EUROPEAN MEDICINES AGENCY
SCIENCE MEDICINES HEALTH

25 July 2024
EMA/CHMP/406037/2024 rev. 1
Committee for Medicinal Products for Human Use (CHMP)

Guideline on clinical development of medicinal products for the treatment and prevention of bipolar disorder

Draft

Draft Agreed by Central Nervous System Working Party	June 2024
Adopted by CHMP for release for consultation	25 July 2024
Start of public consultation	9 September 2024
End of consultation (deadline for comments)	31 March 2025

This guideline replaces Guideline on clinical investigation of medicinal products in the treatment and prevention of bipolar disorder (CPMP/EWP/567/98).

Comments should be provided using this EUSurvey form. For any technical issues, please contact the EUSurvey Support.

Keywords	***Bipolar disorder, mania, bipolar depression, acute treatment, maintenance treatment, recurrence prevention***

Official address Domenico Scarlattilaan 6 ● 1083 HS Amsterdam ● The Netherlands
Address for visits and deliveries Refer to www.ema.europa.eu/how-to-find-us
Send us a question Go to www.ema.europa.eu/contact **Telephone** +31 (0)88 781 6000 An agency of the European Union

Guideline on clinical development of medicinal products for the treatment and prevention of bipolar disorder

Table of contents

Definitions

Maintenance of effect: the effect of treatment, seen in the short-term studies is maintained during the whole (hypo) manic/depressive episode. In the literature this treatment phase sometimes is called the continuation period.

Relapse: an increase in symptomatology during the current (index) episode.

Recurrence: a re-emergence of depressive or manic symptoms after a time with no or minimal symptoms and without medication. It is seen as the start of a new episode.

Response: clinically relevant improvement in terms of symptom reduction (to be defined in advance of the study).

Responder: a patient with clinically relevant and pre-defined improvement.

Remission: no or only few signs of a specified condition (e.g., depression) remaining, and at the same time absence of a concomitant increase in symptoms of another condition (e.g., mania or hypomania).

Recovery: a sustained period of remission (usually 2 months as a minimum) representing resolution of the index episode but not of the bipolar disorder itself.

Specifier: extension to a diagnosis that further clarifies the course, severity, or special features of the patient's disorder. More than one specifier may be applied on a patient.

Executive summary

The main aim of the guideline is to provide up to date guidance for the development of medicinal products for the treatment and prevention of Bipolar disorder (BD). This is the first revision of the Note for Guidance on clinical investigation of medicinal products in the treatment and prevention of bipolar disorder issued in 2001.

With the transition of Diagnostic and Statistical Manual of Mental Disorders (DSM) IV into DSM- 5, bipolar and related disorders have been separated from depressive disorders, and BD II is no longer considered a milder form of BD I. Criterion A for manic and hypomanic episodes has changed to emphasise changes in activity and energy as well as mood and new specifiers have been added (e.g. anxious distress and mixed features). This has subsequent implications for identifying patients both in clinical and research settings.

The key requirements for the development of medicinal products for the treatment of bipolar disorder in terms of study design, patient population and outcome measures are described. Specific issues, including children and adolescents are addressed. The importance of specific domains in bipolar disorder including cognitive and anxious/distress symptoms as well as treatment resistance and partial response are discussed and aligned with studies in other psychiatric disorders, i.e. major depressive disorder. Prevention of relapse and recurrent episodes is addressed as well.

This document should be read in conjunction with other relevant EMA and ICH guidelines (see section 3).

1. Introduction (background)

Bipolar Disorder (BD) I and II are chronic psychiatric disorders characterised by recurrent episodes of manic and depressive states with intervening episodes of normal mood. Over the past decades their life-time prevalence remained stable (BD I 0.6% and BD II 0.4%, worldwide). Approximately 40 million people in the world have bipolar disorder (World Health Organisation, 2019). All-cause mortality is increased in bipolar disorder, and it is estimated that the risk of suicide is 10-30 times greater for individuals affected by bipolar disorder relative to the general population. Aetiology and pathophysiology of BD remain still largely unknown, however, genetic research using Genome Wide Association Studies (GWAS) and Copy Number Variations (CNV) methodologies in extremely large samples have shed some light on the neurobiology. Bipolar Disorder is now considered a bridging disease between schizophrenia spectrum and other psychotic disorders and depressive disorders in terms of symptomatology, family history and genetics (DSM-5, 2013).

BD I is characterized by predominant manic episodes. Hypomanic or major depressive episodes are common in BD I, but they are not required for the diagnosis. In contrast, to receive a diagnosis of BD II, patients must have experienced at least one hypomanic episode and at least one major depressive episode. The presence of a manic episode during the course of illness precludes the diagnosis of BD II.

Many medicinal products are available for the treatment of acute manic episode, and some are also indicated for the treatment of major depressive episodes in bipolar disorder, recurrence prevention or maintenance treatment in bipolar disorder. Requirements for efficacy studies addressing short-term treatment of mood episodes and maintenance of effect are discussed in section 4.3.1.

Development of new medicinal product is encouraged that should not only focus on the treatment of acute symptoms and prevention of relapse during the current i.e. index episode but explore also the potential of a medicinal product in the prevention of new episodes i.e. recurrence prevention. However,

prevention of a new episode is not a mandatory part of a registration package for treatment of episodes of BD but is considered as an additional claim (section 4.4.2).

Targeting specific symptoms or domains within bipolar disorder has become of interest in recent years and is thus briefly addressed in this guideline (see section 4.4 and 4.5). The justified claims through the clinical development program will be reflected in the wording of the indication.

Although, BD primarily affects adults, there is increasing evidence that the disorder often begins in adolescence, and diagnostic accuracy is moving towards early detection and a younger patient population. Information from both prospective studies and studies in the offspring of affected parents show that symptoms of irritability, often associated with Attention Deficit Hyperactivity Disorder (ADHD), predispose to BD and that prodromal mood symptoms are present before the onset of BD and in the offspring of patients with BD. However, these symptoms can also occur prodromally to other psychiatric disorders and are therefore not specific to bipolar disorder. Several behavioural, genetic, neuroimaging, electrophysiological- and (immune) biomarkers hold promises for identification of these patients at risk. However, at present the data are still too heterogeneous and not suitable for 'fingerprinting'.

Studies in children and adolescents with established bipolar disorder are addressed in section 4.6.2.

2. Scope

Guidance is provided on medicinal products developed specifically for BD I and II in adults and adolescents, including targeting specific specifiers or symptom domains. Other related disorders of the bipolar spectrum (e.g. cyclothymic disorder) are beyond the scope of this guideline.

3. Legal basis and relevant guidelines

This guideline has to be read in conjunction with the introduction and general principles (4) and part I and II of the Annex I to Directive 2001/83 (as amended). Further is referred to the EMA and ICH guidelines on pharmaceutical development, PK/PD topics, clinical study design, special populations including the elderly and paediatric population:

https://www.ema.europa.eu/en/human-regulatory/research-development/scientific-guidelines

Among those of specific interest for bipolar disorder are:

- EMA/CHMP/185423/2010, Rev.3 Guideline on clinical investigation of medicinal products in the treatment of depression
- Guideline On The Non-Clinical Investigation Of The Dependence Potential Of Medicinal Products (EMEA/CHMP/SWP/94227/2004)

4. Specific considerations when developing products for the treatment and prevention of bipolar disorder episodes

4.1. Clinical Pharmacology studies

4.1.1. Pharmacodynamics

Bipolar disorder (BD) is characterised by recurrent episodes of manic and depressive states with intervening episodes of normal mood. Due to this heterogenic symptomatology in BD, there is still a lack of adequate animal models for bipolar disorder. Specifically, most animal models mimic only mania or depression and only a few include the cyclical nature of the human condition.

Behavioural models should help understanding the biological mechanisms that contribute to mood cycling, which is the hallmark of BD. The development of such models in BD is encouraged.

The use of potential biomarkers such as receptor occupancy PET studies (e.g. D2 and 5HT 2A) and imaging techniques to detect structural or functional changes in the brain may be useful in dose finding and proof of concept studies in early drug development. For specific claims on efficacy on domains, e.g. cognition, appropriate preclinical studies supporting the mechanism of action and effect shown on clinical outcomes should be provided. Gender aspects should be taken into account and potential differences interpreted for clinical outcome (see section 4.4.3).

As with unipolar depression, novel mechanisms of actions and novel pathways associated with quicker onset of action should be specifically investigated to provide the appropriate support for the clinical efficacy.

4.1.2. Pharmacokinetics

Studies should be performed to characterise the pharmacokinetics of the new medicinal product (see Guideline on Pharmacokinetic studies in man, 3CC3a) and where possible this information should be used to study the relationship between dose, exposure and response. The extent to which the drug crosses the blood-brain barrier should be established. Population PK analyses may be used to investigate relevant covariates e.g. weight, age, sex (gender), healthy vs. patient population, concomitant medications, etc. which could potentially influence the pharmacokinetics of the drug. The choice of dose for the clinical programme should be adequately justified. If appropriate, pharmacokinetic studies in patients with hepatic and /or renal impairment should be performed (see CPMP/EWP/560/95/Rev. 1 Corr. 2**).

4.1.3. Interaction studies

In general, the guideline on drug interactions should be followed to investigate possible pharmacokinetic interactions with other drugs and food. Interactions with alcohol and other relevant CNS active medicinal products should be investigated. It is recommended that pharmacodynamic interactions between the test drug and any other drug that may be prescribed simultaneously in clinical practice are explored and discussed either through dedicated studies or literature data. ICH guideline M12 on drug interaction studies should be followed.

4.2. Assessment of therapeutic efficacy

In order to assess therapeutic efficacy in the context of 'bipolar disorder', the treatment effect on manic or depressive symptoms in the acute phase and in the maintenance treatment period should be demonstrated. Even if a product is only effective in one situation e.g. in the treatment of acute manic episode instead of the entire spectrum of the disorder, including depressive episodes, it is important to know whether it provokes the opposite pole of bipolar disorder (e.g. induces a switch from mania to depression). Therefore, longer term or maintenance data are required in case of mere "acute treatment" claims (see 4.3.1.2). Switching should be carefully monitored and rates reported for all studies in bipolar disorder.

Recurrence prevention is a specific claim and discussed in detail in sections 4.2.3 and 4.4.2.

Results should be discussed in terms of both clinical relevance and statistical significance, and the effect should be shown to be insensitive to the analysis used. When an effect is quantified in terms of change from baseline to end of treatment using a validated measurement tool, this effect has to be

addressed also as rates of responders and remitters. The statistical significance and the clinical relevance of the effect together form the basis for the benefit/risk assessment.

4.2.1. Target of estimation in bipolar disorder

The scientific question(s) of interest, i.e. what the study seeks to address, and consequently the target(s) of estimation (estimand) should be clearly specified. Study planning, design, conduct, analysis, and interpretation must be aligned with the estimand. Reference is made to ICH E9 (R1) addendum on estimands and sensitivity analysis in clinical trials (EMA/CHMP/ICH/436221/2017).

The estimand attributes as well as relevance of intercurrent events depend on the specific context and might differ depending on the condition of the disorder (BD I, BD II), the type of treatment (monotherapy, add-on treatment), and the treatment goal [treatment of acute symptoms in current (index) episode (manic or depressive), maintenance of effect during current episode) and prevention of new episodes with long-term treatment]. Intercurrent events to be considered include, but are not limited to, treatment discontinuation, changes in medication such as use of alternative antidepressants/antipsychotics or other medications and changes in background therapy (mood-stabilizer, psychotherapy, anxiolytic medication, hypnotic medication). In addition, depending on the population selected, death due to committed suicide might require incorporation into the estimand definition.

Generally, and in particular for the treatment of depressive episodes in the context of bipolar depression, reference is made to the section on the target of estimation in the guideline for unipolar depression (EMA/CHMP/185423/2010, Rev. 3). Considerations also apply for bipolar depression, but even more detailed definition of the population attribute of the estimand is required given the variety of different treatment settings. However, an inherent difference to the unipolar setting that warrants further consideration with respect to the estimand, is that patients can switch from one pole of the disorder to the other (i.e. from depression to mania and vice versa).

Intercurrent events not consisting in or following from a switch from one pole of the disorder to the other can be broadly treated in analogy with the guideline on unipolar depression, a treatment-policy strategy is preferred for treatment discontinuation, and changes in background therapy. For switch to alternative treatments a treatment-policy strategy may also be considered, but a hypothetical strategy may be more relevant.

Patients switching from one pole of the disorder to the other represent events of specific interest in studies in bipolar disorder. Both poles of the disorder are undesirable health states, and as such a switch is considered a treatment failure. As the switch might be associated with a "good" score on the main outcome measurement tool(s) of a study it affects the interpretation of the observed main outcome. Hence, it represents an intercurrent event. The undesirable nature of this event should be taken into account in the estimand and, in line with it, a composite strategy defining the switches as treatment failures should be pursued. Implementing a composite strategy without dichotomizing the outcome (and consequent loss of information) is possible but it requires methods and choices that might need to be discussed in the context of CHMP scientific advice. Secondary estimands that include alternative strategies or ways of defining the outcome that otherwise incorporate this intercurrent event might also be of interest.

Overall and regardless of the specific setting, the choice of estimand and the aligned methods of estimators (see section 4.3.3.) are still areas of ongoing discussion and research. Applicants are encouraged to discuss the estimand and an aligned study design and aligned method of estimation (analytical approach) at CHMP Scientific Advice or Protocol Assistance procedure.

4.2.2. Placebo effect and strategies to address high placebo response

Some specific variables related to bipolar disorder appear to be more frequently associated with the placebo effect, for example low symptomatic severity, higher fluctuation in mixed episode, patients with first episode, and rapid cycling with spontaneous remission. Appropriate population selection, careful screening process and appropriate training of investigators for example in terms of scoring may help to reduce high placebo response during the study. Meta-analysis of RCTs in acute mania do suggest manic patients to be less responsive to placebo as compared to depression.

Use of a placebo run-in period (single- or double-blind) is considered problematic with regard to the generalisability of the results to the population treated in clinical practice, since patients included in the studies may not correspond to the target population. This concern also applies to study designs implementing a second randomisation step of placebo non-responders and also potentially placebo responders.
Enrichment strategies that identify and segregate placebo responders from the primary analysis could be provided in the case of phase II studies and not for confirmatory phase III studies. In such a case, additional efficacy analyses, which will include the whole population and allow comparisons between placebo responders and non-responders, should be also submitted. For such studies, further discussion on the relevant estimand may be required in particular with respect to the population attribute.

If a placebo run-in or enrichment strategies are planned a CHMP scientific advice is recommended for a case-by-case consideration.

4.2.3. Investigation of relapse and recurrence

Given the nature of bipolar disorder—acute episodes of both mania/hypomania and depression—and the inherently recurrent nature of the disorder, an optimal treatment should include both acute treatment and long-term maintenance strategies.

The aim of the demonstration of antimanic or antidepressant efficacy is to observe improvement in a generally accepted scale for the acute phase, corresponding to the current (index) manic or depressive episode.

In mania and in bipolar depression, usually a response criterion of decrease of 50 % or more on an appropriate assessment scale is applied to define treatment response (see also section 4.3.2). The next step of the clinical programme should be to investigate maintenance of the initial anti-manic or anti-depressive effect, throughout the current manic or depressive episode (relapse prevention). The duration of the maintenance phase is usually set at about 12 weeks for mania and 3 to 6 months for bipolar depression, to correspond with the average duration of an episode of either mania or depression. In any individual, however, the duration of an episode varies considerably and may be more (or less). As this might affect the interpretation of the results, the 6 months cut-off point is not used for regulatory purposes. Instead, the guideline focuses on showing effect during the current episode and/or prevention of the next episode.

For MAA it should be shown that a short-term effect can be maintained during the current (index) episode (relapse prevention) (section 4.3.1).

Prevention of the next episode(s) or recurrence prevention is a worthwhile treatment goal. It is encouraged to evaluate this in specific studies (section 4.4). Patients in full remission, preferably those in recovery i.e. sustained remission, should be randomised to test product or placebo and/or suitable comparator. Study duration will be dependent on the frequency of episodes in the study population and

should be justified accordingly. Recurrence should be prespecified as either a manic or a depressive episode that fulfils current DSM-5 criteria and a certain degree of severity on a validated rating scale.

For a given patient in the everyday clinical practice, the duration of treatment depends on the rate and the number of his/her recurrences. Patients with a history of higher frequency of manic or depressive episodes should be included in the recurrence prevention investigation and the recent recurrence rate should be taken into account when planning the duration and power of the study.

4.2.4. Study population

Patients should be diagnosed and classified according to the criteria of DSM 5 or ICD 11 and subsequently classified according to the specifiers, severity of disease based on number of symptoms, their severity and degree of functional disability (e.g. mild, moderate, severe), and longitudinal course of their disease (e.g. in partial remission, in full remission). The diagnosis should be made by a qualified psychiatrist and confirmed with the use of a structured assessment tool e.g. the SCID-I (Structured Clinical Interview for DSM 5 Axis I Disorders) or other instruments. Further descriptive parameters (co-morbidity, substance abuse), demographic characteristics and a detailed disease history (duration of disease, duration and number of episodes, naïve or exposed status, and previous treatment outcome) should be documented.

In contrast to proof-of-concept studies, for confirmatory efficacy studies it is important to maximise external validity without unduly compromising internal validity. This may impact the selection of a study population. Inclusion and exclusion criteria should be clearly justified.

Patients in an acute phase (with florid symptoms) may generally show a much larger response to pharmacological treatment compared to patients in a chronic phase. Therefore, if patients in different phases of the disease are included in the same clinical study, stratified randomisation is recommended to ensure that patients in different stages are evenly represented throughout treatment arms. It is recommended to include at least 20% of patients with a disease history of less than 5 years. The inclusion criteria should define the patient population in terms of severity of symptoms. Cut-off scores can be based on efficacy measurement scales (e.g. Young Mania Rating Scale, YMRS).

As compliance is a problem in these patients, compliance controls and screenings for additional illicit psychotropic substances are recommended.

4.2.5. Comedication

Polypharmacy is often applied in clinical practice for treatment of bipolar disorder. In the interests of the external validity of the study, most types of co-medication should be permitted and exclusion criteria for concomitant medications should be justified. A wash out of excluded comedication should be applied.

When monotherapy is claimed, treatments that may affect the test treatment should be excluded.

Standardised psychotherapy, psychoeducation, support or counselling may be offered as supplementary treatment, but their use should be prospectively defined in the protocol and documented in the study report, including a discussion on the effects on treatment.

4.2.6. Extrapolations

Patients included in the studies will be diagnosed as having bipolar disorder using accepted diagnostic criteria, DSM-5 or ICD-11. If mania, hypomania or depression in the context of another disorder are strived for, specific studies in this patient population should be provided.

A major depressive episode as seen in BD II also occurs in the framework of major depressive disorder. Based on current evidence, and due to the different disease characteristics (or different nature) of the two disease entities (BD and MDD), extrapolation of short term and maintenance of efficacy in adults from bipolar depression to unipolar depression does not seem possible. A switch from depression to mania should always be considered in BD, but not in MDD. Extrapolation from unipolar depression to bipolar depression need to be considered on a case-by-case basis. Some specific issues, like duration of the episodes, switching rates and population selection and safety data, as addressed in other sections of this guideline need to be taken into consideration.

Data from rapid cyclers or other patient populations defined by a specifier as listed in diagnostic manuals cannot be extrapolated to the whole group of patients with bipolar disorder (see section 4.5.1).

For studies required in paediatric patients and possible extrapolations reference is made to section 4.6.2.

4.3. Methodological features

4.3.1. Study design

Study planning, design, conduct, analysis, and interpretation must be aligned with the estimand of interest and special focus should be on collecting all relevant data for the targeted estimand(s) (primary and supplementary) (see sections 4.2.1. and 4.3.3.).

Confirmatory short-term studies should be double-blind, randomised, parallel group, placebo-controlled studies. Crossover designs are unsuitable for studies in patients with BD.

A two-arm non-inferiority study is not an option as the sole basis for demonstrating efficacy as a reliable non-inferiority margin is difficult to determine. E.g.in some studies, lithium, which is believed to be the gold standard for the treatment of bipolar disorder, failed to separate from placebo.

Screening and run-in periods

Following screening, qualitative and quantitative baseline assessments should be conducted in a short run-in period. Regarding the wash out of prior medication, tapering in a single blind placebo run-in period of sufficient duration aiming at elimination of prior treatment without substantial worsening of symptoms should be applied. Typically, a few days will be appropriate for run-in. Placebo responders should not be excluded from randomisation. In some instances, screening, baseline assessments, randomisation and start of study medication may be performed in a single day, especially if patients are severely ill.

4.3.1.1. Short-term studies

4.3.1.1.1. Acute manic episodes

The preferred duration for demonstrating efficacy in acute mania is a 3-week clinical study. A meta-analysis of BD studies at the individual patient data (IPD) level has demonstrated that efficacy in the first week of treatment is predictive of further response.

The occurrence of switching to depression should also be investigated. Criteria for "switching" (e.g. change in symptoms) have to be defined in advance of the studies.

Improvement should be documented as the difference between baseline and post-treatment score in signs and/or symptoms but should also be expressed as the proportion of responders or remitters. In mania, a 50% improvement of a patient on a usual rating scale is accepted as a clinically relevant response. Other definitions of responder may be used, e.g. other grades of response or proportion of patients with full remission. Criteria for response and remission must be pre-specified and justified in the study protocol.

4.3.1.1.2. Major depressive episode in the framework of bipolar disorder

Study drugs may be used as monotherapy, or in combination. However, this has consequences for the claim. For a monotherapy claim, patients need to be off lithium or other mood stabilisers at baseline for a substantial period of time because of possible rebound phenomena. Moreover, they should not be lithium resistant (depression under lithium therapy) as this has consequences for the claim requested.

Studies in the acute phase should be 6 to 8 weeks in duration.

Switching to hypomania or mania is an important issue. Switching criteria need to be defined so that not only mania meeting diagnostic criteria is considered. Incidence of switching needs to be established in relation to placebo.

4.3.1.2. Long-term studies

Due to the character of the disorder, longer studies are necessary to demonstrate that the acute effect is maintained during an episode (relapse prevention).

4.3.1.2.1. Mania

For mania, maintenance of effect during the episode has to be shown. A placebo control during this period will be difficult. Therefore, an active comparator is acceptable, provided that assay sensitivity is taken into account. A possible design is a comparison of placebo, test product and active control for three weeks followed by a two-arm phase for the remaining nine weeks, comparing only test product and active control. At least 2 assessment time-points are used in this study: one at 3 weeks, comparing the results of active comparator and test product with placebo and one at 12 weeks comparing the results of active comparator and test product. The choice of comparator should be justified and non-inferiority to the comparator in maintenance of effect should be demonstrated. Non-inferiority margin should be pre-defined and justified in the study protocol.

4.3.1.2.2. Major depressive episode in the framework of bipolar disorder

Showing maintenance of effect for a depressive episode in bipolar disorder - like in major depression- is needed. A duration of 3 to 6 months should be sufficient. Randomized withdrawal study is the preferred design. Please refer to the Guideline on clinical investigation of medicinal products in the treatment of depression (EMA/CHMP/185423/2010, Rev.3) for details.

4.3.2. Efficacy endpoints

Efficacy should be assessed by adequate, validated rating scales. The choice of instruments should be justified from the test quality criteria (reliability, validity and sensitivity to change). For the assessment of response, specific developed rating instruments are necessary. In addition, global assessment of severity and improvement is needed.

The rates of response/remission need to be evaluated in order to adequately assess clinical relevance, besides statistically significant results in rating scales.

4.3.2.1. Manic episodes

The Young Mania Rating Scale (YMRS), the Manic State Rating Scale, and the Bech-Rafaelsen Mania Scale (MAS) are measurements with well-known reliability and validity. These clinician-rated scales can be used for assessing improvement in manic episodes and maintenance studies. The development of new scale is encouraged, although the traditional instruments are considered valid for standard claims in BD.

Furthermore, a global scale, e.g. the CGI – BP, can be used as secondary endpoint. Such a scale gives an impression of the clinical relevance of the effect seen on the specific scales used. Other tools like actigraphs and electronic devices are currently applied to provide additional information on wellbeing of patients.

The use of lifecharting may also be used as a secondary endpoint.

Reliable depression scales, e.g., Hamilton Rating Scale for Depression (HDRS), Montgomery-Åsberg Rating Scale (MADRS), Bech-Rafaelsen Melancholia Scale (MES) should be used to assess switching from mania to depression (reference is made to the depression guideline).

Choice of endpoints should be indicated in the protocol, and the clinical relevance of expected effects should be defined by the degree of symptom reduction and discussed in the protocol with reference to other comparable data or publications available. In acute manic patients, onset of action and time to response is an important parameter and should be investigated.

4.3.2.2. Major depressive episode in the framework of bipolar disorder

The conventional depression scales (used in unipolar depression) like the HDRS, preferably the 17-item scale, or the MADRS are recommended; however, other validated scales might be acceptable as well. See also the Guideline on clinical investigation of medicinal products in the treatment of depression.

4.3.3. Statistical considerations

Generally, efforts should be made to collect all relevant data for the primary and important additional estimands (e.g. follow-up regardless of intercurrent events) to minimize the need to rely on untestable assumptions in the analysis and interpretation of the study results. Still, handling of missing data is of particular concern, as a relevant amount of missing data can be expected.

Given the similarities to the unipolar depression setting with respect to the relevant estimand(s), advice given in section 4.3.3 of the unipolar depression guideline (EMA/CHMP/185423/2010, Rev.3) also applies in this setting. Mainly, use of multiple imputation-based approaches (e.g. placebo-based multiple imputation) aligned to the targeted estimand accompanied by relevant sensitivity analyses is advised. However, further considerations are warranted with respect to the switch from one pole of the disorder to the other.

When several confirmatory conclusions on efficacy (e.g. effect for acute and maintenance treatment in 3+9-week study for treatment of mania) are addressed within a single study the approach to properly control for multiplicity should be pre-specified (usually co-primary assessment).

4.4. Specific Claims

4.4.1. Treatment resistance and partial response

There are no globally accepted definitions of treatment resistance or partial response in bipolar disorder in general. With respect to treatment resistance and partial response in bipolar depression, reference is made to the EMA guideline on depression. Studies aiming to address patients with treatment resistance or partial response should incorporate clear inclusion criteria defining the target patient population and the chosen approach should be justified. CHMP scientific advice is recommended.

4.4.2. Recurrence prevention

For demonstrating prevention of recurrence of (hypo)manic/ depressive symptoms (new episodes), long-term studies are needed versus placebo and/ or a suitable comparator, e.g. lithium.

A randomised withdrawal study is a possible study design. In randomised withdrawal studies, efficacy usually is expressed as rate of patients worsening (relapsing) and/or time to this event. Both efficacy criteria are of interest and should be submitted. The choice of one of them as primary and the relevance in clinical terms will depend on the target population which is selected based on pre-defined criteria and will need to be justified. The choice of comparator has to be clearly justified and the risks should be carefully considered.

The duration of the study should be long enough to demonstrate an effect on recurrence of either a manic or a depressive episode.

The number of patients should be sufficient to address both recurrences of depression and of mania. It is recommended to use the same rating scales to document recurrence of any mood episode as those used in the acute treatment studies.

Criteria for stabilisation, prodromal signs and recurrence of both (hypo)manic and depressive symptoms should be predefined and have to be discussed in the protocol with reference to other comparative data or publications available.

The claim for recurrence prevention does not have to be supported by short-term efficacy also in acute (hypo)manic/depressive symptoms during an episode itself.

4.4.3. Improvement in cognitive function

To support a claim for efficacy on cognitive aspects in patients with BD, specific studies should be performed, targeted to examine this effect specifically beyond the effect on mood symptoms. If effect on cognitive function is claimed, specific effect on cognitive function needs to be shown that can clearly be disentangled from the overall effect on mood. Cognitive complaints are present during both, a manic or depressive episode, and may vary according to the mood state. The patient population should be clearly defined in terms of a range of relevant measures of cognitive functioning and state of the disorder, e.g. manic/depressed, in remission etc. To avoid confounding by changes in mood, patients in remission or in partial remission may be the most appropriate population. As in schizophrenia, a relatively younger patient population might be more appropriate for testing effect on cognition to avoid confounding by age. However, in this case, extrapolation to older patient population will need to be justified.

The effect of treatment on cognitive functioning should be demonstrated as the difference between baseline and endpoint on a validated cognitive functioning scale. Whatever tool is used, mere reduction on specific items of a larger test battery is not acceptable. The outcome should be such that relevance to the patients functioning is clear.

4.5. Bipolar disorder with specifiers

The DSM-5 includes a number of specifiers for bipolar and related disorders with defined diagnostic features. A few of these specifiers with regulatory precedence or recent interest within the field are briefly discussed below.

If a claim for a sub-population as defined by a specifier is pursued, a dedicated study with specific inclusion criteria and adequate endpoints is required.

4.5.1. Rapid cycling

Bipolar disorder is specified as rapid cycling when at least four mood episodes (manic, hypomanic or major depressive) have occurred in the previous 12 months. Rapid cycling pattern is associated with poorer prognosis and in general functional impairment of these patients is higher than in patient with non-rapid cycling bipolar disorder. It should be noted that rapid cycling pattern is often transient, and patients can convert – sometimes after many years – to non-rapid bipolar disorder.

Evidence-based data are too limited to provide further advice, but as rapid cycling is defined as four or more acute mood episodes within the past 12 month, it might be considered to define efficacy by a clinically relevant reduction of cycles. Next to the acute efficacy, the potential of a medicinal product to prevent relapsing is important in rapid cycling bipolar disorder, necessitating long-term studies.

Data from rapid cyclers cannot be extrapolated to the whole group of patients with bipolar disorder.

4.5.2. Mixed features

When a manic, hypomanic or depressed episode presents together with at least three symptoms from the opposite pole, the mixed features specifier is used. Patients with mixed features are more susceptible for rapid cycling pattern, have higher risk for suicidality and poorer psychosocial functioning compared to patients without mixed features episodes. Episodes with mixed features – if left untreated - also have a longer duration.

A separate claim specifically in patients with mixed features would require dedicated studies. Considering that various patient populations are possible (e.g. depressive episode with manic features and vice versa), the included patient population will need to be clearly defined in the inclusion criteria and may affect the final label recommendations.

Used endpoints should cover both the mood episode itself and the features of the opposite end.

4.5.3. Peripartum onset

The specifier 'with peripartum onset' in the DSM-5 refers to a mood episode with an onset either during pregnancy or in the four weeks following delivery. The most common psychopathology among postpartum bipolar patients is depression. Postpartum psychosis is estimated to occur in 10-50% of postpartum bipolar patients and increases the risk for suicide and infanticide.

Whether mood episodes with peripartum onset are distinct from mood episodes without peripartum onset is still a matter of debate, however based on identified differences in for example hormone contributions, a claim in post-partum bipolar disorder should be supported by specifically designed studies in this sub-population.

4.6. Special populations

4.6.1. Older people

In general, no specifically designed studies in elderly patients are necessary. However, information for this group concerning dose recommendations and safety is necessary. Therefore, the number of elderly patients included in the studies should suffice to generate these data. Safety data should be analysed separately with special attention to safety concerns relevant to this age group e.g. (cardio-) vascular and stroke events, as well as diabetes and CNS effects.

4.6.2. Children and adolescents

There is insufficient evidence for the existence of BD in childhood (<13 years of age). No studies are recommended in this age group.

In regular studies for BD, adolescents can be included as of 13 years of age. Sufficient patients should be included to allow a separate analysis for this younger patient group (13-18 years of age). Full extrapolation of efficacy and safety data from adults is not considered appropriate. Short term-efficacy data need to be generated either in a subgroup analysis or a dedicated study with adolescents. Maintenance of effect and long-term efficacy studies may not be necessary and extrapolation from adults could be acceptable, provided that robust evidence of short term-efficacy is available from both adults and adolescents with comparable effect sizes. Any extrapolation would have to be justified on a case-by-case basis and be supported by appropriate PK data/models. Long-term safety data still need to be generated. Post-marketing long-term safety studies in adolescents should be structured to include also efficacy endpoints to support accepted extrapolation of long-term efficacy. Please refer to the Reflection paper on the use of extrapolation in the development of medicines for paediatrics (EMA/189724/2018).

If a claim for delaying onset of established disease in children or adolescents with prodromal symptoms is pursued, CHMP scientific advice is recommended.

4.6.3. Sex issues / differences

In contrast to anxiety disorders and unipolar mood disorder, bipolar disorder does not differ in its incidence between men and women. No consistent sex differences have been found in a number of variables including rates of depressive episodes, severity of the illness, response to treatment and suicidal behaviour in BD. However, women are at increased risk of bipolar type II, and hypomanic, rapid cycling and mixed episodes. Additionally, the onset of bipolar occurs later in women and they are more susceptible to the seasonal pattern of mood disturbance. Important differences also exist in patterns of comorbidity with women suffering more often from anxiety and men presenting substance-related disorders. Up to now there is no robust evidence on sex differences in response to mood stabilizers.

Hence, during the development of medicines for BD, predefined analyses of gender specific groups are welcomed.

4.7. Safety evaluation

In general, the content of ICH E1 should be taken into consideration.

Identified adverse events (AEs), including serious AEs and AEs leading to withdrawal, should be characterised in relation to duration of treatment, dosage, duration of the AE age, frailty, and other relevant variables. Adverse event scales should be standardised for use in studies with psychotropic drugs (e.g., UKU Side Effect Rating Scale (SERS)). Clinical observations should be supplemented by appropriate laboratory tests and cardiac recordings (e.g., ECG). AE rates should be presented for the test treatment, placebo and active comparators.

As treatment durations, including the long–term open label studies, will generally be longer for the test treatment as compared to other treatments (e.g., placebo), the data should be presented in a suitable way for comparisons of event rates.

Special efforts should be made to assess potential AEs that are characteristics of the class of drugs being investigated in view of actions on specific receptor sites. Particular attention should be paid to anti-dopaminergic, anti-cholinergic or cholinergic, anti-histaminergic, serotoninergic and adrenergic, and to glutamatergic or anti-GABAergic AEs, if relevant.

4.7.1. Specific adverse events to be monitored

4.7.1.1. Psychiatric adverse events

Psychiatric adverse events typically represent a large proportion of the AEs reported in studies in patients with bipolar disorders. These events may be related to the disorder itself, to comorbid conditions, and to the study medication. In order to explore the risk of an adverse effect on the severity of the disorder being treated, the proportion of patients deteriorating during treatment should be documented using the primary efficacy measure.

As part of the adverse event data, undesirable psychiatric effects including depression and anxiety should be measured using validated rating scales.

4.7.1.2. Adverse effects on cognitive functioning

A detrimental effect on cognition should be monitored using validated rating scales, which may be identical to those used to support an efficacy claim. Effects on cognition, reaction time, driving and severity of sedation should also be studied. In the adolescent population specific issues such as memory, learning, school performance, etc. should be studied in relation to both the safety and efficacy perspective.

4.7.1.3. Overdose

Depending on the mechanism of action risks and effects of overdose should be studied particularly with regard to serotonin-syndrome, QT-prolongation and delirium.

4.7.1.4. Suicide

The potential for the test product to precipitate suicidal thoughts and behaviour should be actively evaluated in all age groups using validated rating scales (e.g. InterSePT Scale for Suicidal Thinking, Columbia Suicidality Severity Rating Scale (C-SSRS), the SIBQ (Suicidal Ideation and Behaviour Questionnaire) or other validated instruments). Rates of suicidal events (from suicidal ideation to

completed suicide) should be presented and narrative summaries of suicidal patient statements or behaviours should be provided.

4.7.1.5. Metabolic risk factors

The effects on weight, glucose metabolism and lipid metabolism should be actively measured using standard laboratory measures. The metabolic profile of the test product should be thoroughly characterised in comparison with placebo and active comparator(s).

4.7.1.6. Haematological adverse events

Special attention should be paid to incidence of neutropenia, agranulocytosis and aplastic anaemia.

4.7.1.7. Endocrinological adverse events and sexual dysfunction

Special attention should be paid to effects on sexual functioning, libido, galactorrhoea, and gynaecomastia. Investigation of neuro-endocrinological parameters relating to prolactin is necessary. In the adolescent population effects on growth and sexual maturation require specific attention and should be closely monitored.

4.7.1.8. Cardiovascular adverse events

Due to the known cardiovascular effects associated with drugs used in the treatment of BD, cardiac adverse events should be actively monitored. Reported adverse events that might represent orthostatic hypotension or arrhythmia (including syncope, loss of consciousness, etc.) should be presented where relevant. The effect on QT-interval prolongation should be investigated in accordance with the ICH E14 and ICH E14/S7B guideline.

4.7.1.9. Extrapyramidal symptoms (EPS)

There is concern that patients with affective disorders show a higher sensitivity to suffer from acute extrapyramidal side effects and a higher incidence of tardive dyskinesia compared to patients with schizophrenia. Therefore, if antipsychotics are used for augmentation or as treatment option in patients with bipolar disorders, rates of extrapyramidal symptoms should be presented. In addition, the extent and severity of EPS should be actively measured using validated and specifically designed rating scales. Dose – response relationships of EPS should be explored. During the wash-out phase prior to acute studies, possible tardive EPS should be measured to distinguish this from acute EPS due to the test treatment.

Tardive dyskinesia occurs late in treatment and is reported for both atypical and typical antipsychotics.

4.7.1.10. Serotonin syndrome/ Neuroleptic malignant syndrome

Serotonin syndrome (SS) can be caused by excessive serotonergic agonism in central and peripheral nervous system serotonergic receptors and has been described for many antidepressants. The clinical symptoms include neuromuscular hyperactivity, autonomic hyperactivity and altered mental status.

Neuroleptic malignant syndrome (NMS) consists of similar clinical symptoms and has been reported for all antipsychotics.

4.7.1.11. Rebound/ withdrawal phenomena/ dependence

When pharmacological treatment is stopped, rebound and/or withdrawal phenomena may occur. Studies should be designed in such a way, that these phenomena can be studied. In some of the short-term and long-term clinical studies, treatment should be stopped abruptly, and patients should be followed for a suitable duration, in other studies careful tapering off might be more appropriate depending on the mechanism of action of the medicinal product. Occurrence of rebound and/or withdrawal phenomena should be scored at the appropriate time.

Animal studies will be needed to investigate the possibility of dependence in new classes of medicinal products or when there is an indication that dependence may occur.

Depending on the results of these studies further studies in humans may be needed.

4.7.1.12. Long term safety

Since a depressive episode can have a duration of up to 2 years, the total clinical experience should include data on exposure of a large and representative group of patients in line with the guideline on population exposure of at least 12 months.

4.7.1.13. Elderly patients

Certain adverse events such as anticholinergic effects, delirium, sedative effects, cardiovascular and hypotensive effects, dizziness, falls, effect on food intake and functional decline, have been observed in elderly patients and these should be monitored in the studies designed for elderly patients.

4.7.1.14. Children and adolescence

Rather than relying on spontaneous AE reporting, potential treatment-emergent adverse events such as somnolence, sexual disturbances, weight gain, affective symptoms and suicidality, discontinuation/rebound symptoms, etc. should be clearly defined and actively monitored. Validated questionnaires/ scales/ tests should be used for the assessment of adverse events.

Long-term effects on learning, development, growth and sexual function may be studied post-marketing, but appropriate protocols should be available when the use in children is applied for.

5. References

American Psychiatric Association: *Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition*. Arlington, VA: American Psychiatric Association, 2013.

Beyer DKE, Freund N. Animal models for bipolar disorder: from bedside to the cage. Int J Bipolar Disord. 2017 Oct 13;5(1):35. doi: 10.1186/s40345-017-0104-6. PMID: 29027157; PMCID: PMC5638767.

Brietzke E, Rosa AR, Pedrini M, Noto MN, Kapczinski F, Scott J. Challenges and developments in research of the early stages of bipolar disorder. Braz J Psychiatry. 2016 Oct-Dec;38(4):329-337.
Craddock N, Sklar P. Genetics of bipolar disorder. Lancet. 2013;381(9878):1654.

De Crescenzo F, Economou A, Sharpley AL, Gormez A, Quested DJ. Actigraphic features of bipolar disorder: A systematic review and meta-analysis. Sleep Med Rev. 2017 Jun;33 58-69. doi:10.1016/j.smrv.2016.05.003. PMID: 28185811.

Diflorio A, Jones I (2010). Is sex important? Gender differences in bipolar disorder. International Review of Psychiatry 22(5): 437-452.

Dome P, Rihmer Z, Gonda X. Suicide risk in bipolar disorder: a brief review. Medicina 2019;55:403
Fountoulakis KN, Kontis D, Gonda X, Yatham LN. A systematic review of the evidence on the treatment of rapid cycling bipolar disorder. Bipolar Disord. 2013 Mar;15(2):115-37. doi: 10.1111/bdi.12045. PMID: 23437958.

Goldstein BI, Birmaher B, Carlson GA, DelBello MP, Findling RL, Fristad M, Kowatch RA, Miklowitz DJ, Nery FG, Perez-Algorta G, Van Meter A, Zeni CP, Correll CU, Kim HW, Wozniak J, Chang KD, Hillegers M, Youngstrom EA. The International Society for Bipolar Disorders Task Force report on pediatric bipolar disorder: Knowledge to date and directions for future research. Bipolar Disord. 2017 Nov;19(7):524-543. doi: 10.1111/bdi.12556.

Kokras N, Dalla C. Sex differences in animal models of psychiatric disorders. Br J Pharmacol. 2014 Oct;171(20):4595-619. doi: 10.1111/bph.12710. Epub 2014 Jul 1. PMID: 24697577; PMCID: PMC4209934.

Logan RW, McClung CA. Animal models of bipolar mania: The past, present and future. Neuroscience. 2016 May 3;321:163-188. doi: 10.1016/j.neuroscience.2015.08.041. Epub 2015 Aug 24. PMID: 26314632; PMCID: PMC4766066.

Martínez-Arán A, Vieta E, Reinares M, Colom F, Torrent C, Sánchez-Moreno J, Benabarre A, Goikolea JM, Comes M, Salamero M. Cognitive function across manic or hypomanic, depressed, and euthymic states in bipolar disorder. Am J Psychiatry. 2004 Feb;161(2):262-70. doi: 10.1176/appi.ajp.161.2.262. PMID: 14754775.

McIntyre RS, Berk M, Brietzke E et al. Bipolar disorders. Lancet 2020;396: 1841-56.

McIntyre RS, Alda M, Baldessarini RJ, Bauer M, Berk M, Correll CU, Fagiolini A, Fountoulakis K, Frye MA, Grunze H, Kessing LV, Miklowitz DJ, Parker G, Post RM, Swann AC, Suppes T, Vieta E, Young A, Maj M. The clinical characterization of the adult patient with bipolar disorder aimed at personalization of management. World Psychiatry. 2022 Oct;21(3):364-387. doi: 10.1002/wps.20997. PMID: 36073706; PMCID: PMC9453915.

Salagre E & Vieta E (2022). Biomarkers in bipolar disorder: an overview, in Machado-Vieira R, Soares JC (ed.) *Biomarkers in Bipolar Disorders.* Academic Press, pages 1-18.

Sigitova E, Fišar Z, Hroudová J, Cikánková T, Raboch J. Biological hypotheses and biomarkers of bipolar disorder. Psychiatry Clin Neurosci. 2017 Feb;71(2):77-103. doi: 10.1111/pcn.12476. Epub 2017 Jan 6. PMID: 27800654.

Spearing MK, Post RM, Leverich GS, Brandt D, Nolen W. Modification of the Clinical Global Impressions (CGI) Scale for use in bipolar illness (BP): the CGI-BP. Psychiatry Res. 1997 Dec 5;73(3):159-71. doi: 10.1016/s0165-1781(97)00123-6. PMID: 9481807.

Welten CC, Koeter MW, Wohlfarth TD, Storosum JG, van den Brink W, Gispen-de Wied CC, Leufkens HG, Denys DA. Early Nonresponse in the Antipsychotic Treatment of Acute Mania: A Criterion for Reconsidering Treatment? Results From an Individual Patient Data Meta-Analysis. J Clin Psychiatry. 2016 Sep;77(9):e1117-e1123. doi: 10.4088/JCP.15r10051. PMID: 27780320.

Yildiz A, Vieta E, Tohen M, Baldessarini RJ. Factors modifying drug and placebo responses in randomized trials for bipolar mania. Int J Neuropsychopharmacol. 2011 Aug;14(7):863-75.