'ICD-11 sia nel capitolo sulle malattie del sistema genito-urinario, il PMDD è

⁶ Cfr. <https://icd.who.int/browse/2024-01/mms/en#578635574%2Funspecified>

elencato anche nel sottogruppo dei disturbi depressivi – come nel DSM-5 - a causa del suo impatto sull'umore (Reed *et al.* 2019).

1.3.2. DISTURBI BIPOLARI

Come per i disturbi depressivi, anche per i disturbi bipolari il DSM-5-TR ha portato modifiche di piccolo rilievo. Alcune di queste hanno riguardato lo specificatore di gravità “lieve” per l’episodio maniacale e il “disturbo bipolare e altri disturbi correlati con altra specificazione”.

Lo specificatore di gravità "lieve" per l'episodio maniacale è stato modificato a causa di una incoerenza con il criterio C dell'episodio maniacale stesso. Il criterio C infatti richiede che l'alterazione dell'umore sia abbastanza grave da causare una marcata compromissione del funzionamento sociale o lavorativo; pertanto, nel DSM-5-TR il qualificatore "lieve" può essere attribuito all’episodio maniacale nel caso di disturbo bipolare se vengono soddisfatti *solo i criteri minimi dei sintomi* (First *et al.* 2022, APA 2022a).

Nel DSM-5-TR è stata aggiunta, come per i Disturbi depressivi, una (quinta) specificazione per la categoria nosografica "disturbo bipolare e altri disturbi correlati con altra specificazione": l’episodio maniacale sovrapposto (APA2022e).

Un cambiamento fondamentale introdotto dall’ICD-11 è l’introduzione della distinzione - già presente nel DSM a partire dalla sua quarta edizione – tra Bipolar type I e Bipolar type II, che rompe quindi in due categorie nosografiche separate il disturbo affettivo bipolare (F31) dell’ICD-10. Questo è stato fatto in seguito al crescente sostegno della validità di questa distinzione. A supporto di questa differenziazione ci sono infatti evidenze che riguardano la risposta alla monoterapia con antidepressivi, le misurazioni neuropsicologiche, le influenze genetiche e i risultati di neuroimaging (Reed *et al.* 2019).

Nell’ICD-11 è stato separato il sottogruppo dell'ICD-10 dei Disturbi dell'umore persistenti (F34), costituito da Distimia e Ciclotimia nelle relative categorie diagnostiche; inoltre è disponibile lo specificatore “rapid cycling” per le diagnosi di disturbo bipolare⁷, non presente nell’ICD-10. Infine, come per gli Episodi depressivi, le categorie nosografiche singole di episodio maniacale e di episodio ipomaniacale sono state rimosse: di conseguenza,

⁷ <https://icd.who.int/browse/2024-01/mms/en#2021041320>

un episodio maniacale/ipomaniacale può essere diagnosticato solo nel corso di un disturbo bipolare (Krawczyk *et al.* 2020).

1.4. LIMITI DEI SISTEMI NOSOGRAFICI: CATEGORIZZARE VS INDIVIDUALIZZARE

“la psicologia oggettivante, tipica della diagnostica nosografico-descrittiva, non lascia spazio al privato del singolo [...] perché dovendo oggettivare, deve de-individualizzare” (Liccione 2019, p. 66).

Fino ad ora abbiamo svolto una trattazione dei disturbi dell'umore facendo perlopiù riferimento alla tradizione nosografico-descrittiva propria dei sistemi diagnostici. È importante però ora sottolineare i limiti di questi sistemi. Le persone, infatti, non essendo meri oggetti, non possono essere presi (solo) come tali, come inevitabilmente fanno le scienze “neuro-” e i sistemi diagnostici, che tendono a ridurre i soggetti al solo piano biologico – le prime - o a descrivere in modo impersonale le esperienze delle persone – i secondi -. Gli oggetti, in quanto meri enti di natura, sono categorizzabili, le persone non sono (soltanto) categorizzabili perché non sono “meri” enti di natura. A differenza dei meri enti di natura, l'essere umano appartiene a un doppio piano, come sostiene Paul Ricœur, quello del *bios* – e pertanto risulta diagnosticabile quindi categorizzabile – e quello del *logos*, per cui non solo fa esperienza, qualitativamente sempre individuale, ma è costitutivamente portato a raccontarla.

Rispetto agli oggetti (che sussistono, sono semplicemente presenti), quell'ente particolare che è l'esserci, vive in una dimensione temporale che lo “deietta” nel passato ma soprattutto lo proietta nel futuro; in ciò consiste la differenza ontologica di Heidegger⁸. Secondo Heidegger l'esserci è temporalità costituita di passato (gettatezza), presente e futuro (progettualità): in questo senso l'esserci è – per utilizzare una espressione heideggeriana - “progetto-gettato”. Se l'esistenza si caratterizza come apertura al mondo, e l'esserci non è semplicemente presente, perché si progetta (“ha da essere”) ne deriva che l'esistenza non risulta mai categorizzabile, coglibile in modo definitivo. Il racconto del soggetto – prodotto della coscienza riflessiva o narrativa – si modifica costantemente nel corso della sua vita, a

⁸ Per una chiarificazione di questa fondamentale distinzione heideggeriana cfr. il capitolo “La differenza ontologica”, Liccione 2019.

seconda delle nuove possibilità d'azione che di volta in volta gli si presentano⁹. Secondo l'approccio fenomenologico-ermeneutico adottato dalla psicoterapia cognitiva-neuropsicologica, il racconto di sé si modifica perché l'essere umano, pur essendo sempre se stesso, è anche sempre diverso¹⁰.

Categorizzare significa cogliere l'individuo in ciò che di comune ha con altri individui, cioè coglierlo in "terza persona", secondo medesimezza; diagnosticare un soggetto significa considerarlo impersonalmente sotto il profilo dei sintomi e dei segni che condivide con la descrizione della classe nosografica specifica. La diagnosi non permette quindi di comprendere il soggetto rispetto a ciò che egli ha di più proprio, cioè l'esser mio dell'esperienza (ipseità). Soltanto a partire dal racconto della sua esperienza preriflessiva il soggetto – eventualmente aiutato in ciò dal terapeuta - può far emergere in prima persona la sua identità. Individualizzare significa allora permettere all'esperienza preriflessiva di essere raccontata attraverso una narrazione coerente:

Se l'identità deve essere una relazione fra due poli, e se la caratteristica dell'identità personale è il riconoscersi sempre come se stessi, allora entrambi i poli, per formare l'identità, devono avere come specificità ineludibile la sensazione di «proprietà», di appartenenza a Sé... in quanto «se stesso». L'esperienza vissuta e quella raccontata devono condividere un primo e indispensabile livello di identità specifica: l'individuo deve sentirsi l'autore di entrambe [...] uno di questi poli è sicuramente l'ipseità, l'essere mio dell'esperienza. L'altro polo, seguendo Ricoeur, è il racconto di Sé (Liccione 2019, p. 68).

All'interno di questo paragrafo abbiamo considerato i limiti intrinseci ad ogni diagnosi nosografica: essa non permette al clinico di comprendere l'esperienza dell'altro e non permette al paziente di cogliersi in "prima persona", ma aiuta a spiegare la psicopatologia¹¹, oltre a essere essenziale per la comunicazione fra specialisti di discipline diverse. Lungi dall'intenzione di chi scrive sottrarre importanza alla classificazione nosografica: una corretta diagnosi nosografica risulta essenziale a fini terapeutici, soprattutto farmacologici, e di ricerca.

⁹ Questa costante rfigurazione della propria narrazione è coerente con la concezione post-moderna dell'identità. Il costante cambiamento e superamento di sé sono infatti l'impronta dell'identità postmoderna (cfr. Liccione 2019, pp. 65s).

¹⁰ L'identità assoluta (A = A) in questo senso può andare bene per gli oggetti, ma non per l'essere umano – ad esclusione del codice genetico che però coglie soltanto il lato biologico, non quello psicologico - che pur essendo sempre se stesso è sempre diverso cfr. Liccione 2019, pp. 66ss. L'identità personale è una identità specifica tra chi ha vissuto l'esperienza in prima persona e l'autore del racconto di Sé.

¹¹ Secondo il modello cognitivo-neuropsicologico, ciò che distingue una patologia organica (o "storica") da una funzionale (o "non storica") è l'eziopatologia del disturbo. Se i motivi, meglio delle cause organiche possono render conto di una determinata psicopatologia allora parleremo di patologie storiche, se è vero l'opposto invece, parleremo di patologie non storiche; laddove il criterio di distinzione fondamentale è dato dal ruolo che gioca l'identità personale del paziente nel veicolare questo o quel disturbo.

La spiegazione (che coinvolge tutte le discipline “neuro-”) e la comprensione (che si basa sulla fenomenologia e l’ermeneutica) non solo possono essere utilizzate entrambe in ambito psicopatologico, ma devono entrare in relazione dialettica (costruttiva), cercando una traduzione dei reciproci linguaggi. Dato che l’oggetto di studio è lo stesso – l’essere umano - si tratterebbe “di una traduzione fra due linguaggi specialistici, che cercano di dire diversamente la stessa cosa, o meglio quasi la stessa cosa” (Jervolino, 2006, p. 404, citato da Liccione 2019). L’essere umano infatti, come scrive Ricœur, appartiene sia al piano del *bios* sia al piano del *logos*:

Qual è questo essere che rende possibile questa doppia appartenenza del motivo alla forza e al senso, alla natura e alla cultura, al bios e al logos? Bisognerebbe riflettere sulla posizione stessa del corpo nella natura: esso è a un tempo corpo tra i corpi (cosa tra le cose), e modo di esistere di un essere capace di riflettere, di riprendersi e di giustificare la sua condotta. L’uomo è tale proprio perché appartiene al tempo stesso al regime della causalità e a quello della motivazione, quindi della spiegazione e della comprensione. (Ricœur 2003, p. 165, citato da Liccione 2019).

2. L’APPROCCIO DIMENSIONALE AI DISTURBI DELL’UMORE: LE CONDIZIONI SPETTRALI

Nella trattazione di questo e dei prossimi paragrafi andremo a presentare gli approcci dimensionali prevalenti ai disturbi dell’umore privilegiando una prospettiva storico-genetica, anziché trasversale-comparativa. Riteniamo, infatti, che un approccio genetico alla nosografia, permetta meglio di coglierne gli sviluppi, giacché come è evidente dalle diverse edizioni del DSM, gli sviluppi della nosografia – più di quanto accade per le altre scienze – non dipendono solo dagli avanzamenti scientifici, ma anche dal contesto storico-culturale.

Negli ultimi tre decenni, c’è stato un crescente interesse per le forme atipiche, subcliniche e/o sub-soglia dei disturbi unipolari e bipolari (Benvenuti *et al.* 2015). È dunque cresciuta l’attenzione per un approccio dimensionale ai disturbi dell’umore, dato che il modello categoriale veniva considerato troppo restrittivo e rigido; sono nati in questo contesto i modelli spettrali. Secondo questi modelli, identificando precocemente la costellazione fenomenica sottosoglia nei pazienti affetti da disturbo bipolare è possibile offrire degli strumenti prognostici e terapeutici più raffinati che favoriscono un miglioramento globale della qualità della vita del paziente (Harnic *et al.* 2011).

Come abbiamo già accennato nel paragrafo 1.1. Kraepelin osservò che la mania e la depressione non erano entità separate, ma piuttosto due estremità di uno spettro clinico. I suoi studi longitudinali su pazienti maniaco-depressivi evidenziarono diversi elementi chiave: molti pazienti che manifestano sintomi depressivi sviluppavano successivamente sintomi maniacali, e viceversa; altri pazienti affetti da depressione potevano sperimentare episodi di ipomania; alcuni pazienti presentavano un decorso ciclico con alternanza di depressione e stati non chiaramente maniacali, ma con una predisposizione temperamentale simile a chi sperimentava episodi maniacali o ipomaniacali (Kraepelin 1921, p. 2; Akiskal & Pinto 1999). Il superamento della categoria nosografica kraepeliniana della psicosi maniaco-depressiva con il DSM-III ha però portato all'eliminazione delle molte precedenti sfumature presenti nella concezione spettrale della psicosi maniaco-depressiva. La nuova dicotomia disturbo unipolare-disturbo bipolare, come scrivono Benvenuti *et al.* (2014): “ha rappresentato un passo importante non solo per la classificazione ma anche e soprattutto per la comprensione e il trattamento dei disturbi dell'umore” tuttavia – sottolineano gli autori - “già dai primi anni '80 si rese evidente come molti dei soggetti altrimenti diagnosticati come unipolari potevano avere forme di tipo bipolare”. Fu di fatto questa osservazione a portare allo sviluppo del modello dimensionale di spettro bipolare o “soft bipolar spectrum” (Akiskal 1983; Akiskal & Mallya 1987). Akiskal e Pinto (1999) rispetto alla divisione tra disturbo bipolare I e II introdotta nel DSM-IV (APA, 1994) presentarono una ulteriore sottotipizzazione in sette categorie del disturbo bipolare sulla base di una diversa gravità della componente ipomaniacale (cfr. Tabella 1). La condizione bipolare IV è considerata la forma più stabile ed attenuata di ipomania ma è quella sulla quale sono anche stati pubblicati il minor numero di dati (Harnic *et al.* 2011).

Bipolare I	Depressione e mania
Bipolare I e 1/2	Depressione e ipomania protratta
Bipolare II	Depressione e ipomania
Bipolare II e 1/2	Depressione e ciclotimia
Bipolare III	Depressione e ipomania farmaco-indotta
Bipolare III e 1/2	Bipolarità associata all'uso di sostanze stimolanti
Bipolare IV	Depressione e temperamento ipertimico

Tabella 1. Spettro bipolare di Akiskal e Pinto (Harnic *et al.* 2011).

I due psichiatri sostenevano che esisteva una percentuale significativa di pazienti unipolari il cui disturbo dell'umore mostrava caratteristiche bipolari:

This article has emphasized the depressive manifestations of a range of bipolar conditions, which are best regarded as pseudo-unipolar. [...] The main thesis of this article has been that many major depressions in the DSM-IV schema are, in reality, part of the bipolar spectrum (Akiskal & Pinto 1999)¹².

Per gli autori il riconoscimento di caratteristiche bipolari nel disturbo depressivo risulta in particolare fondamentale a fini farmaco-terapeutici.

Il modello di Akiskal mirava a limitare gli errori diagnostici e terapeutici senza tuttavia eliminare l'approccio categoriale e la dicotomia unipolare-bipolare (Benvenuti *et al.* 2014). Pur concordando esplicitamente con Akiskal e Pinto sul fatto che un vasto numero di pazienti cosiddetti unipolari manifestasse sintomi ipomaniacali lievi, Cassano nel 2004 ipotizzava la necessità di eliminare la diagnosi di unipolarità:

we argue that the field would benefit from a unitary and continuous approach to the assessment of both manic/hypomanic and depressive symptoms. On the basis of this unitary conceptualization of mood disorders, we hypothesized that patients with recurrent major depression without discrete lifetime hypomanic episodes would nonetheless report lifetime hypomanic/manic symptoms and that the number of lifetime hypomanic/manic symptoms would be related to the number of lifetime depressive symptoms [...] Our data suggest that unipolar disorder and bipolar disorder are not two discrete and dichotomous phenomena but that mood fluctuations—up and down—are common to both conditions (Cassano *et al.* 2004)¹³.

Nonostante queste parole, ci sembra che Cassano - come Akiskal - non escluda la presenza della depressione unipolare, non solo perché questa è una categoria degli attuali sistemi diagnostici prevalenti, ma anche perché permette di costituire un continuum su cui si collocano la maggior parte dei pazienti (approccio probabilistico)¹⁴. In altri termini, per i modelli spettrali, non è utile diagnosticare una entità nosografica come la depressione pura o l'episodio maniacale puro se la maggior parte dei pazienti, presentando fluttuazioni dell'umore, non manifestano questi disturbi in forma pura; tuttavia, entrambe le forme pure sono necessarie per definire il continuum riscontrabile nella realtà clinica.

¹² “Questo articolo ha messo in evidenza le manifestazioni depressive di una gamma di condizioni bipolari, che è meglio considerare come pseudo-unipolari. [...] La tesi principale di questo articolo è che molte delle depressioni maggiori secondo il DSM-IV sono, in realtà, parte dello spettro bipolare” (traduzione nostra).

¹³ “Sosteniamo che il campo di studi trarrebbe beneficio da un approccio unitario e continuo alla valutazione sia dei sintomi maniacali/ipomaniacali che di quelli depressivi. Basandoci su questa concettualizzazione unitaria dei disturbi dell'umore, abbiamo ipotizzato che i pazienti con depressione maggiore ricorrente, senza episodi ipomaniacali distinti nel corso della vita, riporterebbero comunque sintomi ipomaniacali/maniacali nel corso della loro vita e che il numero di sintomi ipomaniacali/maniacali nel corso della vita sarebbe correlato al numero di sintomi depressivi nel corso della vita [...] I nostri dati suggeriscono che il disturbo unipolare e il disturbo bipolare non sono due fenomeni discreti e dicotomici, ma che le fluttuazioni dell'umore — sia verso l'alto che verso il basso — sono comuni a entrambe le condizioni” (traduzione nostra).

¹⁴ Un continuum è, per definizione, costituito da due estremità opposte; non importa se esse non si ritrovano mai nella realtà clinica, le polarità servono a spiegare la fenomenologia clinica sulla base dell'intervallo graduale e continuo da queste definito.

Nei primi anni 2000, clinici e ricercatori dell'Università di Pisa, in Italia, e delle Università di Pittsburgh, Columbia (New York) e California (San Diego), negli Stati Uniti, hanno promosso lo "Spectrum Project Collaborative Group" (SPCG)¹⁵ con l'obiettivo di creare e validare strumenti in grado di riconoscere l'ampia gamma di fenomeni che circonda le caratteristiche "centrali" di ciascuna categoria di disturbo dell'umore del DSM e superare la classica dicotomia unipolare-bipolare (Benvenuti *et al.* 2014). Nasce quindi in questi anni, grazie ai contributi di Cassano e collaboratori, il costrutto di "spettro dell'umore" che si va ad affiancare a quello di "spettro bipolare" di Akiskal.

2.1. TEORIZZAZIONE DELLO SPETTRO BIPOLARE

Il temperamento è "la disposizione affettiva fondamentale e caratteristica di ogni persona, il tono emotivo di fondo che precede l'esperienza e ad essa predispone" (Maina *et al.* 2003). Il temperamento si differenzia dal carattere: se il primo è legato al corredo genetico e si manifesta fin da piccoli, il carattere deriva dalle esperienze avute, e dalle risposte più o meno adattive introiettate. La personalità è il risultato dell'interazione fra temperamento e carattere, dunque fra componenti costitutivi e acquisiti (Harnic *et al.* 2011). I manuali diagnostici riconoscono i temperamenti nelle loro forme più globali attraverso il costrutto di "personalità" e nelle forme patologiche come "disturbi di personalità" (Akiskal 1998b). Nei disturbi di personalità i modelli interpersonali sono rigidamente stereotipati con una resilienza adattiva relativamente scarsa. Akiskal (1998b) preferisce il termine temperamento alla personalità perché – scrive - "it is more basic, more primary. It is best considered as a bridge between biology and personality. It is, in essence, the biological personality"¹⁶.

L'idea che il temperamento costituisca la condizione predisponente per lo sviluppo di un disturbo dell'umore è una delle idee centrali che ha portato allo sviluppo dello "spettro bipolare". In altri termini, secondo questo modello, il temperamento può costituire una condizione attenuata o sottosoglia della patologia. Questa concezione era presente *in nuce* già in Kraepelin, che sotto questo profilo riprese la teoria degli umori ippocratico-galenica.

¹⁵ Cfr. <https://www.spectrum-project.org/>

¹⁶ "È più basilico, più primario. È meglio considerato come un ponte tra biologia e personalità. È, in sostanza, la personalità biologica" (traduzione nostra).

2.1.1. I TEMPERAMENTI COME PRODROMI: DA KRAEPELIN A TELLENBACH

Il testo “Manic-depressive Insanity and Paranoia” di Kraepelin costituisce la traduzione inglese, pubblicata nel 1921, dell’ultima sezione della nona edizione dell’opera “Psychiatrie. Ein Lehrbuch für Studierende und Ärzte”. In relazione agli “stati fondamentali” scrive lo psichiatra tedesco:

there are certain temperaments which may be regarded as rudiments of manic-depressive insanity. They may throughout the whole of life exist as peculiar forms of psychic personality without further development; but they may also become the point of departure for a morbid process which develops under peculiar conditions and runs its course in isolated attacks (Kraepelin 1921, p. 118)¹⁷.

È importante sottolineare che con il termine “stato fondamentale” – cioè il temperamento - Kraepelin non si riferisce soltanto a quella parte della personalità che presenta sintomi lievi e sottosoglia, ma si riferisce anche agli stessi sintomi morbosi.

Il temperamento depresso, identificato da Kraepelin nel 12,5% dei pazienti del campione maniaco-depressivo, è caratterizzato da un persistente stato emotivo cupo in tutte le esperienze della vita. Molti pazienti con temperamento depresso sono costantemente tormentati dal senso di colpa, non trovano gioia nella vita sociale e nei piaceri e dal punto di vista della condotta, osserva Kraepelin, sono privi di iniziativa, insicuri.

Il temperamento maniacale (o ipertimico) – identificato nel 9% dei pazienti con psicosi maniaco depressiva – costituisce naturalmente l’opposto, sotto il profilo dell’umore, del temperamento depresso. Il temperamento ipertimico nella sua forma più lieve viene descritto in un passaggio che non può non essere riportato per intero:

"The slightest forms of the disorder lead us to certain personal predispositions still in the domain of the normal. It concerns here brilliant, but unevenly gifted personalities with artistic inclinations. They charm us by their intellectual mobility, their versatility, their wealth of ideas, their ready accessibility and their delight in adventure, their artistic capability, their good nature, their cheery, sunny mood. But at the same time, they put us in an uncomfortable state of surprise by a certain restlessness, talkativeness, desultoriness in conversation, excessive need for social life, capricious temper and suggestibility, lack of reliability, steadiness, and perseverance in work, a tendency to building castles in the air and scheming, occasional unusual activities. Now and then one possibly hears also of periods of causeless depression or

¹⁷ “Ci sono certi temperamenti che possono essere considerati come abbozzi di psicosi maniaco-depressiva. Essi possono esistere per tutta la vita come forme peculiari della personalità psichica senza ulteriori sviluppi; ma possono anche diventare il punto di partenza per un processo morboso che si sviluppa in condizioni particolari e segue il suo corso in attacchi isolati” (traduzione nostra).

anxiety, which usually are traced back to external circumstances, over-work, disappointments" (Kraepelin 1921, p. 130)¹⁸.

In relazione al temperamento "irritabile" (12,4% dei pazienti) - riconducibile al temperamento sanguigno ippocratico-galenico - Kraepelin evidenzia come questo costituisca un miscuglio fra le caratteristiche degli stati fondamentali maniacale e depressivo; i pazienti in generale sono allegri, sicuri di sé, disinibiti; ma si inseriscono periodi in cui sono irritabili e di cattivo umore (Kraepelin 1921, p. 131). Il tratto essenziale è però l'irascibilità: i pazienti sono polemici tendono a litigare, non sopportano alcuna contraddizione e quindi cadono facilmente in dispute con le persone che li circondano. Infine, la loro impulsività li espone al rischio di incidenti.

L'ultimo temperamento descritto da Kraepelin è quello ciclotimico, identificato nel 3/4 % dei soggetti con psicosi maniaco-depressiva. Questo è caratterizzato da frequenti fluttuazioni, più o meno regolari, dello stato psichico verso il lato maniacale e depressivo:

To-day lively, sparkling, beaming, full of the joy of life, the pleasure of enterprise, and pressure of activity, after some time they meet us depressed, enervated, ill-humoured, in need of rest, and again a few months later they display the old freshness and elasticity. (Kraepelin 1921, p. 133)¹⁹.

Come abbiamo anticipato, l'idea che il temperamento costituisca un prodromo della patologia, è stata sviluppata anche da Hubertus Tellenbach nel terzo capitolo del suo contributo *Melancholia. Storia del problema, endogenicità, tipologia, patogenesi, clinica*. Tellenbach tratta del Typus Melancholicus (TM), ovvero quel temperamento che predispone allo sviluppo della Melancholia.

I tratti temperamentali prevalenti del TM sono due: la ricerca dell'ordine e la coscienziosità. Questi tratti vengono individuati attraverso uno studio retrospettivo condotto su 119 pazienti ricoverati nel 1959 per Melancholia nella clinica di Heidelberg; le due

¹⁸ "Le forme più lievi del disturbo ci conducono a certe predisposizioni personali che sono ancora nel dominio del normale. Si tratta qui di personalità brillanti, ma diversamente dotate, con inclinazioni artistiche. Ci affascinano con le loro fluttuazioni del pensiero, la loro versatilità, la loro ricchezza di idee, la loro affabilità e il loro gusto per l'avventura, le loro capacità artistiche, la loro buona natura, il loro umore allegro e solare. Allo stesso tempo, però, ci mettono in uno stato di scomoda sorpresa per una certa irrequietezza, loquacità, incoerenza nella conversazione, eccessivo bisogno di vita sociale, temperamento capriccioso e suggestionabilità, mancanza di affidabilità, costanza e perseveranza nel lavoro, per una tendenza a costruire castelli in aria e complottare e per le occasionali attività insolite. Ogni tanto si sente anche parlare di periodi di depressione o ansia senza causa apparente, che di solito vengono attribuiti a circostanze esterne, lavoro eccessivo, delusioni" (traduzione nostra).

¹⁹ "Oggi vivaci, brillanti, raggianti, pieni di gioia di vivere, di intraprendenza e spinta all'attività; dopo un po' di tempo li ritroviamo depressi, infiacchiti, di cattivo umore, bisognosi di riposo, e poi, ancora qualche mese più tardi, mostrano nuovamente la vecchia freschezza ed elasticità" (traduzione nostra).

caratteristiche, precisa più volte Tellenbach, vengono individuati in *tutti* i pazienti dello studio. Come in Kraepelin anche in questo caso il temperamento viene inteso in senso anche (ma non esclusivamente) patologico.

Secondo lo psichiatra fenomenologo il TM è incline alla costante ricerca dell'ordine, tendenza che chiama "ordinatezza". Trattandosi di un tratto temperamentale con "«ordinatezza» – sottolinea Tellenbach – non intendiamo alcun contrassegno di abnormità"; la patologia subentra infatti allorquando la personalità melancolica diventa "ancorata all'ordine" (Tellenbach 1975, p. 106). La ricerca dell'ordine va ben al di là dalla mera 'pulizia' per estendersi soprattutto alle sfere lavorativa e sociale:

"la vita lavorativa è completamente determinata da diligenza e scrupolosità, coscienza del dovere [...] L'ordine contesse anche i rapporti con il prossimo, soprattutto nel pensiero, talvolta addirittura angoscioso, di mantenere l'atmosfera sgombra da disturbi, frizioni, conflitti, in particolare da *colpevolezza* sotto ogni forma. Nelle relazioni con i superiori ed i colleghi stanno in prima linea la fedeltà, l'essere servizievole e la prontezza all'aiuto [...] L'essere per gli altri ha qui il carattere dell'attaccamento devoto, del disinteressato voler prendersi cura, di saldi vincoli con il coniuge ed in particolare con i figli" (Tellenbach 1975, p. 107).

Attraverso la sollecitudine nei confronti degli altri, l'individuo con TM riesce ad anticipare qualunque possibilità di rimanere in debito, e quindi di vivere il senso di colpa. La ricerca dell'ordine nella sfera lavorativa è caratterizzata da un'elevata pretesa rispetto alle proprie prestazioni, tanto sotto il profilo qualitativo, quanto sotto il profilo quantitativo:

Un possibile pericolo è insito nella congiunzione di queste due attitudini: la meticolosa ordinatezza ed una contemporanea grossa pretesa da sé, anche per quel che riguarda la massa di lavoro. Fintanto che le due tendenze rimangono in una certa misura in equilibrio l'una con l'altra tutto è a posto [...] allorquando una tendenza diventa predominante o è ostacola nella sua realizzazione [...] l'esistenza può cadere in un circolo pericoloso (Tellenbach 1975, p. 114).

Il tipo melanconico ricerca la stabilità e la fissità nell'ambiente circostante; desidera uno spazio definito in cui potersi collocare stabilmente, allontanandosi da ogni possibilità di novità o cambiamento. L'eventualità di una violazione dell'ordine o la mancata creazione di esso fa emergere il secondo tratto del *typus melancholicus*: la coscienziosità.

La coscienza è considerata a priori, e talvolta in modo quasi coatto, espressamente come mezzo per evitare ed impedire ogni colpa, anche la più piccola. [...] Tra i melancolici è ampiamente diffuso il sentimento di non fare o di non aver fatto abbastanza, per quanto è loro richiesto, per amore degli altri [...] Si sente dai pazienti come essi abbiano sempre evitato di fare sbagli, di farsi addebitare una colpa, come essi si siano sempre sforzati di sgomberare il campo anche dalle più piccole mancanze (Tellenbach 1975, pp. 128s).

Nel TM il cambiamento crea senso di colpa poiché i soggetti si ritengono responsabili del fallimento del compito di cui si sono incaricati: il mantenimento dell'ordine. Il soggetto si percepisce come l'unico artefice del mantenimento dell'ordine e ogni evento che lo modifica diventa occasione per rimproverarsi di non aver fatto abbastanza per prevenire anche i più piccoli cambiamenti (Siciliano 2015). Il cambiamento (violazione dell'ordine) è, secondo Tellenbach, la circostanza che solitamente dunque porta allo sviluppo o all'esordio dell'episodio depressivo a partire dallo sfondo temperamentale malinconico:

il *typus melancholicus* [...] cade in prossimità del campo di gravità della Melancolia quando vengono posti in pericolo, disturbati o addirittura distrutti quegli ordinamenti in cui si realizzano i suoi rapporti con il prossimo più vicino. Con essi egli ha conferito un significato alla sua esistenza (Tellenbach 1975, p. 128).

2.1.2. SOFT BIPOLAR SPECTRUM: CONTINUUM TRA I DUE POLI MANIACALE E DEPRESSIVO

Negli ultimi anni è notevolmente aumentato l'interesse per lo studio sui temperamenti, un contributo importante in questa direzione è venuto da Akiskal e collaboratori che hanno creato il modello dello "spettro bipolare". Considerando il temperamento come prodromo di malattia, Akiskal e colleghi avevano l'obiettivo di andare ad identificare con migliore precisione le condizioni sub-cliniche per migliorare la prevenzione e la terapia psicofarmacologica.

È nostro interesse ora dare qualche dato sulla incidenza dei Disturbi dello spettro bipolare. Da una revisione sistematica di Dell'Aglio *et al.* (2013) condotta su 11 studi, la prevalenza del disturbo bipolare nell'arco della vita è risultata essere compresa tra lo 0,1 e il 7,5%, mentre la prevalenza dei disturbi dello spettro bipolare variava tra il 2,4 e il 15,1%. Questo dato è interessante perché mostra come, rendendo i criteri meno restrittivi, e ponendo attenzione a certi sintomi subclinici, il tasso di prevalenza del disturbo bipolare possa cambiare radicalmente.

Akiskal e collaboratori (1998a) andando al di là della mera osservazione clinica, attraverso analisi statistiche, diedero per la prima volta sostegno empirico alla classificazione di tutti e quattro "gli stati fondamentali" di Kraepelin trattati nel paragrafo precedente. L'operationalizzazione dei quattro temperamenti individuati dallo psichiatra tedesco viene effettuata attraverso la scala Semi-Structured Affective Temperament Interview (TEMPS-1)

somministrata a 1010 soggetti. Lo studio, che intendeva misurare la consistenza interna della scala, sebbene non ottenne risultati sufficienti per quanto riguarda il temperamento irritabile, supportò l'esistenza dei temperamenti depressivo, maniacale (o ipertimico) e ciclotimico:

It is also noteworthy that, except for the irritable scale, discriminant analysis demonstrated the ability of TEMPS-I to classify subjects. This ability is also confirmed by a negative correlation between hyperthymic and depressive factor scores in the additional exploratory factor analysis (Akiskal *et al.* 1998a).

La TEMPS-I ha ottenuto una coerenza interna da buona a eccellente. Le analisi statistiche condotte hanno individuato i tratti più discriminanti delle quattro scale del temperamento affettivo permettendo di definire i punteggi *cut-off*. Dai risultati delle analisi statistiche, per quanto riguarda la depressione, l'item DEP-7 "Preoccupied with inadequacy, failure and negative events to the point of morbid enjoyment of one's failures"²⁰ assieme all'item DEP-5 "Conscientious or self-disciplining" (coscienzioso e autodisciplinato), è risultato essere il più indicativo di appartenenza al gruppo con temperamento depressivo; quest'ultimo item va a supporto, sostengono gli autori dello studio, dei tratti anancastici del temperamento fatti emergere da Tellenbach (1975). Per quanto riguarda il temperamento ipertimico, tutti i tratti di questo costrutto, evidenzia anzitutto Akiskal, sono altamente intercorrelati; in questo caso gli item HYP-5 "Warm, people-seeking or extroverted" (caloroso, socievole, estroverso) e HYP-1 "Cheerful, overoptimistic or exuberant" (allegro, eccessivamente ottimista, esuberante) sembrano essere i più importanti per predire l'appartenenza al gruppo. In relazione al temperamento ciclotimico, infine, gli item più altamente correlati nell'analisi fattoriale sono risultati essere CYC-1 "Biphasic dysregulation characterized by abrupt endoreactive shifts from one phase to the other, each phase lasting for few days at a time, with infrequent euthymia"²¹ e CYC-2 "Lethargy alternating with eutonia" (letargia alternata a eutimia).

Nel 2005 Akiskal e colleghi diedero nuovo supporto ai quattro temperamenti predisponenti al disturbo dello spettro bipolare (distimico, ipertimico, ciclotimico, irritabile) con la validazione psicometrica del questionario Temperament Evaluation of the Memphis, Pisa, Paris, and San Diego Autoquestionnaire (TEMPS-A). TEMPS-A è un questionario self-report costituito da 110 item a risposta dicotomica vero/falso che va a valutare il

²⁰ "preoccupato per l'inadeguatezza, il fallimento e gli eventi negativi, al punto da trarre un piacere morboso dai propri insuccessi" (traduzione nostra).

²¹ "Disregolazione bifasica caratterizzata da improvvisi cambiamenti di natura endogena da una fase all'altra, ciascuna fase della durata di pochi giorni, con infrequente eutimia" (traduzione nostra).

temperamento in riferimento all'intera vita del soggetto. Il questionario è stato somministrato a 398 pazienti con disturbo dell'umore e ha ottenuto un'elevata affidabilità *test-retest* che varia da 0,58 per l'irritabile a 0,68, 0,69 e 0,70 rispettivamente per il ciclotimico, distimico e ipertimico; la consistenza interna è risultata eccellente, con un valore alpha di Cronbach che va da 0,76 per il distimico a 0,88 per il ciclotimico (Akiskal *et al.* 2005).

Uno studio prospettico condotto da Placidi *et al.* (1998) ha osservato la stabilità dei tratti temperamentali²² in un campione di studenti italiani di 14-18 anni; il campione è stato monitorato per due anni. Dai risultati dello studio emerge che il temperamento ciclotimico è il più stabile fra i temperamenti e costituisce un fattore di rischio per lo sviluppo di episodi depressivi. Come evidenzia inoltre la revisione di Harnic *et al.* (2011), la presenza di un temperamento affettivo di tipo ciclotimico o ipertimico nei pazienti affetti da forme depressive è un fattore di rischio per lo sviluppo del disturbo dello spettro bipolare; questi pazienti, hanno cioè una maggiore probabilità rispetto ai soggetti con gli altri temperamenti (irritabile, distimico) di avere episodi maniacali o ipomaniacali e, di conseguenza, dovrebbero essere considerati un sottotipo "soft" del disturbo bipolare, poiché non presentano una semplice depressione "pura".

Sempre nell'ottica del *soft bipolar spectrum*, la mania potrebbe rappresentare sia un'estensione di tratti ipertimici sia un'inversione di tendenza da un temperamento di base di tipo distimico (Harnic *et al.* 2011, Akiskal *et al.* 1998c). Per quest'ultimo caso, ovvero quando la mania si sviluppa a partire da un temperamento distimico, Akiskal *et al.* (1998c) coniano il termine di "mania disforica". Si tratta di una forma di mania che presenta manifestazioni depressive sottosoglia (almeno due sintomi depressivi). Dai risultati dello studio sembra che la disforia maniacale sia presente in un terzo dei soggetti con mania (il 37%). Inoltre, dai risultati emerge che il temperamento ipertimico sembra conferire una "protezione" contro la formazione di sintomi depressivi durante un episodio maniacale.

²² Il temperamento tende a rimanere stabile nel corso della vita, tuttavia può cambiare durante lo sviluppo.

2.2. TEORIZZAZIONE DELLO SPETTRO DELL'UMORE

Il “mood spectrum” è un costrutto molto simile al “soft bipolar spectrum”, sebbene i due modelli dimensionali, avendo una storia diversa, abbiano prodotto strumenti di valutazione differenti. Le origini dello spettro dell'umore, come abbiamo già accennato, sono da ricondurre al periodo a cavallo tra gli anni '90 del secolo scorso e i primi anni 2000 con lo sviluppo dello "Spectrum Project Collaborative Group" (Cassano *et al.* 2002). Questo progetto ha proposto di considerare la fluttuazione dell'umore come il fondamento costituzionale sia degli episodi maniacali conclamanti sia di quelli depressivi. Attraverso lo spettro dell'umore sono stati sviluppati la SCI-MOODS e la sua versione self-report (MOODS-SR), derivata dalla SCI-MOODS. La SCI-MOODS è una intervista strutturata che indaga la presenza/assenza dei sintomi durante il corso della vita. Del MOODS-SR sono presenti tre versioni, in base al periodo di vita valutato: corso della vita; ultimo mese, ultima settimana²³. Tutte le versioni sono state sviluppate in inglese e in italiano e presentano 161 item. Il MOODS-SR è composto da quattro domini: dominio Neurovegetativo, dominio dei livelli di energia, dominio dell'umore e dominio cognitivo. Il dominio neurovegetativo valuta i disturbi e i cambiamenti ritmici degli stati d'animo, degli atteggiamenti alimentari, dell'attività sessuale e del sonno, comprese le variazioni ritmiche dei sintomi affettivi; il dominio dei livelli di energia valuta i cambiamenti nelle attività quotidiane, con particolare attenzione al lavoro, agli hobby e alla vita sociale; il dominio dell'umore esplora l'intero spettro dei sintomi depressivi e maniacali, dei segni e delle manifestazioni sottosoglia della fluttuazione dell'umore; il dominio cognitivo infine valuta i cambiamenti nella sfera cognitiva che si verificano insieme alla fluttuazione dell'umore (Benvenuti *et al.* 2015).

La revisione sistematica di Benvenuti e colleghi (2015) ha considerato tutti gli studi pubblicati tra il 1997 e l'agosto 2014 che hanno usato questi due strumenti, per valutarne i risultati. La ricerca ha prodotto 43 articoli, incluse cinque revisioni della letteratura o articoli metodologici e un caso clinico. Gli studi hanno coinvolto campioni di quattro paesi diversi (Italia, Stati Uniti, Spagna e Giappone). I primi due studi sulla SCI-MOODS hanno testato la sua validità e affidabilità. L'affidabilità dell'intervista ha dimostrato essere eccellente: l'affidabilità *inter-rater* dei domini variava tra lo 0,93 e lo 0,94 e la consistenza interna dei domini variava tra lo 0,79 e lo 0,92 (Fagiolini *et al.* 1999, Rucci & Maser 2000).

²³ I questionari sono consultabili dal sito <https://www.spectrum-project.org/questionnaires/moods.html>

Un altro studio importante è quello di Cassano *et al.* (2004). Uno degli obiettivi della costruzione del MOODS-SR era quello di creare uno strumento in grado di riconoscere caratteristiche ipomaniacali lievi e atipiche tanto nei pazienti bipolari quanto in quelli unipolari. Come ipotizzato da Cassano e collaboratori (2004), lo strumento ha identificato lievi caratteristiche ipomaniacali in 117 pazienti con diagnosi di depressione unipolare ricorrente in remissione, basandosi sui criteri diagnostici del DSM-IV. Una presenza significativa della componente ipomaniacale è stata riscontrata tra i pazienti che presentavano un numero maggiore di sintomi depressivi nel corso della vita. Questa correlazione fu un risultato importante dato che agli individui era stata diagnosticata una depressione unipolare pura. Nei risultati della metanalisi discussi ed elencati da Benvenuti *et al.* (2015), quest'area è dunque la prima in cui lo spettro dell'umore ha dimostrato la sua utilità: la possibile manifestazione di ipomania atipica o sottosoglia rilevabile attraverso il MOODS-SR (1). Le altre due aree riguardano: 2. le caratteristiche dello spettro non rilevate da altri strumenti ma clinicamente rilevanti; queste possono manifestarsi a ondate durante la vita, a volte insieme, a volte da sole, a volte raggiungendo la gravità di un disturbo conclamato, a volte complicando il corso delle malattie somatiche; 3. punteggi più alti sui fattori della MOODS-SR rispetto ai "disturbi psicomotori", all'"instabilità mista" e al "rischio suicidario" delineano sottotipi di pazienti caratterizzati dalle forme più gravi dei disturbi dell'umore, dal rischio maggiore di sintomi psicotici e dalla qualità della vita inferiore dopo la remissione dell'episodio conclamato.

2.3. CRITICITÀ E PROSPETTIVE FUTURE

L'approccio dimensionale ai disturbi dell'umore è certamente utile per la ricerca e per la clinica se usato in modo complementare a quello categoriale, proprio degli attuali sistemi diagnostici. I modelli spettrali sono infatti in grado di considerare quelle caratteristiche sintomatologiche che non soddisfano i criteri diagnostici del DSM-5-TR e dell'ICD-11. Questi modelli in questo modo – ipotizzando un continuum tra polo idiografico e nomotetico - avvicinano al polo idiografico la diagnosi. L'approccio dimensionale rispetto a quello categoriale - e ciò vale non solo per quello spettrale ai disturbi dell'umore, ma per tutti i disturbi - si fa meglio carico delle sfumature dei disturbi mentali permettendo di coglierne

la variabilità individuale. L'approccio spettrale, come complemento di quello categoriale, sembra in particolare necessario nei disturbi dell'umore, vista la grande variabilità – riconosciuta fin dagli albori della nosografia moderna - delle caratteristiche dello spettro bipolare-unipolare. Tale approccio, permettendo l'identificazione dei fenotipi all'interno dei disturbi affettivi, ci sembra pertanto avere idealmente due utilità fondamentali: consente di migliorare la prevenzione, sulla base della conoscenza dei sintomi sottosoglia e delle caratteristiche temperamentali (1); consente di migliorare il trattamento attraverso la sua individualizzazione (2). Sotto questo profilo l'approccio spettrale ai disturbi dell'umore ha manifestato la sua efficacia nella capacità di rilevare le caratteristiche ipomaniacali sottosoglia presenti nella depressione unipolare che, se non rilevate, potrebbero limitare o persino rendere iatrogeno lo *switch* verso il bipolarismo a causa del trattamento psicofarmacologico²⁴.

A fronte di questi vantaggi, devono essere fatti rilevare i punti critici dei modelli spettrali dei disturbi dell'umore. Anzitutto, la ricerca sull'efficacia del trattamento che si avvale di questi modelli è, da quanto conosciamo, sostanzialmente limitata (1); questa è infatti basata perlopiù su un approccio categoriale. Ciò è dovuto anzitutto alla difficoltà di trovare un campione consistente con caratteristiche sfumate. Questo primo punto a sfavore dell'approccio dimensionale si ritorce contro il secondo punto a favore dell'approccio dimensionale: se manca la ricerca sull'efficacia, come è possibile migliorare il trattamento? Inoltre, non esiste ancora un largo consenso su come definire le dimensioni dei disturbi dello spettro dell'umore/bipolare e sui valori cut-off (2). In terzo luogo, le categorie proposte, essendo più sfumate e complesse, sono meno intuitive rispetto alle entità nosografiche dei manuali diagnostici, questo può sfavorire l'applicazione clinica di questi modelli (3). Infine, è possibile che un'attenzione eccessiva alle condizioni pre-morbose possa portare al rischio di sovradiagnosi (4).

²⁴ Mi riferisco all'uso degli antidepressivi in soggetti che hanno un episodio depressivo all'interno di un disturbo bipolare che - seguendo l'approccio spettrale - potrebbe non essere riconosciuto anche per la mancata soddisfazione di tutti i criteri diagnostici necessari. A riguardo scrive Miklowitz: “[gli antidepressivi] posseggono un effetto collaterale maggiore nel disturbo bipolare [rispetto alla depressione unipolare]: possono elicitare uno stato ipomaniacale, maniaco o misto e possono causare una rapida cicizzazione del disturbo” (Miklowitz 2005, p. 123). Una metanalisi di studi clinici randomizzati controllati e studi di coorte ha mostrato che i tassi di "switching" dell'umore, ovvero il passaggio rapido da uno stato depressivo a uno maniaco o ipomaniacale, in un follow-up medio di cinque mesi sono approssimativamente del 15,3% nelle persone con disturbo bipolare trattate con antidepressivi rispetto al 13,8% in quelle senza trattamento antidepressivo. Le linee guida attuali sconsigliano l'uso di antidepressivi in pazienti bipolari (Goes 2023).

Appare dunque evidente la necessità, in primo luogo, di migliorare la ricerca sull'efficacia del trattamento dei disturbi dell'umore basata su un approccio dimensionale perché solo in questo modo i modelli spettrali possono rivelare la loro (potenzialmente grande) utilità clinica. Per compiere questo passo è però necessario un maggiore consenso sulle categorie proposte da questi modelli. In questo senso è richiesta anche una maggiore ricerca sulle proprietà psicometriche di validità e affidabilità delle scale proposte. Attualmente, il contributo più significativo dei modelli spettrali risiede nella loro capacità di supportare la diagnosi differenziale quando integrati con l'approccio categoriale dei manuali diagnostici prevalenti.

3. DIAGNOSI DEI DISTURBI DELL'UMORE: SFIDE E LIMITI

Nel prossimo sottoparagrafo ci concentreremo su alcuni degli strumenti di screening più usati per la diagnosi dei Disturbi depressivi e dei Disturbi bipolari. Nei sottoparagrafi successivi andremo a considerare alcuni *field trials* eseguiti sulle recenti revisioni dei manuali diagnostici per vedere quali siano stati i risultati in termini di affidabilità e accuratezza delle linee guida per le categorie diagnostiche oggetto della tesi. Esporremo infine alcuni dei *bias* del clinico e del paziente che, entrando in gioco nel processo diagnostico, rischiano di inficiarne l'esito.

3.1. METODOLOGIE DIAGNOSTICHE

Nella metanalisi da noi condotta sull'accordo tra valutatori nei disturbi affettivi, che andremo a presentare nel capitolo 1, i due strumenti più utilizzati per la diagnosi dei disturbi dell'umore sono stati: SCID (in 10 studi) e CIDI (in 5 studi). È da notare che non tutti gli studi che hanno usato questi strumenti hanno svolto diagnosi solo sui disturbi dell'umore. Dunque, non escludiamo naturalmente che queste interviste semi-strutturate e strutturate, rispettivamente SCID e CIDI, non siano utilizzate per le diagnosi dei disturbi affettivi - visto che costituiscono due delle interviste strutturate e semi-strutturate più diffuse per i disturbi

mentali *in generale* - ma nemmeno possiamo indicarle come gli strumenti più utilizzati per la diagnosi dei disturbi dell'umore.

Lo strumento di screening più utilizzato e validato per la depressione è il questionario self report "Patient Health Questionnaire-9" (PHQ-9), come emerge da una recente metanalisi sugli strumenti di screening per la diagnosi dei disturbi depressivi (El-den *et al.* 2018). Il PHQ-9 si compone di 9 domande. Ogni domanda è valutata su una scala da 0 a 3, a seconda della frequenza del sintomo. Il punteggio totale fornisce una categorizzazione che va da "nessuna depressione" a "depressione grave". Il PHQ-9 presenta un'elevata sensibilità (88%) e specificità (88%) per gli episodi di depressione maggiore; tuttavia, questo strumento non è adatto a distinguere il disturbo bipolare dal disturbo depressivo maggiore (Hirschfeld 2014). Oltre al PHQ-9 vi sono molti altri strumenti di screening comunemente utilizzati nella pratica clinica della depressione; tra questi citiamo il Beck Depression Inventory-II (BDI-II) a 21 item e il Quick Inventory of Depressive Symptomatology (QIDS).

È importante mettere in luce l'infelice dato che emerge dalla metanalisi di Dagani *et al.* (2017). Questa ha identificato un intervallo di quasi sei anni tra la comparsa dei sintomi o degli episodi di disturbo bipolare e la presa in carico del paziente dal punto di vista trattamentale. Questo enorme ritardo dipende anche, come notano i ricercatori, dalla frequente errata diagnosi di disturbo bipolare. Altri studi meno recenti, hanno indicato un ritardo medio della corretta diagnosi di 7-10 anni (Suppes *et al.* 2001, Mantere *et al.* 2004). Il riconoscimento del disturbo bipolare è difficile e molto frequente è l'errata diagnosi (Stensland *et al.* 2010, Bruchmüller & Meyer 2009). A causa della presenza degli episodi depressivi, esso viene, di fatti, troppo spesso confuso con la depressione unipolare, venendo diagnosticato come tale solo con il sopraggiungere di un episodio maniacale. Queste evidenze indicano l'urgenza di migliorare il processo diagnostico del disturbo bipolare nella pratica clinica, per permetterne una diagnosi precoce che possa rendere la prognosi dei soggetti più favorevole. Un passo essenziale in questa direzione è l'utilizzo nella pratica clinica di strumenti di screening ampiamente validati e con alti valori di sensibilità e specificità.

Gli strumenti di screening più studiati e riportati nelle revisioni (McIntyre *et al.* 2020, Hirschfeld 2014, Rucci *et al.* 2013) per i disturbi bipolari sono il Mood Disorders Questionnaire (MDQ) e l'Hypomania Checklist-32 (HCL-32), entrambi sono questionari self-report. Per l'MDQ, la sensibilità riportata da una recente metanalisi (Wang *et al.* 2019) è dell'80% e la specificità del 70%, ma i dati variano a seconda degli studi, ad esempio in

un'altra metanalisi Zimmermann e Galione (2011) riportano una sensibilità piuttosto bassa, del 61,3%.

Per l'HCL-32, la sensibilità riportata è dell'80% e la specificità del 51% (Angst *et al.* 2005). L'HCL-32 è suddiviso in tre sezioni. La prima sezione indaga lo stato emotivo attuale del soggetto e il suo livello abituale di attività, energia e umore. La seconda riguarda specificatamente i sintomi di mania e ipomania; la terza valuta alcuni comportamenti associati, come, ad esempio, se, durante uno stato maniacale, il soggetto tenda a litigare più spesso, consumare più caffè o alcol, o fumare di più (Hirschfeld 2014).

Dato che questi strumenti sono tutti test di screening, non possono essere utilizzati da soli per fare diagnosi. Essi sono pensati per essere utilizzati durante la visita medica dal medico di famiglia o nel primo colloquio di assessment dallo psichiatra o dallo psicologo e, in caso di risultato positivo, indicano la necessità di una valutazione clinica attenta e approfondita. È da precisare inoltre che uno strumento di screening per la depressione come il PHQ-9 non affronta la sfida diagnostica relativa alla diagnosi differenziale tra disturbo unipolare e bipolare (Hirschfeld 2014); il nocciolo di questa sfida è, quindi, costituito dall'eventuale individuazione dell'episodio ipomaniacale. Sotto questo profilo, il MOODS-SR, basato sul modello dello Spettro dell'umore, sembra promettente. Ciò è dovuto al fatto che - come già è stato precedentemente evidenziato - esso è stato sviluppato per rilevare le forme più leggere dei disturbi bipolari nonché i sintomi isolati dello spettro maniacale-ipomaniacale (Rucci *et al.* 2013). Per migliorare la diagnosi differenziale tra disturbo depressivo maggiore e disturbo bipolare II, potrebbe essere somministrato insieme a un test di screening sulla depressione, anche uno strumento di screening per il disturbo bipolare, come il già citato MDQ. Questo non sembrerebbe ancora sufficiente per effettuare una corretta diagnosi. La già citata metanalisi di Zimmermann e Galione (2011) condotta su 20 studi sulla sensibilità e specificità dell'MDQ ha rilevato, infatti, che la sensibilità del questionario era relativamente modesta, pari a 61,3%. Di conseguenza, se i clinici seguissero lo screening valutando più approfonditamente solo i pazienti con punteggi positivi, potrebbero non diagnosticare il disturbo bipolare in circa un terzo dei casi. Pertanto, gli strumenti di screening possono costituire solo la prima fase del processo diagnostico, a cui deve seguire necessariamente - se il soggetto risulta positivo al test - una intervista strutturata o semi-strutturata come la SCID o la CIDI. È opportuno, inoltre, durante la raccolta anamnestica, valutare la storia familiare di malattia (Rucci *et al.* 2013).

3.2. AFFIDABILITÀ E VALIDITÀ DELLA DIAGNOSI

Come scrive Price (2017, pp. 59ss), nell'ambito delle misurazioni psicologiche una definizione di validità internazionalmente riconosciuta è “a judgment or estimate of how well a test or instrument measures what it is supposed to measure”²⁵. La validità in psicometria è la capacità di un indicatore di rappresentare empiricamente il costrutto che intende misurare. La validità è costituita da più componenti:

- la validità di contenuto: indaga se il campione di item che abbiamo inserito nel test è un campione rilevante e rappresentativo della totalità delle operazionalizzazioni possibili del costrutto. Un modo per valutare la validità del contenuto è chiedere a un gruppo di professionisti di esaminare la definizione del costrutto e il modo in cui è stato operazionalizzato attraverso gli item.
- la validità di criterio: quella caratteristica di un test psicologico che consente di valutare quanto i suoi risultati possono essere utilizzati per prevedere un comportamento specifico di una persona.
- la validità di costrutto: consiste in una valutazione dell'appropriatezza delle deduzioni basate sui punteggi di un test che misura un costrutto specifico; in altre parole, è il giudizio relativo alla questione se il test misura o meno il costrutto che intende misurare (Chiorri 2020, pp. 42ss).

La forma più importante di validità è la “validità di costrutto”, in quanto costituisce la base su cui si fondano tutte le altre forme di validità (Price 2017, pp. 59ss). Le altre due forme di validità possono essere considerate, infatti, delle componenti specifiche all'interno del modello unificato di validità di costrutto. Sotto questo profilo, stabilire prove per la validità di criterio e di contenuto contribuisce in modo più generale a sostenere la validità di costrutto di un determinato strumento o indicatore. La validità dei punteggi dei test non è semplicemente espressa da un'unica misura di sintesi statistica (ad esempio, la correlazione tra un punteggio del test e un criterio esterno), ma piuttosto da prove multiformi acquisite da questioni relative a criteri, contenuti e costrutti (Price 2017, *ibid*). È importante notare, infine, che la validità dei punteggi dei test può essere vista come un continuum che va da debole ad accettabile a forte.

²⁵ “Un giudizio o una stima di quanto bene un test o uno strumento misura ciò che è inteso a misurare” (traduzione nostra).

L'affidabilità, invece, si riferisce al grado in cui i punteggi di test o altri strumenti sono privi di errori di misurazione (Price 2017, pp. 206ss). In altre parole, la stima dell'affidabilità (*reliability*) è direttamente collegata alla precisione o imprecisione della misurazione, quindi all'errore di misurazione (Chiorri 2020, pp. 38ss). Possiamo distinguere due errori di misurazione di un test o di una scala: errori casuali, cioè le “fluttuazioni casuali del punteggio osservato attorno al punteggio vero” (Chiorri 2020, *ibid*), e gli errori sistematici, ovvero errori costanti, i *bias*. Di questi ultimi, in relazione al processo diagnostico, tratteremo nel sottoparagrafo 3.2.2.

Due modalità per quantificare l'affidabilità di uno strumento diagnostico sono l'affidabilità *test-retest* e l'affidabilità *inter-rater*. L'affidabilità *test-retest* valuta la stabilità dei punteggi in somministrazioni successive dello strumento; se, come postulato, la sintomatologia esaminata non cambia tra le due valutazioni, il punteggio della scala dovrebbe essere sostanzialmente identico; ci dovrebbe essere, quindi, una elevata correlazione tra le due valutazioni. L'affidabilità *inter-rater* indica il grado di accordo fra valutatori indipendenti in relazione alle caratteristiche di uno stesso soggetto. Se l'affidabilità *inter-rater* è alta, gli esperti possono essere fiduciosi del fatto che individui con una formazione simile sono portati a produrre punteggi simili su soggetti dello stesso tipo. Se l'affidabilità è bassa, è meno probabile che i punteggi ottenuti possano essere riprodotti in un gruppo diverso di valutatori.

Dato che la diagnosi costituisce la base della scelta del trattamento, la bassa affidabilità diagnostica minaccia la validità sia della ricerca quanto della pratica clinica. Vista la scarsa presenza di biomarcatori specifici identificabili, diagnosticare disturbi mentali è spesso difficile; ciò costringe infatti i clinici a fare affidamento sulla valutazione delle caratteristiche del paziente, che però sono suscettibili di interpretazione e di distorsioni cognitive (Rocha Neto *et al.* 2023a).

In linea generale, l'intervista diagnostica, per essere valida e affidabile, deve permettere una interpretazione corretta della nomenclatura del manuale diagnostico a cui fa riferimento; inoltre, le domande dell'intervista devono essere facilmente comprensibili, devono essere presentate in modo standard e le risposte ottenute devono essere registrate correttamente; le risposte devono essere, quindi, valutate secondo i criteri diagnostici del sistema diagnostico di riferimento (Robins & Cottler 2004). Affinché l'intervista diagnostica sia valida e affidabile, è molto importante che la valutazione diagnostica si basi almeno in parte su scale e test che presentano proprietà psicometriche di validità e affidabilità adeguate.

Chiaramente anche le entità nosografiche dei manuali diagnostici come l'ICD e il DSM hanno una certa validità e affidabilità. Il paragrafo successivo sarà dedicato proprio a questo tema.

3.2.1. I *FIELD TRIALS* SULLE SINDROMI AFFETTIVE

Com'è noto, i manuali diagnostici devono essere periodicamente revisionati per garantire che le loro definizioni diagnostiche riflettano le conoscenze scientifiche attuali sui disturbi mentali. I *field trials* ("studi sul campo") consentono di migliorare la base empirica delle classificazioni in ambito psicopatologico, contribuendo dunque a migliorare l'assistenza clinica dei pazienti. Essi si dividono in *developmental field trials* ovvero "di sviluppo" e *summative field trials*, ovvero "sommativi" o "conclusivi". I *developmental field trials* sono una componente importante nel processo di revisione diagnostica basato sull'evidenza, poiché permettono di raccogliere i dati clinici in merito alla affidabilità, validità e utilità clinica delle entità nosografiche proposte nello sviluppo di una nuova versione di un manuale diagnostico; questi studi sul campo aiutano a perfezionare – ad esempio determinando le soglie diagnostiche ottimali – le bozze proposte prima che vengano finalizzate (First 2016). Gli studi sul campo conclusivi hanno invece lo scopo di valutare la qualità complessiva del sistema diagnostico già pubblicato. Essi forniscono stime sull'affidabilità diagnostica per le diverse categorie nosografiche, fornendo in questo modo informazioni per future revisioni dei manuali diagnostici (First 2016).

I *developmental field trials* sono stati inaugurati durante il processo di sviluppo del DSM-III-R, ma sono stati eseguiti anche per il DSM-IV, l'ICD-10 e l'ICD-11. Per quanto riguarda gli studi sul campo di sviluppo del DSM-5 invece, sebbene inizialmente presentati come *developmental*, si sono rivelati - nota First - essenzialmente come *summative*. Ciò significa che i risultati di affidabilità non hanno avuto un ruolo importante nello sviluppo dei criteri diagnostici finali del DSM-5 (First 2016).

Presentiamo ora il *field trial* di Regier e collaboratori (2013), parte dello studio di Clarke *et al.* 2013, condotto sul DSM-5, di cui si voleva valutare l'affidabilità *test-retest*. Questo studio come sottolinea First (2016) è un *field trial summative*. Lo studio è stato condotto per ottenere una stima di kappa come misura del grado di accordo tra due clinici che valutano indipendentemente la presenza o l'assenza di diagnosi selezionate del DSM-5

per lo stesso paziente in due occasioni diverse. I clinici che hanno partecipato alle ricerche come valutatori sono stati 279, mentre i pazienti valutati sono stati 2246 (Clarke *et al.* 2013). Sono stati selezionati undici centri accademici negli Stati Uniti e in Canada, a ciascuno dei quali sono state assegnate diverse diagnosi specifiche usualmente trattate in quella sede. I pazienti, una volta recati nel centro, sono stati sottoposti a test di screening e per poi essere divisi in base alla diagnosi del DSM-IV. I pazienti sono stati quindi assegnati casualmente a due clinici per un colloquio diagnostico; i clinici non erano a conoscenza di nessuna diagnosi precedente.

Lo studio ha riportato il kappa per 15 categorie diagnostiche per soggetti adulti e per 8 categorie nosografiche dell'età infantile/adolescenziale. La maggior parte delle diagnosi testate ha mostrato un'affidabilità da buona a molto buona. Per quanto riguarda i disturbi affettivi e in particolare per il disturbo bipolare I negli adulti, l'affidabilità *test-retest* è risultata essere buona (kappa 0,40-0,59). Il disturbo Bipolare II non ha avuto un campione sufficiente per consentire una stima esatta di kappa, mentre l'affidabilità per il disturbo depressivo maggiore (MDD), non è risultata affatto buona, ma “discutibile” (kappa 0,20-0,39) tanto negli adulti quanto nei bambini/adolescenti; Nei *field-trial* condotti sul DSM-IV – i cui criteri per MDD rimangono uguali - l'affidabilità di MDD risultava migliore; Liu (2016) ipotizza che la discrepanza tra i risultati dei precedenti *field trials* sul DSM-IV e quello di Regier e colleghi (2013) sia dovuta all'introduzione di nuovi specificatori per il MDD, con i quali i clinici erano poco famigliari. Due specificatori introdotti nel DSM-5 sono infatti “con caratteristiche miste” e “con ansia”. Questi specificatori potrebbero aver reso più sfumati i confini tra il MDD e altri disturbi, come il disturbi d'ansia generalizzato e il disturbo bipolare, facendo diminuire l'affidabilità *test-retest*; questa spiegazione non viene però ipotizzata dagli autori dello studio, che indicano tuttavia la necessità di migliorare l'affidabilità per MDD.

Lo studio di Kogan *et al.* (2021) è un *developmental field trial* sull'ICD-11 basato sulla valutazione di vignette cliniche standardizzate via internet; lo studio intendeva valutare le linee guida diagnostiche (CDDG) per i disturbi dell'umore dell'ICD-11. Ad essere valutata è stata l'accuratezza dei giudizi diagnostici dei 1892 clinici partecipanti oltre all'utilità clinica percepita. I ricercatori hanno confrontato l'accuratezza dei clinici nell'utilizzare le CDDG dell'ICD-11 con quella dei clinici che hanno utilizzato le CDDG dell'ICD-10 nel diagnosticare disturbi classificati all'interno del gruppo dei disturbi affettivi e nel differenziare i disturbi dell'umore da altri disturbi con sintomatologia condivisa. I

partecipanti, una volta entrati nella piattaforma, venivano inseriti casualmente in uno dei gruppi ed erano chiamati a dare un giudizio diagnostico basandosi su un elenco che comprendeva tutte le diagnosi di disturbo dell'umore e le diagnosi con sintomatologia sovrapposta. L'accuratezza delle CDDG dell'ICD-11 per i disturbi dell'umore è risultata simile a quella delle CDDG dell'ICD-10. Notevoli eccezioni riguardano le difficoltà nel differenziare le nuove categorie ICD-11 di disturbo bipolare tipo I e tipo II e una minore accuratezza rispetto all'ICD-10 nell'applicazione degli specificatori di gravità depressiva dell'ICD-11, i quali sono stati revisionati in seguito ai risultati dello studio. Le linee guida dell'ICD-11 hanno dimostrato una più accurata rilevazione degli episodi depressivi nel contesto del disturbo depressivo ricorrente. L'utilità clinica per il disturbo bipolare di tipo I e di tipo II è risultata inferiore rispetto al disturbo bipolare affettivo dell'ICD-10; ciò - concludono i ricercatori - potrebbe dipendere della scarsa familiarità dei clinici con l'utilizzo delle CDDG ICD-11 nella pratica clinica (Kogan *et al.* 2021).

3.2.3. BIAS COGNITIVI DEL CLINICO E DEL PAZIENTE

Lo studio dei *bias*, che affonda le sue radici nei lavori di Amos Tversky e Daniel Kahneman a partire dagli anni '70, si concentra su quegli errori del pensiero sistematici – appunto i *bias* – che derivano dall'utilizzo di scorciatoie cognitive, le cosiddette "euristiche". Le euristiche, tipiche del pensiero clinico maturo, riducono il tempo, le risorse e lo sforzo cognitivo necessari per il processo decisionale clinico permettendo di prendere decisioni in condizioni di incertezza (Crumlish & Kelly 2009). L'uso di euristiche si contrappone al metodo decisionale ipotetico-deduttivo, processo strutturato che si basa su osservazioni e ragionamenti logici per formulare e verificare le ipotesi sul possibile corso delle azioni. Quest'ultimo metodo implica l'assenza di *bias*, la disponibilità di tempo e risorse ottimali. Il problema delle euristiche, rispetto al metodo ipotetico-deduttivo, sta nel fatto che appunto – nonostante di solito risultino efficaci - possono determinare dei *bias*; questi, essendo errori sistematici, sono però prevedibili. È quindi importante che il clinico che utilizza le euristiche sia consapevole dei loro limiti e sia in grado di riconoscerle.

Gli errori diagnostici in psicopatologia non sono certamente dovuti solo all'uso di euristiche da parte del clinico. Molteplici sono infatti le fonti di errore, ne elenchiamo solo

alcune: la soggettività del paziente, i criteri del manuale diagnostico²⁶, il grado di affidabilità e di validità delle categorie diagnostiche, il grado di affidabilità e di validità delle scale utilizzate.

Crumlish e Kelly (2009) riportano i seguenti *bias* diagnostici correlati alle euristiche che più frequentemente entrano in gioco nel processo diagnostico:

- **Rappresentatività:** L'errore di rappresentatività si verifica quando varianti atipiche di un disturbo vengono ignorate perché il clinico si basa su una presentazione prototipica. Ad esempio, la schizofrenia, che ha come sintomi prototipici deliri e allucinazioni, può essere diagnosticata anche in assenza dei sintomi positivi; infatti, un paziente con soli sintomi negativi e disorganizzati può essere diagnosticato con la schizofrenia. Saremmo di fronte a un errore cognitivo se si diagnosticasse la schizofrenia per la semplice presenza di un sintomo prototipico.
- **Disponibilità:** L'euristica della disponibilità è determinata dalla facilità con cui al clinico viene in mente un esempio in relazione a una scelta da intraprendere²⁷. L'errore si verifica, ad esempio, quando si sceglie un intervento perché è stato recentemente discusso in un congresso.
- **Ancoraggio:** L'euristica dell'ancoraggio è la tendenza a concentrarsi molto presto su caratteristiche ritenute importanti di una presentazione sintomatologica, quindi a formulare un'ipotesi precoce e a non modificarla alla luce di informazioni successive. Da questa tendenza deriva l'errore d'ancoraggio. L'esempio che riportano i due clinici è lampante: le lettere di invio di un collega spesso indirizzano inevitabilmente la diagnosi e possono quindi portare ad attivare tanto l'euristica dell'ancoraggio quanto quella della disponibilità determinando spesso i relativi *bias*.
- **Bias di conferma:** è la tendenza a cercare solo informazioni che supportino, piuttosto che smentiscano, un'ipotesi iniziale, o a interpretare selettivamente le informazioni acquisite dopo la formazione dell'ipotesi in modo che queste la confermino.

²⁶ Ad esempio, ci sono criteri che rimando direttamente al giudizio del clinico. A riguardo Crumlish & Kelly scrivono: "The decision to diagnose a DSM-IV-TR mental disorder, when symptoms are clearly elicited, depends on the psychiatrist's judgement as to what constitutes 'clinically significant distress or impairment in social, occupational, or other areas of functioning' [...] Clearly, the degree of impairment judged to be 'clinically significant' may vary from psychiatrist to psychiatrist. The psychosocial dimension of diagnosis distinguishes psychiatry from other disciplines and adds to diagnostic complexity".

²⁷ la probabilità di un evento viene sovrastimata perché quell'evento viene facilmente in mente, oppure al contrario viene sottostimata perché non è immediatamente accessibile ("disponibile") alla mente.

Alcune delle spiegazioni avanzate dalla letteratura per chiarire il fenomeno della errata diagnosi del disturbo bipolare sono le seguenti:

- I pazienti bipolari sperimentano ripetuti episodi depressivi prima di avere il loro primo episodio (ipo)maniacale; questo può portare a una diagnosi iniziale di depressione unipolare, trascurando la presenza del disturbo bipolare (Hirschfeld & Vornic 2004).
- Gli stati d'animo (ipo)maniacali, caratterizzati da euforia, iperattività e impulsività, spesso non spingono i pazienti a cercare un trattamento (Hirschfeld 2001).; inoltre, anche se già in terapia, i pazienti tendono a non riportare spontaneamente questi sintomi (Hirschfeld *et al.* 2003). Ciò può rendere difficile per i medici identificare il disturbo bipolare.

Tra i fattori che giocano un ruolo importante negli errori diagnostici per il disturbo bipolare ci sono i *bias* del clinico. Sembra che l'euristica della rappresentatività (ovvero del prototipo), si attivi spesso nel processo diagnostico del *Bipolar Disorder* (BD), producendo errori diagnostici. La conclusione dello studio con l'uso di vignette cliniche di Bruchmüller e Meyer (2009), è che il criterio del "ridotto bisogno di sonno" sia un criterio prototipico nella diagnosi di BD e questo produce un mancato riconoscimento dei casi che non presentano questa caratteristica. A 400 psicoterapeuti è stata inviata per mail una di queste 5 vignette: una presentava i criteri per la depressione unipolare; 4 presentavano tutti i criteri per la diagnosi di disturbo bipolare in accordo con il DSM-IV e l'ICD-10. In 2 di queste ultime 4 vignette era presente la caratteristica di "ridotto bisogno di sonno", mentre nelle altre due il sonno era normale. I partecipanti che hanno risposto allo studio (185) erano chiamati a esprimere il proprio giudizio in relazione alla diagnosi. Dai risultati dello studio è emerso che quando nelle vignette bipolari era presente il criterio del ridotto bisogno di sonno, ciò aumentava la probabilità di diagnosi di BD. Per le vignette bipolari nel complesso, il disturbo Bipolare (BD) è stato diagnosticato solo nel 38% dei casi. Nel 62% è stata posta un'altra diagnosi, solitamente di depressione unipolare (53%). Al contrario, per la vignetta che presentava i criteri di depressione unipolare, quasi tutti i partecipanti hanno risposto correttamente. Ciò conferma che il problema della diagnosi errata è specifico per il disturbo bipolare. Il fattore 'sonno' ha avuto una significativa influenza sulla diagnosi. La probabilità di una diagnosi bipolare era infatti molto più alta quando erano fornite informazioni aggiuntive sul sonno ridotto rispetto al sonno normale. Le due vignette bipolari che menzionavano un sonno ridotto avevano rispettivamente il 73% e il 38% di diagnosi di BD. Al contrario, le altre due vignette bipolari con sonno normale avevano solo il 18% e il 23%

di diagnosi di BD. Lo studio ha quindi mostrato come il *bias* della rappresentatività entri significativamente in gioco nella diagnosi di disturbo bipolare.

Per quanto riguarda i fattori relativi al paziente con disturbo bipolare che possono influenzare il processo diagnostico, questi ultimi, come abbiamo già accennato, tendono a non riportare spontaneamente i sintomi della (ipo)mania quali l'euforia, l'iperattività e l'impulsività e ciò può rendere difficile ai medici identificare il disturbo bipolare (Hirschfeld *et al.* 2003). Anche i pazienti con depressione possono presentare dei *bias* durante la raccolta anamnestica. Le diagnosi di depressione si basano spesso sui ricordi di una persona riguardo alle sue recenti esperienze emotive e ciò implica che il paziente riferisca come si sia sentito nell'ultima settimana. I soggetti con depressione però sembrano avere un *bias* di memoria che sfavorisce una rappresentazione accurata delle loro esperienze emotive quotidiane (Urban *et al.* 2018). Lo studio di Urban e colleghi è stato condotto su 1.657. I partecipanti hanno compilato un questionario che andava a valutare l'eventuale presenza di sintomi depressivi nei 12 mesi precedenti. Circa 6 mesi dopo, i partecipanti hanno riportato le loro esperienze emotive della giornata ogni giorno per 8 giorni e hanno poi ricordato queste esperienze nell'ultimo giorno. Le differenze tra le emozioni riportate ogni giorno durante la settimana e quelle ricordate alla fine della settimana sono state confrontate tra i partecipanti con e senza storia di depressione. Entrambi i gruppi hanno sovrastimato la frequenza con cui sono state sperimentate le emozioni negative durante la settimana, ma questo *bias* era significativamente maggiore per i partecipanti con una storia di depressione rispetto al gruppo di controllo. Nella conclusione dello studio gli autori dell'articolo scrivono:

Research presented here shows that sadness, anxiety, and anger are elevated to the same extent among people with a history of depression and that recollections of these emotions, particularly anger, are likely to be overestimated. Therefore people with a history of depression are not only more likely to experience more sadness, anxiety, and anger during daily life, but are also more susceptible to remembering these specific emotions as occurring even more (Urban *et al.* 2018)²⁸.

Questo *bias* di memoria nei soggetti con depressione potrebbe influenzare il giudizio diagnostico del clinico, in particolare l'uso degli specificatori di gravità.

²⁸ "La ricerca presentata in questo lavoro mostra che tristezza, ansia e rabbia sono ugualmente elevate nelle persone con una storia di depressione e che i ricordi di queste emozioni, in particolare della rabbia, tendono a essere sovrastimati. Di conseguenza, le persone con una storia di depressione non solo hanno maggiori probabilità di sperimentare più frequentemente tristezza, ansia e rabbia durante la vita quotidiana, ma sono anche più suscettibili a ricordare che queste specifiche emozioni si siano verificate con una frequenza ancora maggiore" (traduzione nostra).

CAPITOLO 1. LO STUDIO

1.1. OBIETTIVI

Intendiamo ora presentare la revisione sistematica con metanalisi da noi condotta sull'accordo tra valutatori (*inter-rater reliability*) rispetto alle diagnosi nosografiche dei disturbi dell'umore. Lo studio è stato condotto secondo le linee guida del Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA). La selezione degli articoli non ha incluso soltanto i disturbi dell'umore ma anche tutti gli altri disturbi mentali psichiatrici; indicheremo, nel sottoparagrafo 1.2 di questo capitolo, i criteri di esclusione e di inclusione degli studi su cui è stata eseguita la metanalisi.

La scelta di condurre una metanalisi sull'affidabilità diagnostica dei disturbi dell'umore deriva da due ordini di ragioni. La prima riguarda le numerose evidenze che indicano come il disturbo bipolare sia troppo spesso sottoposto ad errore diagnostico (Stensland *et al.* 2008, Bruchmüller & Meyer 2009) e che vi sia un'enorme divario temporale tra l'esordio del disturbo e quello della sua corretta diagnosi (Suppes *et al.* 2001, Mantere *et al.* 2004, Dagani *et al.* 2017). Inoltre, il risultato del *field trial* (Regier *et al.* 2013) condotto sul DSM-5 in relazione alla diagnosi di disturbo depressivo maggiore è stato molto insoddisfacente. Risultano pertanto essenziali ulteriori studi sul processo diagnostico dei disturbi dell'umore affinché questo possa essere migliorato. La seconda ragione che ci ha spinti a condurre questa metanalisi è il *gap* della letteratura sull'argomento; a nostra conoscenza, infatti, non esistono revisioni sistematiche con metanalisi aventi come oggetto di studio specificatamente l'*inter-rater reliability* nei disturbi dell'umore.

Come già menzionato nel paragrafo 3.2 dell'introduzione, l'affidabilità *inter-rater* riguarda il grado di accordo tra valutatori. La questione dell'accordo tra valutatori sorge a causa della variabilità tra gli osservatori umani. La misurazione del grado con cui i valutatori assegnano lo stesso punteggio alla stessa variabile è chiamata *inter-rater reliability* (McHuge 2012).

Per esprimere l'accordo tra valutatori il valore più usato è il kappa di Cohen. In passato l'accordo tra valutatori veniva misurato con la "percentuale di concordanza" (*percent agreement*), calcolata dividendo il numero di punteggi concordanti con il numero totale dei punteggi. Nel 1960, Jacob Cohen criticò l'utilizzo della percentuale di concordanza a causa della sua incapacità di tenere conto dell'accordo casuale. Introdusse quindi il kappa di Cohen,

sviluppato per tenere conto della possibilità che gli esperti, a causa dell'incertezza, "indovinino" su almeno alcune variabili. In altri termini, Cohen ipotizzò che in una situazione sperimentare i valutatori chiamati a dare il loro giudizio non sappiano con certezza quale punteggio attribuire ad almeno alcune variabili; l'accordo che si verificherebbe sarebbe pertanto "spurio". La statistica sottostante al valore kappa è stata dunque sviluppata per tenere sotto controllo il fattore di accordo casuale (McHuge 2012). Il valore di Kappa può variare da -1 a +1, dove 0 rappresenta l'accordo che ci si può aspettare dal caso, e 1 rappresenta un accordo perfetto tra i giudici. I valori inferiori a 0 sono molto rari (McHuge 2012).

In letteratura, per definire la forza di kappa, solitamente vengono seguiti gli intervalli indicati da Landis e Koch (1977):

- 0,8 – 1,00: accordo quasi perfetto
- 0,6 – 0,8: accordo sostanziale
- 0,4 – 0,6: accordo moderato
- 0,2 – 0,4: accordo discreto
- 0 – 0,2: accordo lieve
- < 0: scarso accordo.

Tuttavia, come sottolinea McHuge (2012), nei contesti sanitari, "any kappa below 0.60 indicates inadequate agreement among the raters and little confidence should be placed in the study results".

Questa revisione sistematica intendeva calcolare il valore dell'accordo tra valutatori per ogni macrocategoria diagnostica (disturbi dell'umore, disturbi d'ansia, disturbi psicotici, disturbi di personalità ecc.) presente negli studi inclusi. L'autore di questo elaborato ha scelto di occuparsi dei disturbi dell'umore; i risultati della metanalisi condotta sui disturbi dell'umore verranno esposti nel paragrafo 1.3 di questo capitolo.

1.2. MATERIALI E METODI

La ricerca, condotta secondo i criteri PRISMA (Liberati *et al.* 2015), è stata effettuata in uno dei principali databases psichiatrici, PubMed (National library of medicine) utilizzando la seguente chiave di ricerca: “(Inter-rater reliability) AND (psychiatry* OR psychiatric diagnosis OR schizophrenia OR psychosis OR personality disorder OR bipolar disorder OR anxiety disorder OR obsessive-compulsive disorder OR OCD OR post-traumatic stress disorder OR PTSD) NOT (Alzheimer OR Parkinson OR multiple sclerosis OR neurolog*)”. Come risulta visibile dalla stringa i disturbi degli studi su cui è stata eseguita la metanalisi sono disturbi psichiatrici, non disturbi neurologici. Sono stati inclusi tutti gli articoli originali scritti in lingua inglese dal 1973 al 21 marzo 2024, con abstract e testi completi disponibili.

I criteri di inclusione con cui è stata effettuata la selezione sono stati i seguenti:

- Articolo originale;
- età media dei pazienti ≥ 18 anni;
- diagnosi basata su un manuale diagnostico;
- *topic* dell’articolo centrato sull’*inter rater reliability* di diagnosi psichiatriche specifiche e/o generali.

I criteri di esclusione sono stati i seguenti:

- reviews, metanalisi, commenti, lettere, case reports, analisi combinate, commenti, case studies, protocolli di studio;
- età media dei soggetti < 18 anni;
- studi che non presentavano risultati specifici per l’*inter-rater reliability*;
- studi condotti su animali;
- articoli non scritti in lingua inglese;
- studi condotti su altre misure di affidabilità (es. *test retest*).

Quattro autori e un quinto autore come supervisore, hanno successivamente controllato ed estratto i dati dagli articoli inclusi nella piattaforma Rayyan. Per ogni studio selezionato sono stati estrapolati i seguenti dati:

- Numero di valutatori;
- numero dei soggetti del campione;
- manuale diagnostico;
- categoria diagnostica generale e/o specifica;

- modalità e/o strumento di valutazione;
- kappa di Cohen;
- intervallo di confidenza;
- errore standard;
- livello di formazione dei valutatori (basso/medio/alto).

L'errore standard, quando non presente negli studi, è stato estrapolato mediante analisi statistiche attraverso il numero campionario, la deviazione standard o il valore dell'intervallo

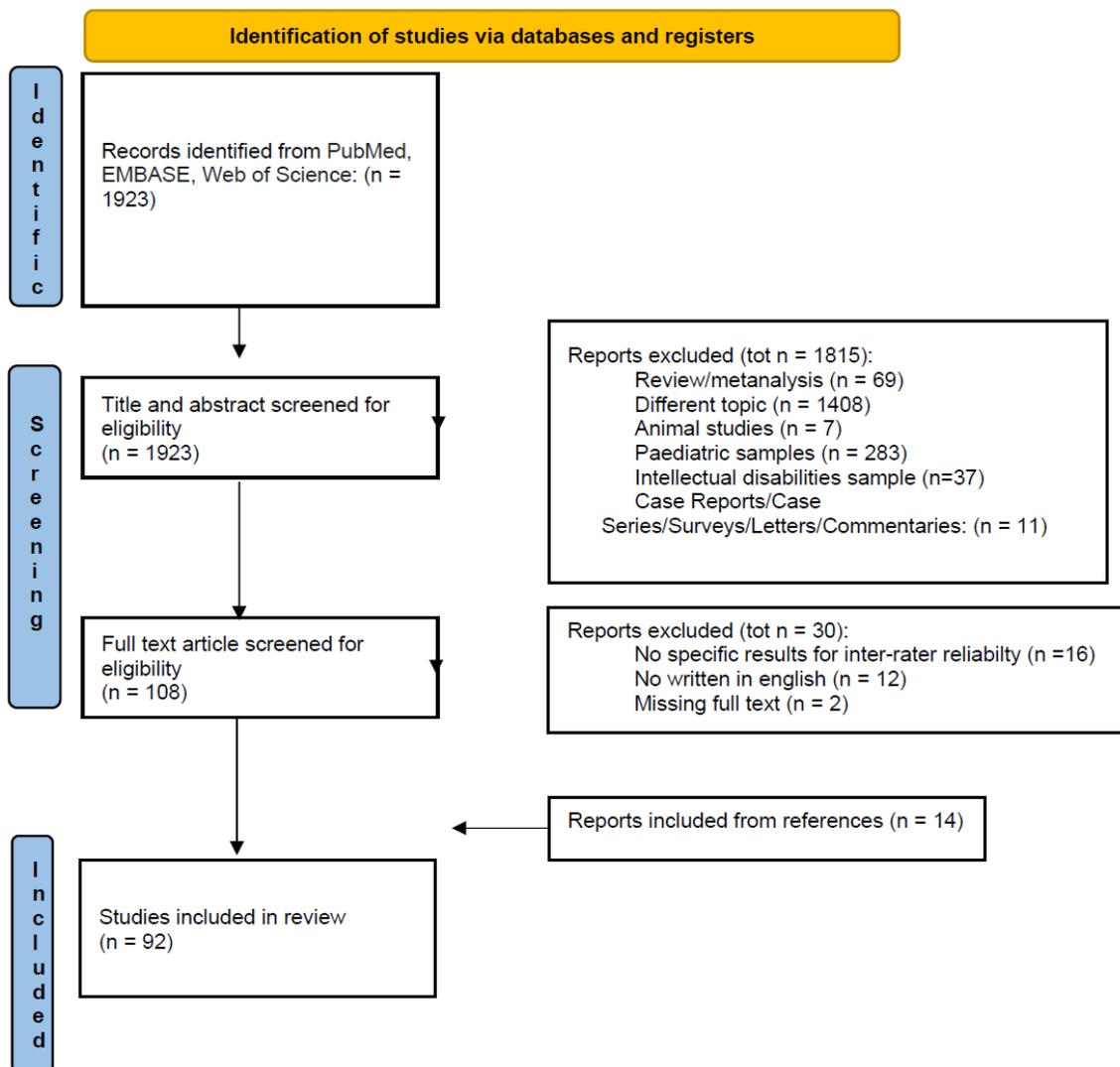


Figura 1. Diagramma di flusso della selezione degli articoli condotta secondo i criteri PRISMA.

di confidenza. È da sottolineare che, come abbiamo già anticipato, per ogni studio è stato estratto il valore di kappa per ogni diagnosi che rispettasse i criteri di inclusione; non è stata dunque estrapolata la media di kappa dei disturbi inclusi nella metanalisi. Il valore di kappa

di ogni disturbo specifico è stato poi conteggiato all'interno della metanalisi specifica per la categoria generale del disturbo (disturbi dell'umore, disturbi d'ansia, disturbi psicotici, disturbi di personalità ecc.). In questa seconda fase di screening per diverse ragioni 30 articoli sono stati esclusi (assenza del valore di affidabilità *inter-rater*, lingua diversa dall'inglese, no *full text*). La ricerca è stata infine ampliata con 14 studi attraverso l'elenco bibliografico degli studi preliminarmente selezionati (78 studi); come si può vedere dal diagramma di flusso riassuntivo (figura 1), il numero definitivo degli articoli inclusi è pertanto pari a 92. Dei 92 studi inclusi, 38 studi presentavano categorie diagnostiche dei disturbi dell'umore con l'*inter-rater reliability*; in questo lavoro presenteremo i risultati della metanalisi condotta sulle 95 categorie diagnostiche dei disturbi affettivi presenti nei 38 studi. La metanalisi sui disturbi dell'umore è stata condotta solo sui valori con kappa di Cohen, nonostante dalla fase di screening fossero stati selezionati altri cinque studi condotti sui disturbi dell'umore con l'*inter-rater reliability* espressa in ICC. Il numero di studi con ICC è risultato tuttavia insufficiente per condurre una seconda metanalisi.

La valutazione della qualità dello studio è stata effettuata secondo il Cochrane risk-of-bias 2 (ROB-2) per valutare il rischio *bias* negli studi randomizzati.

Dato che l'eterogeneità è risultata molto elevata, sia per quanto riguarda l'intervallo di confidenza, sia rispetto alla dimensione del campione degli studi, e questo era qualitativamente rilevabile già dopo la fase di estrapolazione dei dati, l'analisi dei dati ai fini della metanalisi non è stata condotta attraverso il modello a effetti fissi, bensì attraverso il modello a effetti casuali. Nel *random effect model* "si assume che la variabilità tra i risultati ottenuti dagli studi primari sia dovuta non solo a un effetto di tipo casuale legato al campionamento, ma anche a delle differenze reali tra gli studi", in altre parole si assume che "il dato numerico aggregato fornito dalla metanalisi non rappresenti [...] la stima di un unico effetto comune [come nel modello a effetti fissi], bensì la stima della media di una distribuzione di effetti differenti" (Bioletto *et al.* 2022). Nel modello a effetti casuali viene così ridotto il peso del campione del singolo studio e viene allargato l'intervallo di confidenza dell'*effect size* globale.

L'analisi dei dati è stata condotta attraverso i software RStudio, che è stato utilizzato per la creazione dei grafici *Forest plot* e *Funnel plot*, e attraverso Jamovi.

1.3 RISULTATI

Gli studi inclusi nella revisione sistematica sui disturbi dell'umore sono stati 38, per un totale di 95 categorie diagnostiche su cui è stata eseguita la metanalisi. L'applicazione di ROB-2 agli studi ha dato un rischio *bias* che varia da basso (35 studi) a medio (3 studi).

Nella tabelle seguenti esponiamo alcuni dei dati estratti dagli studi. In nota alla tabella 7 riportiamo gli acronimi delle diagnosi e degli strumenti di valutazione delle tabelle 2-7.

Studio	Diagnosi	Forma di valutazione	Manuale diagnostico	Kappa	Soggetti	N° valutatori e livello di formazione	Rischio <i>bias</i>
Ransing <i>et al.</i> (2020)	AD	PHQ-2	PHQ-2	0.68	108	2 – alto: psicologo clinico e infermiere; training	basso
	AD	PHQ-9	PHQ-9	0.732			
Avrichir <i>et al.</i> (2002)	DYS	SCID-I/P 2.0	DSM-IV	0.76	81	2 – alto: clinici	basso
Ptok <i>et al.</i> (2001)	MDD	Questionario sulla storia familiare	ICD-10	0.72	230	2 – medio: laici (studenti di medicina); training	basso
Medina-Mora <i>et al.</i> (2019)	MDD	CDDG per l'ICD-11	ICD-11	0.6	153	2 – alto: clinici (psichiatri e psichiatri in formazione); training	medio
Lauterbach <i>et al.</i> (2004)	MDD	AMDP-DS;	ICD-10	0.67	50	3 – alto: clinici	basso

Tabella 2. Dati estrapolati dagli studi.

Studio	Diagnosi	Forma di valutazione	Manuale diagnostico	Kappa	Soggetti	N° Valutatori e livello di formazione	Rischio bias
Smith-Nielsen <i>et al.</i> (2018)	PD	SCID-5	DSM-5	0.89	70	2 – alto: clinici; training	basso
	PD		ICD-10	0.94			
Dijksman <i>et al.</i> (2016)	MDD	DS-1; DD-2; DS-3-5	DSM-IV	0.67	100	2 – alto: clinici	basso
	DD NOS			0.35			
Rocha Neto <i>et al.</i> (2023b)	BD	LEAD	ICD-10	0.75	99	2 – medio: laici (studenti di medicina); training	medio
	DD NOS			0.68			
Shabani <i>et al.</i> (2021)	BD	SCID-5-CC	DSM-5	0.8	245	2 – alto: clinici	basso
	DD NOS			0.69			
Lobboestael <i>et al.</i> (2011)	MDD	SCID-I	DSM-IV	0.66	151	2 – alto: clinici; training	basso
	DYS			0.81			
Malhotra <i>et al.</i> (2015)	MD	Piattaforma online	ICD-10; DSM-IV	0.87	274	2 – alto: clinici; training	basso
	BD			0.93			
	MDDr			0.86			
Gutkovich <i>et al.</i> (2013)	MDD	SCID IV v. russa	DSM-IV	0.76	35	2 – alto: clinici; training	basso
	DYS			1			
	DD NOS			0.65			
Kralj <i>et al.</i> (2017)	BD	DIGS v. croata	ICD-10	0.956	150	3 – alto: clinici	basso
	MDD			0.493			

Tabella 3. Dati estrapolati dagli studi.

Studio	Diagnosi	Forma di valutazione	Manuale diagnostico	Kappa	Soggetti	N° valutatori e livello di formazione	Rischio di bias
Onofa <i>et al.</i> (2019)	BD I	ICD-11	ICD-11	0.83	132	2 – alto: clinici; training	basso
Kortmann <i>et al.</i> (1988)	MD	DSM-III	DSM-III	0.78	40	2 – basso: psichiatra olandese in un contesto etiope e studente di medicina	medio
Wittchen <i>et al.</i> (1991)	DYS	CIDI	ICD-10; DSM-III	0.96	575	2 – alto: 53 clinici, 55 non clinici; training	basso
	MDD			0.97			
	MDDr			0.93			
	BD I			0.92			
	BD II			0.94			
Preisig <i>et al.</i> (1999)	BD	DIGS v. francese	DSM-IV	0.85	136	2 – alto: clinici; training	basso
	DD NOS			0.93			
Pang <i>et al.</i> (1997)	MD	ICD-10	ICD-10	0.65	238	2 – alto: clinici	basso
Andrews <i>et al.</i> (1995)	MDD	CIDI;	DSM-III	1	29	2 – alto: clinici; training	basso
	MDD	SCAN		0.37			
Quintana <i>et al.</i> (2004)	BD	CIDI 2.1 v. portoghese	ICD-10	0.74	186	2 – medio: 13 studenti di medicina con training; 3 clinici	basso
	MDD			0.94			
	DYS			0.87			
Sharma <i>et al.</i> (2004)	MDD	GMHAT-PC	ICD-10	0.82	56	2 – alto; clinici	basso

Tabella 4. Dati estrapolati dagli studi.

Studio	Diagnosi	Forma di valutazione	Manuale diagnostico	Kappa	Soggetti	N° valutatori e livello di formazione	Rischio bias
Dahl <i>et al.</i> (1995)	MDD	SPIFA/SPIFA	DSM-IV	0.70	164	2 – alto; clinici (psichiatri e medici di famiglia)	basso
	BD			0.68			
	MDD	SPIFA/SCID	DSM-IV	0.71			
Roca <i>et al.</i> (2007)	MDD	DIGS v. spagnola	DSM-IV	0.877	91	2 – alto: clinici	basso
	BD I			0.933			
	MDD	LEAD		0.704			
	BD I			0.825			
Okasha <i>et al.</i> (1993)	MDD	PSE v. arabo-inglese	ICD-9	0.86	100	3 – alto: clinici	basso
	MDD		ICD-10	0.62			
	MDD		DSM-III-R	0.85			
Regier <i>et al.</i> (1994a)	MD	ICD-10	ICD-10	0.79	1969	2; alto: clinici	basso
	MDD			0.69			
	BD			0.83			
	MDDr			0.72			
Regier <i>et al.</i> (1994b)	MD	ICD-10	ICD-10	0.66	491	2 – alto: clinici	basso
	MDD			0.48			
	BD			0.69			
	MDDr			0.53			
Andreasen <i>et al.</i> (1992)	MDD	CASH	DSM-III-R	0.65	30	2 – alto; clinici; training	basso

Tabella 5. Dati estrapolati dagli studi.

Studio	Diagnosi	Forma di valutazione	Manuale diagnostico	Kappa	Soggetti	N° valutatori e livello di formazione	Rischio di bias
Osorio <i>et al.</i> (2019)	BD II	V. brasiliana SCID-5-CV	DSM-5	0.94	180	2 – alto: clinici	basso
	MDD			0.82			
	MD			0.98			
Chmielewski <i>et al.</i> (2015)	MDD	SCID-I/P	DSM-5	0.92	1000	2 – non valutabile (non viene riportato se siano clinici o laici; training)	basso
	DYS			0.75			
	BD I			0.73			
	BD II			0.64			
Reed <i>et al.</i> (2018)	BD I	CDDG per l'ICD-11	ICD-11	0.84	1806	2 – alto: clinici	basso
	BD II			0.62			
	MDD			0.64			
	MDDr			0.74			
	DYS			0.45			
Nagar <i>et al.</i> (2018)	MDD	SCID-II	DSM-IV	0.29	170	2 alto: coppia formata da due clinici di cui uno solo ha ricevuto il training	basso
	DYS			0.03			
Cheniaux <i>et al.</i> (2008)	BD	CIDI	DSM-IV	0.72	200	2 – alto: clinici	basso
	MDD			0.41			
	BD		ICD-10	0.72			
	MDD			0.54			
Zanarini <i>et al.</i> (2000)	MDD	SCID-I	DSM-IV	0.8	27	2 – medio: studenti di master o phd; training	basso
	DYS			0.76			

Tabella 6. Dati estrapolati dagli studi.

Studio	Diagnosi	Forma di valutazione	Manuale diagnostico	Kappa	Soggetti	N° valutatori e livello di formazione	Rischio bias
Roy <i>et al.</i> (1997)	BD I	SCID	DSM-III-TR	0.81	133	2 – alto: psichiatri (ricercatori e clinici)	basso
	BD II			0.69			
	MDDr			0.76			
	MDD			0.85			
Sartorius <i>et al.</i> (1995)	MD	ICD-10-DCR	ICD-10-DCR	0.83	4322	2; alto: clinici	basso
	BD			0.82			
	MDD			0.73			
	DYS			0.58			
Sundqvist <i>et al.</i> (2008)	MDD	DIBS; SCAN	DSM-III-R	0.95	58	2 – basso: infermieri	basso
	BD			0.7			
	MD			0.84			
Joo <i>et al.</i> (2004)	BD	DIGS v. coreana	DSM-IV	1	24	2 – medio: infermieri; training	basso
		MDD			0.83		
Maj <i>et al.</i> (2000)	MDD	CIDI	DSM-IV	0.82	150	2 – alto: clinici; training	basso
	Manic Disorder			0.71			

Tabella 7. Dati estrapolati dagli studi. AD (Antenatal Depression); PD (Postnatal Depression); DYS (Dysthymia); MD (Mood Disorder); MDD (Major depressive Disorder); MDDr (Major Depressive Disorder recurrent); BD (Bipolar Disorder), DD NOS (Depressive Disorder Not Otherwise Specified). Patient Health Questionnaire (PHQ); SCID (Structured Clinical Interview for DSM); AMPD-DS (Alternative Model of Personality Disorders – Depression Scale); DIBS (Diagnostic Instrument for Brain Studies); DIGS (Diagnostic Interview for Genetic Studies); CIDI (Composite International Diagnostic Interview); SPIFA (Structured Psychiatric Interview for General Practice); GMHAT-PC (Global Mental Health Assessment Tool – Primary

Care); CDDG (Clinical Descriptions and Diagnostic Guidelines) per l'ICD. SCAN (Schedules for Clinical Assessment in Neuropsychiatry); PSE (Present State Examination); CASH (Comprehensive Assessment of Symptoms and History); ICD DCR (Diagnostic Criteria for Research); LEAD (Longitudinal Expert All Data); LEAD (Longitudinal Expert All Data).

Lo studio di Regier *et al.* 1994, dato che presentava due campioni diversi è stato diviso in due parti ai fini della metanalisi: Regier *et al.* 1994a si riferisce alla valutazione con l'ICD-10 del campione globale, Regier *et al.* 1994b si riferisce alla valutazione dell'ICD-10 all'interno del campione canadese e americano.

Il valore medio di kappa emerso dalla metanalisi è 0.741 (IC 95%: 0.706, 0.776) (cfr. tabella 8) Questo valore, in accordo con Landis e Koch, indica un accordo tra valutatori sostanziale per quanto concerne l'*inter-rater reliability* dei disturbi dell'umore.

	Estimate	se	Z	p	CI Lower Bound	CI Upper Bound
Intercept	0.741	0.0179	41.4	<.001	0.706	0.776

Tabella 8. Risultati generali della metanalisi elaborati dal software Jamovi. È presente l'*effect size* globale e il suo intervallo di confidenza.

La metanalisi condotta ha avuto una eterogeneità molto elevata, come mostra l'indice di eterogeneità di Higgins ($I^2= 95\%$), il quale rappresenta la percentuale di variazione totale tra gli studi attribuibile all'eterogeneità piuttosto che al caso. Inoltre, il valore di $Q=1905.292$ con $p < 0.001$ conferma la significativa eterogeneità statistica tra gli studi (cfr. tabella 9).

Heterogeneity Statistics

Tau	Tau ²	I ²	H ²	R ²	df	Q	p
0.148	0.022 (SE= 0.0042)	94.89%	19.559	.	94.000	1905.292	<.001

Tabella 9. Eterogeneità della metanalisi, con l'indice Higgins (I^2) e il *p value*.

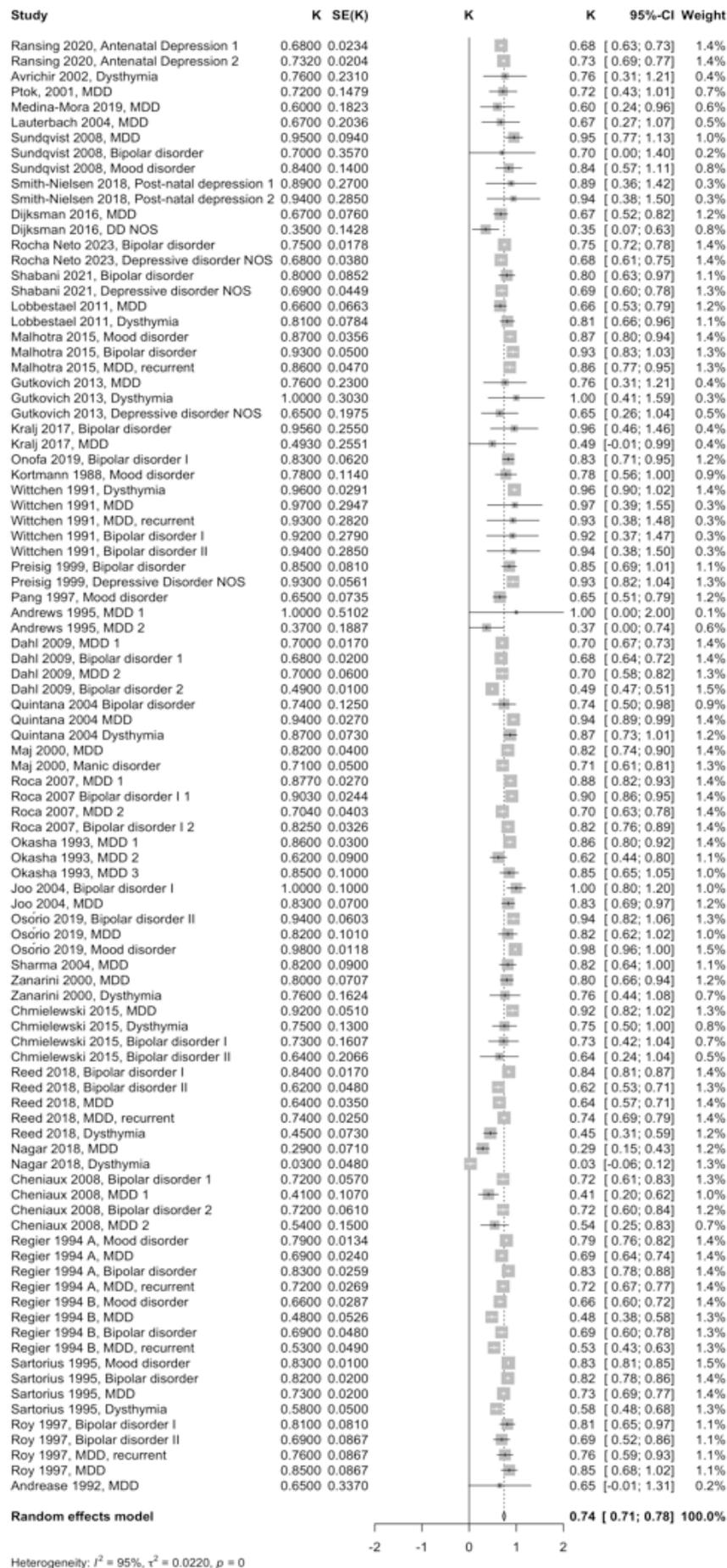


Figura 2. Forest plot prodotto con il software RStudio.

Dal *forest plot* (figura 2) emerge che le diagnosi più significative, che presentano cioè un valore kappa superiore o uguale all'*effect size* globale, sono 53 su 95. Considerata la media di kappa delle diagnosi di ogni singolo studio, gli studi più significativi, cioè con una media di kappa superiore o uguale a 0.74 sono stati i seguenti 23 studi su 38: Wittchen *et al.* 1991; Smith-Nielsen *et al.* 2018; Joo *et al.* 2004; Osório *et al.* 2019; Preisig *et al.* 1999; Malhotra *et al.* 2015; Quintana *et al.* 2004; Sundqvist *et al.* 2008; Onofa *et al.* 2019; Roca *et al.* 2007; Sharma *et al.* 2004; Reed *et al.* 2018; Gutkovich *et al.* 2013; Kortmann *et al.* 1988; Zanarini *et al.* 2000; Okasha *et al.* 1993; Roy *et al.* 1997; Avrighir *et al.* 2002; Maj *et al.* 2000 Chmielewski *et al.* 2015; Regier *et al.* 1994A; Shabani *et al.* 2021, Sartorius *et al.* 1995.

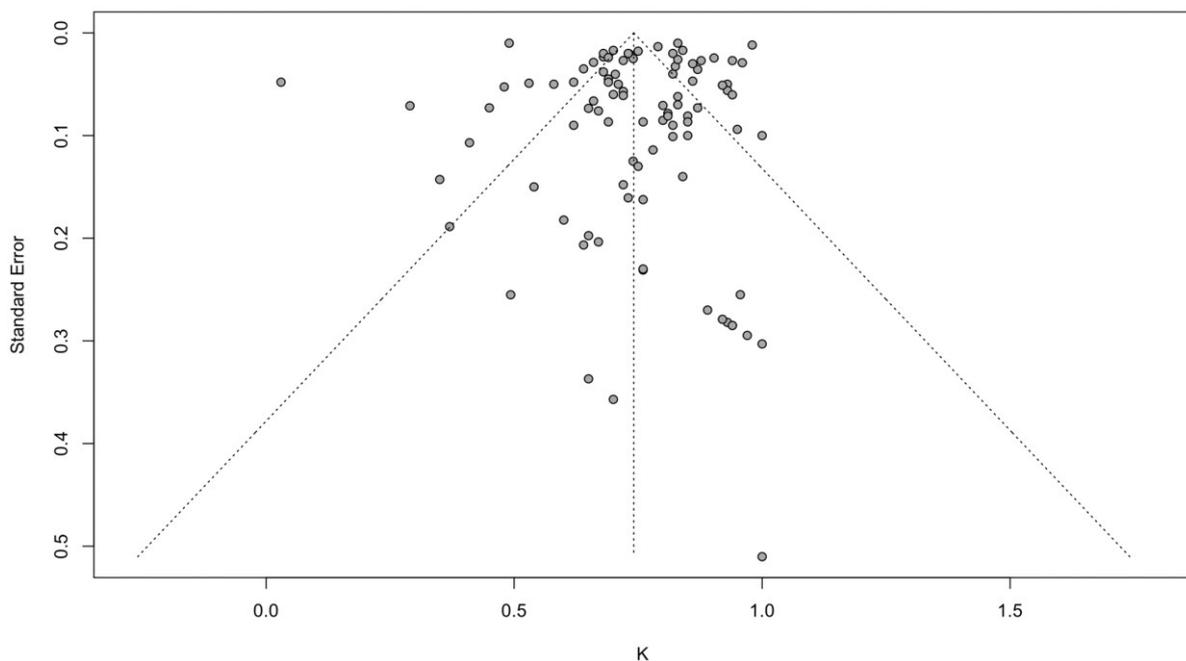


Figura 3. *Funnel plot* prodotto con il software RStudio. Il grafico indica l'eventuale presenza di un *bias* di pubblicazione.

Il *funnel plot* (figura 3) mostra la presenza di *bias* di pubblicazione in quanto manca una distribuzione equa dei punti attorno alla linea centrale. La maggior parte degli studi si trova infatti nella parte superiore del grafico, suggerendo la potenziale presenza di *bias* di pubblicazione.

CAPITOLO 2. DISCUSSIONE

Gli strumenti validati utilizzati in almeno 3 studi sui 38 analizzati sono: SCID (Structured Clinical Interview for DSM); CIDI (Composite International Diagnostic Interview) e DIGS (Diagnostic Interview for Genetic Studies). In particolare, 6 dei 19 studi più significativi hanno utilizzato l'intervista strutturata SCID; 3 hanno utilizzato l'intervista semi-strutturata DIGS e 3 l'intervista semi-strutturata CIDI. Fra i 19 studi meno significativi sono state comunque utilizzate queste interviste standardizzate, ma in numero inferiore: in 4 studi è stata utilizzata la SCID, in solo uno studio la DIGS e in 2 studi la CIDI. Rileviamo pertanto un'associazione positiva fra l'uso di queste scale e una maggiore forza dell'effetto osservato. L'associazione naturalmente non implica relazione causale dato che altri fattori, come le caratteristiche degli studi potrebbero aver influenzato la dimensione dell'effetto.

Come abbiamo visto, dal *funnel plot* emerge la presenza di un *bias* di pubblicazione. Il *bias* di pubblicazione si verifica nella ricerca scientifica quando la decisione di pubblicare uno studio è influenzata dai risultati ottenuti (Bioletto *et al.* 2022); gli studi con risultati significativi tendono infatti ad essere pubblicati di più rispetto a quelli con risultati non significativi, come dimostrato da Dickersin e colleghi (1992), alterando in questo modo la reale comprensione del fenomeno oggetto dello studio. La presenza di un *bias* di pubblicazione in una metanalisi sta ad indicare che i risultati sono sovrastimati in senso positivo. Nel *funnel plot* gli studi inclusi dovrebbero essere coerentemente distribuiti all'interno del triangolo per scartare l'ipotesi di un *bias* di pubblicazione. Al contrario risultano eccessivamente concentrati nella parte alta del triangolo e al di fuori del triangolo (*outliers*). Questa distribuzione indica come ci siano troppi studi con un errore standard basso, correlato a una numerosità del campione ampia. La probabilità che uno studio venga pubblicato aumenta di fatti con la dimensione del campione in virtù della qualità dell'evidenza che un'elevata dimensione del campione di per sé conferisce (Bioletto *et al.* 2022). La presenza di *outliers* indica invece la presenza di studi che si discostano troppo dall'*effect size* globale.

Ulteriori studi di dovrebbero cercare di ridurre il *bias* di pubblicazione svolgendo una selezione più rigorosa degli studi sulla base del metodo diagnostico da questi utilizzato.

L'analisi del *bias* di pubblicazione ha fornito un numero Fail-safe di 875.871.000 (cfr. tabella 10); tale valore indica che per modificare significativamente i risultati sarebbero necessari un numero estremamente elevato di studi.

Publication Bias Assessment

Test Name	value	p
Fail-Safe N	875871.000	<.001

Tabella 10. Numero Fail-Safe calcolato attraverso il software Jamovi.

Dal calcolo della media di kappa per ogni tipo di categoria nosografica inclusa nella metanalisi, si sono ottenuti i risultati esposti nella tabella 11. Il disturbo con l'accordo tra va-

Disturbo	Kappa	N° osservazioni
PD	0,91	2
BD I	0,85	8
MD	0,80	8
BD II	0,77	5
BD	0,76	14
MDDr	0,75	6
MDD	0,71	34
Manic disorder	0,71	1
AD	0,71	2
DYS	0,69	10
DD NOS	0,66	5

Tabella 11. Categorie nosografiche presenti negli studi inclusi nella metanalisi presentate in ordine crescente sulla base del valore medio di kappa per ogni tipo di categoria.

-lutatori più alto è la Depressione post-partum con kappa pari a 0.91. Contro le nostre ipotesi iniziali, legate alle evidenze della letteratura che riportano un tasso molto frequente di errata diagnosi di disturbo bipolare (Stensland *et al.* 2010, Bruchmüller & Meyer 2009), il BD ha riportato un kappa di 0.76; questo valore si accorda con quello indicato nella metanalisi di Santelman *et al.* (2016)²⁹, avente per oggetto di studio l'accordo tra valutatori nel disturbo schizoaffettivo e nelle sue diagnosi differenziali, compresi dunque i disturbi affettivi. Infatti, nella metanalisi di Santelman, che ha incluso 25 studi, l'*inter-rater reliability* per il disturbo bipolare si attesta su un valore di kappa pari a 0.82. Anche il valore di kappa per il disturbo

²⁹ Rispetto ai disturbi dell'umore le uniche categorie diagnostiche presenti nello studio di Santelman *et al.* (2016) sono il "disturbo bipolare" e la "depressione unipolare".

depressivo maggiore pari a 0,71 è coerente con i risultati di Santelman *et al.* (2016), che indica un kappa di 0.75 per la “depressione unipolare”.

Come ci aspettavamo il “Depressive Disorder Not Otherwise Specified” (DD NOS), ridenominato nel DSM-5 “Unspecified depressive disorder” ha avuto il più basso accordo tra valutatori (kappa 0.66). Questo è probabilmente dovuto alla stessa natura residuale della categoria diagnostica che serve a rispondere a quelle condizioni che non soddisfano pienamente i criteri per uno qualsiasi dei disturbi della classe diagnostica dei disturbi depressivi e nei casi in cui non ci sono informazioni sufficienti per porre una diagnosi più specifica (APA 2014, p. 212).

L’eterogeneità dei risultati è risultata essere molto elevata, attestandosi infatti al 95%, come indicato dall’indice di Higgins; si può escludere pertanto che le differenze fra gli studi siano dovute solo al caso. Le spiegazioni dell’elevata eterogeneità possono essere diverse. Ne elenchiamo di seguito alcune:

- Metodo diagnostico. Gli studi, nonostante abbiano utilizzato perlopiù interviste standardizzate, hanno utilizzato *tipi* diversi di interviste; alcuni, inoltre, si sono basati sul giudizio clinico (Roca *et al.* 2007, Rocha Neto *et al.* 2023b); altri sulle linee guida dell’ICD (Medina-Mora *et al.* 2019, Reed *et al.* 2018); altri su questionari self-report (Ransing *et al.* 2020, Nagar *et al.* 2018).
- Diagnosi oggetto dello studio. Gli studi inclusi nella metanalisi hanno considerato tipi diversi di categorie diagnostiche relative ai disturbi dell’umore. Non tutti gli studi, in altri termini, hanno presentato un accordo tra valutatori per le stesse categorie diagnostiche.
- Disegno di ricerca. La maggior parte degli studi inclusi ha un disegno di ricerca trasversale; alcuni studi inclusi sono retrospettivi (Sundqvist *et al.* 2008, Rocha Neto *et al.* 2023b).
- Campione. La popolazione differiva negli studi per grandezza del campione (tendenzialmente alto), età media, genere, livello di istruzione, occupazione, nazione di provenienza ecc..
- Valutatori. I livelli di esperienza e di formazione dei valutatori variano all’interno degli studi. Non è stato possibile considerare il livello di esperienza dei valutatori poiché troppi studi non riportavano la media degli anni di esperienza dei valutatori; è stato dato quindi un giudizio ai valutatori dello studio sulla base del loro livello di formazione. A 2 studi

su 38 è stato dato un giudizio “basso” (Sundqvist *et al.* 2008, Kortmann *et al.* 1988), a 5 studi un giudizio “medio”, a 30 studi un giudizio “alto”; uno studio (Chmielewski *et al.* 2015) non era valutabile perché non specificava se i valutatori fossero clinici o laici. Alcuni studi (Sundqvist *et al.* 2008, Rocha Neto *et al.* 2023b) hanno inoltre, svolto la valutazione a partire da cartelle cliniche, senza incontrare direttamente i soggetti.

Nonostante l’elevata eterogeneità, dai risultati della metanalisi emerge un *effect size* globale significativo; esso si è attestato a kappa 0.74, che definisce un accordo “sostanziale” secondo Landis e Koch (1977). Nonostante sia “sostanziale” l’accordo tra valutatori non è tuttavia “quasi perfetto” (kappa 0.80-1.0). Le ragioni per cui l’accordo non è stato “quasi perfetto” potrebbero essere le seguenti:

- Comorbilità: molti individui presentano più di un disturbo mentale contemporaneamente. Distinguere i sintomi e attribuirli a un disturbo specifico può essere complesso e portare a divergenze di giudizio tra i valutatori.
- Assenza di biomarcatori: la letteratura non ha individuato allo stato attuale biomarcatori dei disturbi dell’umore che godano di ampio consenso e quindi utilizzabili nella pratica clinica (Kennis *et al.* 2019; Vega-Núñez *et al.* 2022).
- Interpretazione dei criteri diagnostici: i criteri diagnostici possono essere talvolta vaghi o comunque suscettibili di interpretazioni diverse da parte dei clinici.
- Variabilità della patologia: i disturbi dell’umore hanno manifestazioni molto variabili. Proprio a causa di questa variabilità della patologia sono nati i modelli spettrali dei disturbi dell’umore di Akiskal e Cassano, che mirano a cogliere la diversificata gamma di manifestazioni cliniche che si colloca nel continuum tra un polo maniacale e un polo depressivo “puri”, come abbiamo visto nei sottoparagrafi 2-2.3 dell’introduzione. La multiforme fenomenologia dei disturbi dell’umore rende più difficile un accordo “quasi perfetto” o “perfetto” tra valutatori.
- *Bias* cognitivi: come è stato esposto nel sottoparagrafo 3.2.2, esistono diversi *bias* cognitivi che influenzano inconsapevolmente la valutazione dei sintomi tanto da parte dei clinici, quanto da parte dei pazienti. Una distorsione cognitiva tipica del clinico, che può entrare in gioco per tutti i disturbi mentali, è il *bias* della disponibilità. Dato che alcuni sintomi sono comuni a più di una psicopatologia, ciò può portare il clinico a propendere per una certa diagnosi piuttosto che per un'altra sulla base della diversa “disponibilità alla mente” del clinico di una certa psicopatologia. Per quanto riguarda il

clinico, nella valutazione del disturbo bipolare sembra entrare in azione in particolare il bias della rappresentatività (Bruchmüller e Meyer 2009): i clinici tendono a considerare il “ridotto bisogno di sonno” come un criterio prototipico del disturbo bipolare, e ciò produce errori diagnostici. A causa delle distorsioni cognitive in atto, esperti diversi possono interpretare in modo diverso gli stessi sintomi e ciò può portare ad una riduzione del grado di accordo tra valutatori. Sul versante dei pazienti, è noto che soggetti con disturbo bipolare, soprattutto durante le fasi maniacali e ipomaniacali tendono a non riportare spontaneamente i sintomi della mania e dell’ipomania quali l’umore elevato, l’iperattività e l’impulsività e ciò può rendere difficile ai clinici identificare il disturbo bipolare (Hirschfeld *et al.* 2003). Inoltre, pazienti con depressione tendono, a presentare un *bias* di memoria; le persone con una storia di depressione, infatti, non solo sperimentano più emozioni negative e meno emozioni positive nella vita quotidiana, ma sovrastimano la frequenza delle emozioni negative nei loro ricordi in misura maggiore rispetto alle persone senza una storia di depressione (Urban *et al.* 2018). Questi *bias* oltre a influenzare negativamente sul processo diagnostico, possono far diminuire l’accordo tra valutatori.

Le interviste semi-strutturate e strutturate sono state validate con la finalità di ridurre alcuni dei *bias* più comuni a cui va in incontro il processo diagnostico; le implicazioni cliniche sono pertanto che, per migliorare il processo diagnostico dei disturbi dell’umore, i clinici dovrebbero utilizzare interviste semi-strutturate e strutturate ampiamente validate (come la SCID) in aggiunta a test di screening specifici per i disturbi dell’umore con elevata sensibilità come il PHQ-9 e l’HCL-32, esposti nel sottoparagrafo 3.1. Per la diagnosi differenziale fra disturbo bipolare e disturbo depressivo maggiore, potrebbe inoltre risultare utile l’utilizzo complementare di questionari self-report con un approccio dimensionale, come il TEMPS-A e il MOODS-SR, presentati nel sottoparagrafo 2.1 e 2.2, in grado di rilevare le caratteristiche temperamentali e le diverse manifestazioni della mania, comprese quelle sottosoglia.

Il presente studio è caratterizzato da alcune limitazioni. Come è stato precedentemente esposto, è indubbia la presenza di un bias di pubblicazione sebbene quest’ultimo non interferisca con la significatività dello studio. Inoltre, l’eterogeneità fra gli studi è risultata essere molto elevata. Ulteriori ricerche sarebbero quindi necessarie al fine di ottenere una minore eterogeneità tra gli studi selezionati.

CONCLUSIONE

La letteratura evidenzia un intervallo di 6-10 anni tra l'insorgenza del disturbo bipolare e l'identificazione della corretta diagnosi (Suppes *et al.* 2001, Mantere *et al.* 2004, Dagani *et al.* 2017). Il riconoscimento del disturbo bipolare è difficile e molto frequente è l'errata diagnosi (Stensland 2010, Bruchmüller & Meyer 2009), anche a causa dei *bias* del clinico e del paziente. I risultati del *field trial* (Regier *et al.* 2013) sul DSM-5 in relazione alla diagnosi di disturbo depressivo maggiore sono stati molto insoddisfacenti. Risultano pertanto essenziali studi sulle caratteristiche del processo diagnostico dei disturbi dell'umore affinché questo possa essere migliorato.

In questo elaborato è stata esposta la revisione sistematica con metanalisi sull'accordo tra valutatori nei disturbi dell'umore da noi condotta. Da quanto siamo a conoscenza si tratta della prima metanalisi avente per oggetto di studio specificatamente l'*inter-rater reliability* nelle diagnosi di disturbo dell'umore. La selezione degli articoli, condotta secondo le linee guida PRISMA, ha previsto lo *screening* di 1923 *abstract* di articoli. Gli studi inclusi sono stati 38. L'analisi dei dati è stata condotta attraverso il modello a effetti *random*.

Dai risultati è emerso un *effect size* medio significativo e sostanziale che si attesta a kappa 0.74, con un intervallo di confidenza ristretto (0.71-0.78). Il disturbo con la media di kappa più alta è la Depressione post-partum (kappa 0.91), mentre quello con l'affidabilità meno significativa è il disturbo depressivo senza specificazione (kappa 0.66). Contro le aspettative il disturbo bipolare ha riportato un kappa medio che si attesta a 0.76. I risultati della metanalisi mostrano un'associazione positiva tra l'utilizzo delle scale standardizzate SCID, CIDI e DIGS e la dimensione dell'effetto osservato; l'associazione non è tuttavia generalizzabile. Come suggerito dalla metanalisi di Rocha Neto *et al.* (2023a), che ha come oggetto di studio l'accordo tra valutatori per le diagnosi di schizofrenia e disturbo bipolare, sarebbe auspicabile che gli studi successivi sull'*inter-rater reliability* andassero a indagare l'accordo tra valutatori confrontando l'uso di interviste strutturate con l'uso di metodi non standardizzati (ad esempio il solo giudizio clinico). I risultati di questi studi consentirebbero ad ulteriori revisioni sistematiche con metanalisi di verificare che il procedimento standardizzato migliori l'affidabilità diagnostica dei disturbi affettivi.

Tra i limiti della presente metanalisi vi è in primo luogo la presenza di un *bias* di pubblicazione sebbene quest'ultimo non interferisca con la significatività dello studio. In secondo luogo, l'eterogeneità fra gli studi è risultata essere molto elevata con un indice di

Higgins pari al 95%; in terzo luogo, non sono state utilizzate tecniche statistiche per ridurre l'eterogeneità della metanalisi. Metanalisi future sull'argomento dovrebbero cercare di ridurre l'eterogeneità e il *bias* di pubblicazione svolgendo una selezione rigorosa degli studi sulla base del metodo diagnostico utilizzato.

BIBLIOGRAFIA

Akiskal H.S. (1983). *The bipolar spectrum: new concepts in classification and diagnosis*. In: Grinspoon L (ed.). *Psychiatry update: the American Psychiatric Association annual review*. Vol. 2. Washington, DC: American Psychiatric Press, 1983; 271-92.

Akiskal H.S., Mallya G. (1987). Criteria for the “soft” bipolar spectrum: treatment implications. *Psychopharmacol Bull*, 1987;23:68-73.

Akiskal H.S., Placidi G.F., Maremmani I., Signoretta S., Liguori A. *et al.* (1998a). TEMPS-I: delineating the most discriminant traits of the cyclothymic, depressive, hyperthymic and irritable temperaments in a nonpatient population. *J Affect Disord*. 1998; 51(1):7-19. doi: 10.1016/s0165-0327(98)00152-9.

Akiskal H.S. (1998b). Toward a definition of generalized anxiety disorder as an anxious temperament type. *Acta Psychiatr Scand Suppl*. 1998; 393:66-73. doi: 10.1111/j.1600-0447.1998.tb05969.x.

Akiskal H.S., Hantouche E.G., Bourgeois M.L., Azorin J.M., Sechter D. *et al.* (1998c). Gender, temperament, and the clinical picture in dysphoric mixed mania: findings from a French national study (EPIMAN). *J Affect Disord*. 1998 Sep;50(2-3):175-86. doi: 10.1016/s0165-0327(98)00113-x.

Akiskal H.S., Pinto O. (1999). The evolving bipolar spectrum. Prototypes I, II, III, and IV. *Psychiatr Clin North Am.*, 1999; 22: 517-34.

Akiskal H.S., Akiskal K.K., Haykal R.F., Manning J.S., Connor P.D. (2005). TEMPS-A: progress towards validation of a self-rated clinical version of the Temperament Evaluation of the Memphis, Pisa, Paris, and San Diego Autoquestionnaire. *J Affect Disord*. 85(1-2):3-16. doi: 10.1016/j.jad.2004.12.001.

American Psychiatric Association. *Diagnostic and Statistical Manual of Mental Disorders*. DSM-IV, Washington, DC: APA, 1994.

American Psychiatric Association. *Diagnostic and Statistical Manual of Mental Disorders*. DSM-IV-TR Washington, DC: APA, 2000.

American Psychiatric Association. *Diagnostic and Statistical Manual of Mental Disorders*. DSM-5, Washington, DC: APA, 2013.

American Psychiatric Association. *Manuale Diagnostico e Statistico dei Disturbi Mentali*. DSM-5, Milano: Raffaello Cortina, 2014.

American Psychiatric Association. (2022a). Bipolar I and bipolar II disorders. <https://www.psychiatry.org/getmedia/98fd2c17-93f0-42cd-9f41-755d77b862a5/APA-DSM5TR-BipolarIandBipolarIIDisorders.pdf>

American Psychiatric Association (2022b). *Diagnostic and Statistical Manual of Mental Disorders*, DSM-5-TR, Washington, DC: APA, 2022.

American Psychiatric Association. (2022c). Unspecified Mood Disorder. <https://www.psychiatry.org/getmedia/4d6aab2f-87cf-4b22-afc4-8b9035076051/APA-DSM5TR-UnspecifiedMoodDisorder.pdf>

American Psychiatric Association. (2022d). Other Specified Depressive Disorder. <https://www.psychiatry.org/getmedia/9c456082-b1ed-4a7c-84f7-7b340fa10b01/APA-DSM5TR-OtherSpecifiedDepressiveDisorder.pdf>

American Psychiatric Association. (2022e). Other Specified Bipolar and Related Disorder, <https://www.psychiatry.org/getmedia/ffc643b7-9382-4a29-a8c8-868dabcced10/APA-DSM5TR-OtherSpecifiedBipolarandRelatedDisorder.pdf>

American Psychiatric Association. *Manuale Diagnostico e Statistico dei Disturbi Mentali*. DSM-5-TR, Milano: Raffaello Cortina, 2023.

Andreasen N.C., Flaum M., Arndt S. (1992). The Comprehensive Assessment of Symptoms and History (CASH). An instrument for assessing diagnosis and psychopathology. *Arch Gen Psychiatry*. 49(8):615-23. doi: 10.1001/archpsyc.1992.01820080023004.

Andrews G., Peters L., Guzman A.M., Bird K. (1995). A comparison of two structured diagnostic interviews: CIDI and SCAN. *Aust N Z J Psychiatry*. 29(1):124-32. doi: 10.3109/00048679509075901.

Angst J., *Zur Ätiologie und Nosologie endogener depressiver Psychosen. Eine genetische soziologische und klinische Studie*. Berlin, Heidelberg, New York: Springer, 1996.

Angst J., & Perris C. (1972). The nosology of endogenous depression: Comparison of the Results of Two Studies. *International Journal of Mental Health*, 1(1/2), 145–158. <http://www.jstor.org/stable/41343910>.

Angst, J., Adolfsson, R., Benazzi, F., Gamma, A., Hantouche, E., Meyer, T.D., Skeppar, P., Vieta, E., Scott, J. (2005). The HCL-32: towards a self-assessment tool for hypomanic symptoms in outpatients. *J. Affect. Disord.* 88 (2), 217–233.

Avrichir B.S., Elkis H. (2002). Prevalence and underrecognition of dysthymia among psychiatric outpatients in São Paulo, Brazil. *J Affect Disord.* 2002 May;69(1-3):193-9. doi: 10.1016/s0165-0327(01)00342-1.

Benvenuti A., Miniati M., Callari A., Fagiolini A., Dell'Osso L. (2014). Il concetto di spettro dell'umore e il Mood Spectrum Questionnaire. *Noos*; 20(2):79-86. doi 10.1722/1569.17073.

Benvenuti A., Miniati M., Callari A., Giorgi Mariani M., Mauri M., Dell'Osso L. (2015). Mood Spectrum Model: Evidence reconsidered in the light of DSM-5. *World J Psychiatry*; 5(1): 126-137; <http://www.wjgnet.com/2220-3206/full/v5/i1/126.htm> DOI: <http://dx.doi.org/10.5498/wjp.v5.i1.126>.

Bioletto F., Berton A. M., Principe N., Parasiliti-Caprino M., Gasco V. *et al.* (2022). Strumenti per la lettura critica di metanalisi e revisioni sistematiche: il punto di vista del clinico. *L'Endocrinologo*; 23:620–626; <https://doi.org/10.1007/s40619-022-01185-3>.

Bradley, L., Noble, N., & Hendricks, B. (2023). DSM-5-TR: Salient Changes. *The Family Journal*, 31(1), 5-10. <https://doi.org/10.1177/10664807221123558>.

Bruchmüller, K., Meyer, T.D. (2009). Diagnostically irrelevant information can affect the likelihood of a diagnosis of bipolar disorder. *J. Affect. Disord.* 116, 148–151.

Bobo W.V., Chen H., Trivedi M.H., Stewart J.W., Nierenberg A.A., *et al.* (2011). Randomized comparison of selective serotonin reuptake inhibitor (escitalopram) monotherapy and antidepressant combination pharmacotherapy for major depressive disorder with melancholic features: a CO-MED report. *J. Affect. Disord.* 133(3), 467–476.

Cassano G.B., Frank E., Miniati M., Rucci P., Fagiolini A., *et al.* (2002). Conceptual underpinnings and empirical support for the mood spectrum. *Psychiatr Clin North Am*; 25: 699-712.

Cassano G.B., Rucci P., Frank E., Fagiolini A., Dell’Osso L. *et al.* (2004). The mood spectrum in unipolar and bipolar disorder: arguments for a unitary approach. *Am J Psychiatry*; 161: 1264-1269.

Cheniaux E., Landeira-Fernandez J., Versiani M. (2009). The diagnoses of schizophrenia, schizoaffective disorder, bipolar disorder and unipolar depression: interrater reliability and congruence between DSM-IV and ICD-10. *Psychopathology.* 42(5):293-8. doi: 10.1159/000228838.

Chiorri C. *Fondamenti di psicometria*. Milano: Mc Graw Hill, 2020.

Chmielewski M., Clark L.A., Bagby R.M., Watson D. (2015). Method matters: Understanding diagnostic reliability in DSM-IV and DSM-5. *J Abnorm Psychol.* 124(3):764-9. doi: 10.1037/abn0000069.

Clarke D.E., Narrow W.E., Regier D.A., Kuramoto S.J., Kupfer D.J., Kuhl E.A., *et al.* (2013). DSM-5 field trials in the United States and Canada, Part I: study design, sampling strategy, implementation, and analytic approaches. *Am J Psychiatry.* 2013; 170(1): 43-58. doi: <http://dx.doi.org/10.1176/appi.ajp.2012.12070998>.

Colman A.M., *A Dictionary of Psychology*, 4 ed., Oxford: Oxford University Press, 2015.

Crumlish N., Kelly B.D. (2009). How psychiatrists think. *Advances in Psychiatric Treatment*; 15(1):72-79. doi:10.1192/apt.bp.107.005298.

Dahl A.A., Krüger M.B., Dahl N.H., Karlsson H., Von Knorring L., Stordal E. (2009). SPIFA-A presentation of the Structured Psychiatric Interview for General Practice. *Nord J Psychiatry*. 63(6):443-53. doi: 10.3109/08039480902874769.

Dagani J., Signorini G., Nielsen O., Bani M., Pastore A, Girolamo G., Large M. (2017). Meta-analysis of the Interval between the Onset and Management of Bipolar Disorder. *Can J Psychiatry*. 62(4):247-258. doi: 10.1177/0706743716656607.

Day C.V., Williams L.M. (2012). Finding a biosignature for melancholic depression. *Expert Rev Neurother*. 2012 Jul;12(7):835-47. doi: 10.1586/ern.12.72.

Dell’Aglia J.C. Jr., Basso L.A., Argimon I.I., Arteché A. (2013). Systematic review of the prevalence of bipolar disorder and bipolar spectrum disorders in population based studies. *Trends Psychiatry Psychother*. 35:99–105. doi: 10.1590/S2237-60892013000200002.

Dickersin K., Min Y.I., Meinert C.L. (1992). Factors influencing publication of research results. Follow-up of applications submitted to two institutional review boards. *JAMA* 267:374–378.

Dijksman I., Dinant G.J., Spigt M. (2016). The concurrent validity of a new eDiagnostic system for mental disorders in primary care. *Fam Pract*. 33(6):607-616. doi: 10.1093/fampra/cmw069.

Di Paolo, L., (2008). *Caratterizzazione neuropsicologica dei disturbi dell’umore*, Tesi di Laurea specialistica in Medicina e Chirurgia, Pisa: Università di Pisa.

Dunner D.L., Fleiss J.L., Fieve R.R. (1976). The course of development of mania in patients with recurrent depression. *Am J Psychiatry*. 133:905–8.

El-Den S., Chen T.F., Gan Y.L., Wong E., O’Reilly CL. (2018). The psychometric properties of depression screening tools in primary healthcare settings: A systematic review. *J Affect Disord*. 2018 Jan 1; 225:503-522.

Fagiolini A., Dell’Osso L., Pini S., Armani A., Bouanani S., Rucci P. *et al.* (1999). Validity and reliability of a new instrument for assessing mood symptomatology: the Structured Clinical Interview for Mood Spectrum (SCI MOODS). *Int J Meth Psych Res*; 8: 71-81.

Fieve R.R., Dunner D.L. (1975). *Unipolar and bipolar affective states*. In Flach FF, Draghi SS (eds): *The Nature and Treatment of Depression*. New York, John Wiley & Sons, pp. 145-160.

First M.B. (2016). The importance of developmental field trials in the revision of psychiatric classifications. *Lancet Psychiatry*. 2016 Jun;3(6):579-84. doi: 10.1016/S2215-0366(16)00048-1.

First M. B., Yousif L. H., Clarke D. E., Wang P. S., Gogtay N., Appelbaum P. S., (2022). DSM-5-TR: overview of what's new and what's changed. *World Psychiatry*.;21(2):218-219. doi: 10.1002/wps.20989.

Frattura L., Bassi G., Nardo E., Tonel P., Zavaroni C. (2022). *Aggiornamento 2019 della Classificazione statistica internazionale delle malattie e dei problemi sanitari correlati – Decima revisione (ICD-10), quinta edizione, 2016 – Volume 1 Elenco sistematico*. Trieste: Azienda Sanitaria Universitaria Giuliano Isontina, 2022.

Gaebel W., Stricker J., Kerst A. (2020). Changes from ICD-10 to ICD-11 and future directions in psychiatric classification^[SEP]. *Dialogues Clin Neurosci*. 2020 Mar;22(1):7-15. doi: 10.31887/DCNS.2020.22.1/wgaebel.

Goes F.S. (2023). Diagnosis and management of bipolar disorders. *BMJ*. 12:381; 1-17. Doi: 10.1136/bmj-2022-073591.

Gutkovich Z. (2013). Initial validation of the Russian version of the World Mental Health Structured Clinical Interview for DSM-IV. *Isr J Psychiatry Relat Sci*; 50(1):24-32.

Harnic D., Koukopoulos A., Mazza, M., Janiri, L., Bria, P. (2011). Il temperamento come elemento fondante dello spettro bipolare delle manifestazioni sottosoglia: una breve rassegna della letteratura psichiatrica, *Giorn Ital Psicopat*, 17, 213-224.

Hirschfeld, R.M.A. (2001). Bipolar spectrum disorder: improving its recognition and diagnosis. *J. Clin. Psychiatry* 62 (suppl. 14), 5–9.

Hirschfeld, R.M.A., Lewis, L., Vornik, L.A. (2003). 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*; 64, 161–174.

Hirschfeld, R.M.A., Vornik, L.A., (2004). Recognition and diagnosis of bipolar disorder. *J. Clin. Psychiatry* 65, 5–9.

Hirschfeld R.M.A. (2014). Differential diagnosis of bipolar disorder and major depressive disorder. *J Affect Disord.* 2014 Dec;169 Suppl 1:S12-6. doi: 10.1016/S0165-0327(14)70004-7.

Jervolino D., *Ricoeur: la fenomenologia della memoria*. In M. Cappuccio (a cura di), Neurofenomenologia. Le scienze della mente e la sfida dell'esperienza cosciente, Bruno Mondadori, Milano, 2006, pp. 397-409.

Joo E.J., Joo Y.H., Hong J.P., Hwang S., Maeng S.J., *et al.* (2004). Korean version of the diagnostic interview for genetic studies: Validity and reliability. *Compr Psychiatry*. 45(3):225-9. doi: 10.1016/j.comppsy.2004.02.007.

Kaltenboeck A., Winkler D., Kasper S. (2016). Bipolar and related disorders in DSM-5 and ICD-10. *CNS Spectr.*, 21(4):318-23. doi: 10.1017/S1092852916000079.

Kendler KS. (2020a). The origin of our modern concept of mania in texts from 1780 to 1900. *Mol Psychiatry*. 2020 Sep;25(9):1975-1985. doi: 10.1038/s41380-020-0657-0.

Kendler KS. (2020b). The Origin of Our Modern Concept of Depression-The History of Melancholia From 1780-1880: A Review. *JAMA Psychiatry*. 2020 Aug 1;77(8):863-868. doi: 10.1001/jamapsychiatry.2019.4709.

Kennis M., Gerritsen L., van Dalen M., Williams A., Cuijpers P., Bockting C. (2020). Prospective biomarkers of major depressive disorder: a systematic review and meta-analysis. *Mol Psychiatry*; 25(2):321-338. doi: 10.1038/s41380-019-0585-z.

Kogan C.S., Maj M., Rebello T.J., Keeley J.W., Kulygina M., Matsumoto C. *et al.* (2021). A global field study of the international classification of diseases (ICD-11) mood disorders

clinical descriptions and diagnostic guidelines. *J Affect Disord.* 2021 Dec 1;295:1138-1150. doi: 10.1016/j.jad.2021.08.050.

Kortmann, F. (1988). DSM-III in Ethiopia: A feasibility study. *Eur Arch Psychiatr Neurol Sci* 237, 101–105. <https://doi.org/10.1007/BF00382373>.

Krawczyk P., Świącicki Ł. (2020). ICD-11 vs. ICD-10 - a review of updates and novelties introduced in the latest version of the WHO International Classification of Diseases. *Psychiatr Pol.*;54(1):7-20. Doi: 10.12740/PP/103876.

Kraepelin E. *Lectures on Clinical Psychiatry*. Terza edizione. New York: W Wood 1905.

Kraepelin E., *Manic-depressive insanity and paranoia*. Traduzione inglese di M. Barclay, Edimburgo: E & S. Livingstone, 1921.

Kralj Ž., Dedić M., Kovačević A., Malički M., Dedić J., (2017). Diagnostic Interview for Genetic Studies: validity and reliability of the Croatian version. *Psychiatr Genet.* 27(1):17-22. doi: 10.1097/YPG.0000000000000154.

Landis J.R., Koch G.G. (1977). The measurement of observer agreement for categorical data. *Biometrics.* 1977 Mar;33(1):159-74.

Lauterbach E., Rumpf H.J., Ahrens B., Haug H.J., Schaub R., *et al.* (2005). Assessing dimensional and categorical aspects of depression: validation of the AMDP Depression Scale. *Eur Arch Psychiatry Clin Neurosci.* 2005 Feb;255(1):15-9. doi: 10.1007/s00406-004-0529-7.

Liu X., Jiang K. (2016). Why is Diagnosing MDD Challenging? *Shanghai Arch Psychiatry.* 2016 Dec 25;28(6):343-345. doi: 10.11919/j.issn.1002-0829.216073.

Lobbestael J., Leurgans M., Arntz A. (2011). Inter-rater reliability of the Structured Clinical Interview for DSM-IV Axis I Disorders (SCID I) and Axis II Disorders (SCID II). *Clin Psychol Psychother.* 18(1):75-9. doi: 10.1002/cpp.693.

Maina G., Forner F., Rosso G., Salvi V., Bogetto F. (2003). Caratteristiche temperamentalì in un campione di 80 pazienti depressi. *Giorn Ital Psicopat*; 9:157-62.

Maj M., Pirozzi R., Formicola A.M., Bartoli L., Bucci P. (2000). Reliability and validity of the DSM-IV diagnostic category of schizoaffective disorder: preliminary data. *J Affect Disord*. 57(1-3):95-8. doi: 10.1016/s0165-0327(99)00059-2.

Malhotra S., Chakrabarti S., Shah R., Sharma M., Sharma K., Singh H. (2015). Diagnostic accuracy and feasibility of a net-based application for diagnosing common psychiatric disorders. *Psychiatry Res*. 15;230(2):369-76. doi: 10.1016/j.psychres.2015.09.020.

Mantere O., Suominen K., Leppämäki S., Valtonen H., Arvilommi P., Isometsä E. (2004). The clinical characteristics of DSM-IV bipolar I and II disorders: baseline findings from the Jorvi Bipolar Study (JoBS). *Bipolar Disorder*, 6:395-405.

Marneros A., Angst, J., *Bipolar disorders: roots and evolution*, in Bipolar disorders: 100 years after manic-depressive insanity, a cura di Marneros A. e Angst J., Kluwer Academic Publishers, 2002.

McHugh M.L. (2012). Interrater reliability: the kappa statistic. *Biochem Med (Zagreb)*. 2012; 22(3): 276-82.

McIntyre R.S., Zimmerman M., Goldberg J.F., First M.B. (2019). Differential Diagnosis of Major Depressive Disorder Versus Bipolar Disorder: Current Status and Best Clinical Practices. *J Clin Psychiatry*. 2019 May 14;80(3):ot18043ah2. doi: 10.4088/JCP.ot18043ah2.

McIntyre R.S., Berk M., Brietzke E., Goldstein B.I., López-Jaramillo C., Kessing L.V., *et al.* (2020). Bipolar disorders. *Lancet*. 2020 Dec 5;396(10265):1841-1856. doi: 10.1016/S0140-6736(20)31544-0.

Medina-Mora M.E., Robles R., Rebello T.J., Domínguez T., Martínez N., *et al.* (2019). ICD-11 guidelines for psychotic, mood, anxiety and stress-related disorders in Mexico: Clinical utility and reliability. *Int J Clin Health Psychol*. 19(1):1-11. doi: 10.1016/j.ijchp.2018.09.003.

Meyer, T.D., Koßmann-Böhm, S., Schlottke, P.F. (2004). Do child psychiatrists in Germany diagnose bipolar disorders in children and adolescents? Results from a survey. *Bipolar Disord.* 6, 426–431.

Miklowitz D., *Il disturbo bipolare: una guida per la sopravvivenza*. Roma: Giovanni Fioriti, 2005.

Nagar M., Westen D., Nakash O. (2018). Reliability of DSM and empirically derived prototype diagnosis for mood, anxiety and personality disorders. *Compr Psychiatry.* 85:8-14. doi: 10.1016/j.comppsy.2018.06.001.

Okasha A., Sadek A., al-Haddad M.K., Abdel-Mawgoud M. (1993). Diagnostic agreement in psychiatry. A comparative study between ICD-9, ICD-10 and DSM-III-R. *Br J Psychiatry.* 162:621-6. doi: 10.1192/bjp.162.5.621.

Onofa L., Odunleye M., Kola L., Gureje O. (2019). Reliability and Clinical Utility of ICD-11 Diagnostic Guidelines for Severe Mental Disorders in Nigeria. *Arch Med Res.* 50(8):535-542. doi: 10.1016/j.arcmed.2020.01.004.

Osório F.L., Loureiro S.R., Hallak J.E.C., Machado-de-Sousa J.P., Ushirohira J.M. *et al.* (2019). Clinical validity and intrarater and test-retest reliability of the Structured Clinical Interview for DSM-5 - Clinician Version (SCID-5-CV). *Psychiatry Clin Neurosci.* 2019 Dec;73(12):754-760. doi: 10.1111/pcn.12931.

Pacchiarotti I., Bond D.J., Baldessarini R.J., Nolen W.A., Grunze H., Licht R.W. *et al.* (2013). The International Society for Bipolar Disorders (ISBD) task force report on antidepressant use in bipolar disorders. *Am J Psychiatry.* 2013 Nov;170(11):1249-62. doi: 10.1176/appi.ajp.2013.13020185.

Pang A., Ungvari G., Wong C., Leung T. (1997). Inter-rater reliability of ICD-10 in Chinese patients. *Eur Psychiatry.* 12(1):8-10. doi: 10.1016/S0924-9338(97)86372-9.

Parker G., McCraw S., Blanch B., Hadzi-Pavlovic D., Synnott H., Rees A-M. (2013a). Discriminating melancholic and non-melancholic depression by prototypic clinical features. *J Affect Disord;* 2013;144(3):199-207.

Parker G., McCraw S., Fletcher K., Friend P., Futeran S. Validation of a new prototypic measure of melancholia (2013b). *Compr Psychiatry*. 2013 Oct;54(7):835-41. doi: 10.1016/j.comppsy.2013.02.010.

Parker G., Tavella G., Hadzi-Pavlovic D. (2019). Identifying and differentiating melancholic depression in a non-clinical sample. *J Affect Disord*. 2019 Jan 15;243:194-200. doi: 10.1016/j.jad.2018.09.024.

Placidi G.F., Marenmani I., Signoretta S., *et al.* (1998). A prospective study of stability and change over 2 years of affective temperaments in 14-18-year-old Italian high school students. *J Affect Disord* 1998; 51:199-208.

Perris, C., A., (1966). Study of bipolar (manic-depressive) and unipolar recurrent depressive psychoses. *Acta Psychiatr Scand*. 194(Suppl.), 1–89.

Perugi, G., Tusini, G., Nassimbeni, A., Marenmani I., Akiskal, H. S. (2003). Il disturbo bipolare II: spettro clinico e sindromi correlate, *Journal of Psychopathology*, v. 3, <https://old.jpsychopathol.it/article/il-disturbo-bipolare-ii-spettro-clinico-e-sindromi-correlate/>

Preisig M., Fenton B.T., Matthey M.L., Berney A., Ferrero F. (1999). Diagnostic interview for genetic studies (DIGS): inter-rater and test-retest reliability of the French version. *Eur Arch Psychiatry Clin Neurosci*. 249(4):174-9. doi: 10.1007/s004060050084.

Price, L. R., *Psychometric Methods: Theory into Practice*. New York, London: The Guildford Press, 2017.

Ptok U., Seeher C., Jessen F., Papassotiropoulos A., Heun R. (2001). Inter-rater reliability of family history information on psychiatric disorders in relatives. *Eur Arch Psychiatry Clin Neurosci*. 2001 Dec;251(6):279-83. doi: 10.1007/pl00007546.

Quintana M.I., Andreoli S.B., Jorge M.R., Gastal F.L., Miranda C.T. (2004). The reliability of the Brazilian version of the Composite International Diagnostic Interview (CIDI 2.1). *Braz J Med Biol Res*. 37(11):1739-45. doi: 10.1590/s0100-879x2004001100020.

Ransing R., Deshpande S.N., Shete S.R., Patil I., Kukreti P., *et al.* (2020). Assessing antenatal depression in primary care with the PHQ-2 and PHQ-9: Can It Be carried Out by auxiliary nurse midwife (ANM)? *Asian J Psychiatr.* 2020 Oct; 53:102109. doi: 10.1016/j.ajp.2020.102109.

Reed G.M., Sharan P., Rebello T.J., Keeley J.W., Medina-Mora M. *et al.* (2018). The ICD-11 developmental field study of reliability of diagnoses of high-burden mental disorders: results among adult patients in mental health settings of 13 countries. *World Psychiatry.* 17(2):174-186. doi: 10.1002/wps.20524.

Reed G.M., First M.B., Kogan C.S., Hyman S.E., Gureje O., Gaebel W., Maj M., Stein D.J., *et al.* (2019). Innovations and changes in the ICD-11 classification of mental, behavioural and neurodevelopmental disorders. *World Psychiatry.*,18(1):3-19. doi: 10.1002/wps.20611.

Regier D.A., Kaelber C.T., Roper M.T., Rae D.S., Sartorius N. (1994a e 1994b). The ICD-10 clinical field trial for mental and behavioral disorders: results in Canada and the United States. *Am J Psychiatry.* 2019; 151(9):1340-50. doi: 10.1176/ajp.151.9.1340.

Regier D.A., Narrow W.E., Clarke D.E., Kraemer H.C., Kuramoto S.J., Kuhl E.A., Kupfer D.J. (2013). DSM-5 field trials in the United States and Canada, Part II: test-retest reliability of selected categorical diagnoses. *Am J Psychiatry.* 170(1):59-70. doi: 10.1176/appi.ajp.2012.12070999.

Ricoeur P., *Dal testo all'azione. Saggi di ermeneutica*, trad. it. di G. Grampa, Jaca Book, Milano 2003.

Robins L.N., Cottler L.B. (2004). Making a structured psychiatric diagnostic interview faithful to the nomenclature. *Am J Epidemiol.* 2004 Oct 15;160(8):808-13. doi: 10.1093/aje/kwh283.

Roca M., Martin-Santos R., Saiz J., Obiols J., Serrano M.J., *et al.* (2006). Diagnostic Interview for Genetic Studies (DIGS): inter-rater and test-retest reliability and validity in a Spanish population. *Eur Psychiatry.*22(1):44-8. doi: 10.1016/j.eurpsy.2006.10.004.

Roy M.A., Lanctôt G., Mérette C., Cliche D., Fournier J.P., Boutin P. *et al.* (1997). Clinical and methodological factors related to reliability of the best-estimate diagnostic procedure. *Am J Psychiatry.*154(12):1726-33. doi: 10.1176/ajp.154.12.1726.

Rocha Neto, H.; Moreira, A.L.R.; Hosken, L.; Langfus, J.A.; Cavalcanti, M.T. *et al.* (2023a). Inter-Rater Reliability between Structured and Non-Structured Interviews Is Fair in Schizophrenia and Bipolar Disorders—A Systematic Review and Meta-Analysis. *Diagnostics*, 13, 526. <https://doi.org/10.3390/diagnostics13030526>.

Rocha Neto H.G., Lessa J.L.M., Koiller L.M., Pereira A.M., de Souza Gomes B.M., *et al.* (2023b). Non-standard diagnostic assessment reliability in psychiatry: a study in a Brazilian outpatient setting using Kappa. *Eur Arch Psychiatry Clin Neurosci*. 2023 Dec 12. doi: 10.1007/s00406-023-01730-7.

Rucci P., Maser J.D. (2000). Instrument development in the Italy-USA Collaborative Spectrum Project. *Epidemiol Psychiatr Soc* 2000; 9: 249-256.

Rucci P., Calugi, S., Miniati M., Fagiolini A. (2013). A review of self-report and interview-based instruments to assess mania and hypomania symptoms, *Journal of Psychopathology* 2013;19:1-18.

Sartorius N., Kaelber C.T., Cooper J.E., Roper M.T., Rae D.S. *et al.* (1993). Progress toward achieving a common language in psychiatry. Results from the field trial of the clinical guidelines accompanying the WHO classification of mental and behavioral disorders in ICD-10. *Arch Gen Psychiatry*. 50(2):115-24. doi: 10.1001/archpsyc.1993.01820140037004.

Santelmann H., Franklin J., Bußhoff J., Baethge C. (2016). Interrater reliability of schizoaffective disorder compared with schizophrenia, bipolar disorder, and unipolar depression - A systematic review and meta-analysis. *Schizophr Res*.176(2-3):357-363. doi: 10.1016/j.schres.2016.07.012.

Shabani A., Masoumian S., Zamirinejad S., Hejri M., Pirmorad T., Yaghmaeezadeh H. (2021). Psychometric properties of Structured Clinical Interview for DSM-5 Disorders-Clinician Version (SCID-5-CV). *Brain Behav*. 11(5):e01894. doi: 10.1002/brb3.1894.

Sharma V.K., Lepping P., Cummins A.G., Copeland J.R., Parhee R., Mottram P. (2004). The Global Mental Health Assessment Tool--Primary Care Version (GMHAT/PC). Development, reliability and validity. *World Psychiatry*. 3(2):115-9.

Siciliano A. (2015). Recensione a Hubertus Tellenbach, *Melancholia. Storia del problema, endogenicità, tipologia, patogenesi, clinica*, *Lo Sguardo. Rivista di filosofia* n. 18, v. 2.

Smith-Nielsen J., Matthey S., Lange T., Væver M.S. (2018). Validation of the Edinburgh Postnatal Depression Scale against both DSM-5 and ICD-10 diagnostic criteria for depression. *BMC Psychiatry*. 20;18(1):393. doi: 10.1186/s12888-018-1965-7.

Spoelma M.J., Serafimovska A., Parker G. (2023). Differentiating melancholic and non-melancholic depression via biological markers: A review. *World J Biol Psychiatry*. 24(9):761-810. doi: 10.1080/15622975.2023.2219725.

Stensland M.D., Schultz J.F., Frytak J.R. (2010). Depression diagnoses following the identification of bipolar disorder: costly incongruent diagnoses. *BMC Psychiatry*. 4; 10:39. doi: 10.1186/1471-244X-10-39.

Sundqvist N., Garrick T., Bishop I., Harper C. (2008). Reliability of post-mortem psychiatric diagnosis for neuroscience research. *Aust N Z J Psychiatry*. 2008 Mar;42(3):221-7. doi: 10.1080/00048670701827242.

Suppes T., Leverich G.S., Keck P.E., Nolen W.A., Denicoff K.D., Altshuler L.L. *et al.* (2001). The Stanley Foundation Bipolar Treatment Outcome Network. II. Demographics and illness characteristics of the first 261 patients. *J Affect Disord*, 67:45-59.

Tellenbach, H. *Melancholia. Storia del problema, Endogenicità, tipologia, patogenesi, clinica*. Tr. it. di E. Cipollini, Roma: Il pensiero scientifico, 1975.

Urban E.J., Charles S.T., Levine L.J., Almeida D.M. (2018). Depression history and memory bias for specific daily emotions. *PLoS One*. 13(9); doi: 10.1371/journal.pone.0203574.

Vega-Núñez A., Gómez-Sánchez-Lafuente C., Mayoral-Cleries F., Bordallo A., Rodríguez de Fonseca F., Suárez J., Guzmán-Parra (2022). J. Clinical Value of Inflammatory and Neurotrophic Biomarkers in Bipolar Disorder: A Systematic Review and Meta-Analysis. *Biomedicines*. 10(6):1368. doi: 10.3390/biomedicines10061368.

Wang Y-Y., Xu D.D., Liu R., Yang Y., Grover S., Ungvari G.S., Hall B.J. *et al.* (2019). Comparison of the screening ability between the 32-item Hypomania Checklist (HCL-32)

and the Mood Disorder Questionnaire (MDQ) for bipolar disorder: a meta-analysis and systematic review. *Psychiatry Res.* 273: 461–66.

Wittchen H.U., Robins L.N., Cottler L.B., Sartorius N, Burke JD, Regier D. (1991). Cross-cultural feasibility, reliability and sources of variance of the Composite International Diagnostic Interview (CIDI). The Multicentre WHO/ADAMHA Field Trials. *Br J Psychiatry.*;159:645-53, 658. doi: 10.1192/bjp.159.5.645.

Woody C.A., Ferrari A.J., Siskind D.J., Whiteford H.A., Harris M.G. (2017). A systematic review and meta-regression of the prevalence and incidence of perinatal depression. *J Affect Disord*; 219:86-92. doi: 10.1016/j.jad.2017.05.003.

World Health Organization (1993). *The ICD-10 Classification of Mental and Behavioural Disorders: Diagnostic Criteria for Research*. Geneva: WHO; 1993.

Zanarini M.C., Skodol A.E., Bender D., Dolan R., Sanislow C. *et al.* (2000). The Collaborative Longitudinal Personality Disorders Study: reliability of axis I and II diagnoses. *J Pers Disord.* 14(4):291-9. doi: 10.1521/pedi.2000.14.4.291.

Zimmerman M., Galione J.N. (2011). Screening for bipolar disorder with the Mood Disorders Questionnaire: a review. *Harv Rev Psychiatry*; 19:219-28.

SITOGRAFIA

Institute of Health Metrics and Evaluation (IHME). Global Health Data Exchange (GHDx). <https://vizhub.healthdata.org/gbd-results/> consultato in data 02/07/2024.

Istituto di Psicopatologia, a cura di Tundo A., “Breve storia del disturbo bipolare”: <https://www.istitutodipsicopatologia.it/breve-storia-del-disturbo-bipolare/> consultato in data 4/07/2024.

Istituto Italiano di Statistica (ISTAT). Comunicato stampa “Salute mentale” del 26.07.2018: <https://www.istat.it/comunicato-stampa/la-salute-mentale-nelle-varie-fasi-della-vita/> consultato in data 30/08/2024.

Spectrum Project Collaborative Group: <https://www.spectrum-project.org/> consultato in data 15/07/2024.

Spectrum Project Collaborative Group. Questionari MOODS-SR e intervista strutturata SCI-MOODS: <https://www.spectrum-project.org/questionnaires/moods.html> consultato in data 16/07/2024.

Treccani. Vocabolario 2022 online. Umore: <https://www.treccani.it/vocabolario/umore/?search=um%C3%B3re%2F> consultato in data 3/07/2024.

World health Organization (WHO). ICD-11 for Mortality and Morbidity Statistics: <https://icd.who.int/browse/2024-01/mms/en> consultato in data 8/07/2024.

World health Organization (WHO). ICD-11 for Mortality and Morbidity Statistics. 6A61.9 Bipolar type II disorder, currently in partial remission, most recent episode unspecified. Rapid cycling: <https://icd.who.int/browse/2024-01/mms/en#2021041320> consultato in data 11/07/2024.

World health Organization (WHO). ICD-11 for Mortality and Morbidity Statistics. Single episode depressive disorder, unspecified: <https://icd.who.int/browse/2024-01/mms/en#578635574%2FUnspecified> consultato in data 10/07/2024.