LUNDS UNIVERSITET

Institutionen för psykologi
Psykologprogrammet

COMPARISON OF PSYCHEDELIC THERAPIES VERSUS PLACEBO IN THE TREATMENT OF PSYCHIATRIC ILLNESSES: A SYSTEMATIC REVIEW OF RANDOMIZED TRIALS INCLUDING PLACEBO CONTROLS

Aleksandra Barstowe

Psykologexamensuppsats. 2023

Handledare: Johannes Björkstrand

Examinator: Sean Perrin

Abstract

Psychedelics have long been recognized for their powerful effects on the human psyche and psychedelic-assisted therapy is becoming increasingly acknowledged as an effective therapeutic intervention. Placebos are used in clinical research to validate the efficacy of a therapy or treatment. Researchers are interested in showing that a certain treatment is superior compared to the placebo. An evaluation of the methodological rigor in recent studies looking at the potential of psychedelics for the treatment of psychiatric illnesses is an important step forward to re-open a scientific study of these substances and their medical use for psychiatric patients. The aim of this systematic literature review was to investigate if psychedelic substances have an influence on symptom reduction for patients suffering from mental illness as compared to placebos and if placebo controls are valid in psychedelic trials, i.e., if the blinding is maintained. The review has presented that an overwhelming majority of the evaluated studies showed large and positive effects of psychedelics on participant's symptom changes. Collectively the studies reviewed displayed a large degree of heterogeneity. In all nine studies the blinding was either insufficient or poor. In conclusion, reviewed studies show large and positive treatment outcomes for patients suffering from psychiatric illnesses such as alcohol use disorder, anxiety with or without a life-threatening disease, anxiety and depression during life-threatening cancer, treatment-resistant depression and major depressive disorder. The lack of successful blinding procedures indicates that the methodological shortcomings displayed in the psychedelic studies of the first wave of psychedelic research persist in these most recent studies on the subject.

Keywords: psychedelic-assisted therapy, placebo, psychiatric illness

Sammanfattning

Psykedelika har länge erkänts för sina kraftfulla effekter på det mänskliga psyket och psykedelisk-assisterad terapi blir alltmer erkänd som en effektiv terapeutisk intervention. Placebo används i klinisk forskning för att validera effekten av en terapi eller behandling. Forskare är intresserade av att visa att en viss behandling är överlägsen jämfört med placebo. En utvärdering av den metodologiska rigoriteten i nya studier som tittar på potentialen hos psykedelika för behandling av psykiatriska sjukdomar är ett viktigt steg framåt för att återuppta studien av dessa substanser och deras medicinska användning för psykiatriska patienter. Syftet med denna systematiska litteraturöversikt var att undersöka om psykedeliska substanser har en inverkan på symptomförändring hos patienter som lider av psykisk ohälsa jämfört med placebo och om placebokontroller är giltiga i psykedeliska prövningar, d.v.s. om blindningen upprätthålls. Granskningen har visat att en överväldigande majoritet av de utvärderade studierna visade stora och positiva effekter av psykedelika på deltagarnas symtomförändringar. Tillsammans visade de granskade studierna en stor grad av heterogenitet. I alla nio studierna var blindningen antingen otillräcklig eller dålig. Sammanfattningsvis visar granskade studier stora och positiva behandlingsresultat för patienter som lider av psykiatriska sjukdomar som alkoholmissbruk, ångest med eller utan en livshotande sjukdom, ångest och depression under livshotande cancer, behandlingsresistent depression och egentlig depression. Avsaknaden av framgångsrika blindnings-procedurer indikerar att de metodologiska bristerna som uppvisades i de psykedeliska studierna av den första vågen av psykedelisk forskning kvarstår i dessa senaste studier i ämnet.

Nyckelord: psykedelisk-assisterad terapi, placebo, psykiatrisk sjukdom

Table of contents

Psychedelic therapies versus placebo in the treatment of psychiatric illnesses

Psychedelics

Psychedelics have long been recognized for their powerful effects on the human psyche (Halberstadt et al., 2018) and have been used for therapeutic purposes for thousands of years to give rise to altered states of consciousness in their users (Wheeler & Dyer, 2020). Naturally occurring hallucinogens have an extensive history of use in the Americas. The hallucinogenic substance in Ayahuasca, DMT, has traditionally been used by indigenous groups in Northwestern Brazil and in the Northwestern Amazon. Mescaline, the psychoactive compound in the peyote cactus, is used by indigenous peoples of Northern Mexico, while psilocybin found in many species of hallucinogenic "magic" mushrooms is widely used by indigenous groups in Central Mexico. The widespread use of psychedelics like these in rituals and religious ceremonies could be explained by their capacity to induce states of consciousness that resemble mystical experiences (Dos Santos & Hallak, 2020). The substances are also used worldwide by many as a form of self-medication for mental health reasons (Carhart-Harris & Goodwin, 2017).

Substances that are psychedelics are pharmacologically classified as serotonergic receptor agonists and partial agonists (Halberstadt et al., 2018), also known as serotonergic hallucinogens (Ko et al., 2022). Agonists are substances that bind to receptors and that are capable of triggering a response while partial agonists are capable of binding to and activating a receptor but cannot trigger a full response through this binding (Norlén & Lindström, 2014). In the brain, the serotonin 5-HT_{2A} receptor appears to be the principal target of hallucinogenic substances. It plays a key role in the regulation of cortical functions and cognition, being central to the structure-activity relationship (SAR) of serotonergic hallucinogens (Halberstadt et al., 2018). The SAR is an idea explaining the relationship between a molecule's molecular structure and its biological activity. By using the SAR model scientists can better understand and explain which chemical groups are responsible for evoking specific biological effects in the body (Pottie et al., 2020).

For the classic serotonergic hallucinogens there are three chemical types: plant-derived indoleamines, phenylalkylamines and semi-synthetic ergolines. Plant-based indoleamines are partial agonists of the 5-HT_1, 5-HT_2, 5-HT_6 and 5-HT_7 receptors and include substances such as *N,N*-dimethyltryptamine (DMT), 5-methoxy-DMT (5-MeO-DMT), psilocybin and 4-hydroxy-DMT. Semi-synthetic ergolines are also partial agonists of the 5-HT_1, 5-HT_2, 5-HT_6

and 5-HT_7 receptors. Furthermore, they also act upon the D_1 and D_2 dopamine receptors and adrenergic receptors. An example of a semi-synthetic ergoline is lysergic acid diethylamide (LSD). Phenylalkylamines are selective agonists of 5-HT_2 receptors (5-HT_{2A}, 5-HT_{2B} and 5-HT_{2C}). This class of hallucinogens includes substances such as mescaline, the synthetic 'amphetamines' 2,5-dimethoxy-4-iodoamphetamine (DOI) and 2,5-dimethoxy-4-bromoamphetamine (DOB) (Halberstadt et al., 2018; Vollenweider & Preller, 2020).

There is evidence that activation of 5-HT_{2A} receptors located in cortical and subcortical structures is the unifying mechanism through which psychedelics mediate their behavioral and psychological effects in animals, including humans (Vollenweider & Preller, 2020). In many cases it is necessary to rely on animal behavioral data to gain understanding of the SAR of psychedelics because molecular pharmacology data, for example information on serotonin affinity and functional potency, is not available for humans except from very limited studies carried out half a century ago. Molecular pharmacology data from more recently developed laboratory compounds tested on animals has however shown these substances to be serotonin 5-HT_{2A} agonists or partial agonists and can therefore be used as a basis for clinical inferences even though they lack formal clinical studies in humans (Halberstadt et al., 2018).

In humans, the 5-HT_2 receptors are highly expressed in the apical dendrites of layer 5 pyramidal (LP5) neurons in the cortex and are particularly enriched in the prefrontal cortex (PFC). A smaller proportion of 5-HT_2 receptors are also located presynaptically on thalamocortical afferents projecting to the neocortex. The PFC is important in cognition and control over subcortical regions. Changes in the PFC caused by stress or other psychiatric illnesses are believed to be the cause of deficits we can observe in learning, memory, motivation and reward-seeking that characterize psychiatric illnesses such as depression, bipolar disorder, anxiety, obsessive-compulsive disorder, schizophrenia, alcohol and substance abuse disorder. There is also evidence emerging that the PFC is the convergence point underlying the pathophysiology of many psychiatric illnesses (Vollenweider & Preller, 2020). Several circuits originating in the PFC control behaviors that are relevant to the treatment of psychiatric illnesses such as depression, anxiety and addiction (Vargas et al., 2021).

Increased thickness and cerebral blood flow to and within the PFC have been shown to correlate with the efficacy of pharmaceuticals used for treatment of psychiatric illnesses such as depression. Antidepressants can therefore be said to promote structural plasticity in the PFC.

In the same way it has been shown that psychedelics increase activation of the PFC and regions connected to it, implying that psychedelics also promote neural plasticity (Vargas et al., 2021).

The broad therapeutic utility of psychedelics arguably comes from their ability to impact the structure and function of the layer V pyramidal neurons in the PFC, an area that exerts top-down control over subcortical regions (Carhart-Harris et al., 2013). It is involved in the control of high-level cognitive and emotional processes including the enabling of a sense of self (Vollenweider & Preller, 2020). Disruptions in the functional integrity of the bottom-up cortico-thalamic and top-town cortico-cortical loops, or hyperactivity within this network, has been associated with the dissolution of self-boundaries and alterations of cognitive and emotional processes as can be seen in psychedelic states (Carhart-Harris et al., 2013).

Psychedelic substances may exert different modulatory effects across cortical regions depending on the dose, the specific drug used and presumptively the density of 5-HT_2 receptors in the different neuronal populations (Vollenweider & Preller, 2020), something that should be of value to know when planning and performing psychedelic therapy.

Psychedelic therapy

It has been shown that a single administration of a psychedelic substance in a psychotherapeutic context can have sustained therapeutic effects (Vargas et al., 2021), and psychedelic-assisted therapy is becoming increasingly acknowledged as an effective therapeutic intervention (Daly, 2018). It can be defined as the administration of psychedelics for therapeutic use (Ko et al., 2022) and is a therapy model that refers to the use of one or a few doses of a classic psychedelic in combination with psychotherapeutic support. The therapy often includes a few drug-free preparatory therapy sessions, followed by a drug session and then a few follow-up therapy sessions that allow the patient to integrate the psychedelic experience (Vollenweider & Preller, 2020). These sessions are often led by licensed professionals that have been trained in the administration of psychedelic substances, the monitoring of their use and in the guidance of patients to minimize their distress and support integration of their experience (Mithoefer et al., 2016).

From a patient perspective in the context of psychotherapy, psychedelic substances have the capacity to and often do alter consciousness. Patients can potentially be provided with a deeper healing experience (Belser et al., 2017) during psychedelic-assisted therapy compared to interventions during therapy done at the patient's normal state of consciousness (Wheeler & Dyer, 2020). The use of psychedelics is also often associated with an experience of "ego

dissolution" where the patient feels like their sense of identity or self is going through a disintegration (Letheby & Gerrans, 2017). This in turn is something that Schenberg (2018) hypothesizes could facilitate objectivity in the patient, a change in perspective that can better help them modify maladaptive emotions, thoughts and behavioral patterns often associated with psychiatric illnesses. In this way, psychedelics can be a substitute or a supplement to mainstream interventions, something that can be of extra interest or importance when trying to find ways to treat psychiatric illnesses that have shown resistance to conventional treatment, such as treatment resistant depression (Schenberg, 2018). An important question that remains when treating psychiatric illnesses such as depression or treatment-resistant depression with the aid of psychedelic therapy is whether it is the active psychedelic substance that is providing the deeper therapeutic experience (Belser et al., 2017) or if other factors may be at play such as placebo effects. Notably, the psychedelic experience during treatment is often quite intense, raising the question whether successful placebo controls can be applied in this context.

Placebo

Historically, the broad definition of placebo – an ineffective treatment for a symptom or disorder being treated – meant that almost all therapies available for doctors to prescribe to their patients were in fact placebos. Most of these "medications" were ineffective, with few exceptions. Treatments were developed with the aim of relieving symptoms and curing disease, but they were not based on scientific rationale or an assessment of efficacy. They were instead developed from a combination of metaphysical beliefs, social influences and scientific ignorance. They were also used indiscriminately to treat virtually any kind of symptom or disease (Benedetti, 2014).

Today placebos are widely used in clinical research to validate the efficacy of a therapy or treatment (Benedetti, 2014). The effect of a treatment can be measured through randomized, placebo-controlled clinical trials (RCTs). An effect can here be defined as the difference between the response of the patient after getting a treatment and what would have happened if the patient did not receive any treatment, if the patient instead had received a placebo (Whitlock et al., 2019). The use of the indiscriminate and ineffective treatments described above sometimes led to spontaneous remissions of symptoms and diseases, that were then erroneously interpreted as an effect - as the result of the medical treatment (placebo) being used (Benedetti, 2014).

If, when using RCTs, we assume that treatment effects and placebo effects are independent of each other then treatment effect is defined as the difference between the treatment and placebo response that the participants of the RCT experience. A high placebo response could influence the treatment effect and the conclusions scientists can draw about the efficacy of the treatment investigated (Whitlock et al., 2019). The general assumption is that the placebo response rates in the treatment arm and the placebo arm are equal, something called the additive model (Enck & Klosterhalfen, 2013). In a study designed with two treatment arms however, the placebo response could be different in the placebo arm and the treatment arm (Whitlock et al., 2019), something that some novel findings argue for (Enck & Klosterhalfen, 2013). If the placebo response is higher in the placebo arm, this could lead to the conclusion that the observed treatment effect is lower than the true treatment effect. A high placebo response could also impact the treatment effect if it produces floor- or ceiling effects - that the placebo response is so high that it limits the "window" in which the researcher can observe the treatment effects. Most well-designed RCTs reduce this potential risk by including baseline assessments (Whitlock et al., 2019). A high placebo response and clinical improvement from the use of the ineffective "medicines" of the past was likely a result of the patient's expectations of the "medicine" being beneficial or to changes in their emotional state. We can see remnants of the use of this kind of placebo treatment today in the category of alternative and/or complementary medicine where many people believe in the effectiveness of these treatments despite them not having gone through modern scientific scrutiny of efficacy. Outside of mainstream medicine, the use of this kind of "medication" (placebo) as a form of physiological, psychological and pharmacological treatment is still pervasive (Benedetti, 2014).

In research, the process of blinding is the action taken to ensure that the study participants, the researchers, the people providing the intervention to the participants, those who collect and those who analyze the data are kept unaware of the group assignment of the participants (Forbes, 2013). Successful blinding is important in RCTs as a lack of blinding has been shown to be associated with an exaggeration in the estimation of intervention effects (Pildal et al., 2007). Random and concealed allocation to the comparison groups in an RCT removes selection bias when participants enter a trial. It is also something that reduces information bias and can reduce biased supplemental care of different trial participants (Schulz et al., 2002).

From a patient perspective, the observed or felt effect following the administration of a placebo can be a result of several things. Patients may experience improvement as a result of psychophysiological response to an inert treatment, spontaneous remission, the natural cause of a disease, regression to the mean, biases or co-interventions. Regardless of origin, these changes are called a placebo response in the scientific community, and in the context of RCTs in particular (Benedetti, 2014). Some researchers discriminate, according to Ernst and Resch's (1995) conceptualization, between the psychophysiological placebo responses and the other responses (spontaneous remission, the natural cause of a disease, regression to the mean, biases or co-interventions) and call the former a true placebo and the latter as a perceived placebo. Most responses to "medications" of the past and to alternative and/or complementary medicines of today would most logically be termed as perceived placebos.

Many clinical researchers are interested in showing that a certain treatment is superior compared to the placebo. They are seldom interested in the underlying mechanisms of why a patient or group of patients improves in the placebo group. Rather, they are interested in whether the active treatment works better than the placebo. Recent neurobiological discoveries about true psychobiological placebos are however making scientists aware of and keen on understanding and differentiating true from perceived placebo phenomenon that are a part of clinical trials (Benedetti, 2014). Connected to this newly sparked interest has been the growing debate about the ethics of the administration of placebos in clinical trials in general, and in trials with psychiatric patients in particular. Proponents of the use of placebos in clinical trials argue that their use is critical for maintaining methodological rigor and therefore scientifically critical, especially if the placebo is not contributing to irreversible morbidity or increased mortality rates. On the other hand, if the efficacy of a certain treatment for a particular illness has been shown at least once to be good, those who object the use of placebos in psychiatric clinical trials argue that it is no longer ethical to assign participants to interventions that are expected to be less effective. They suggest using active comparators instead, interventions where efficacy has already been established to be good (Walsh et al., 2002).

An active placebo is a substance that produces noticeable effects that may make the participant receiving it and/or the practitioner administering it believe that they are being treated by a real medication, instead of by an ineffective substance (Farlex Partner Medical Dictionary, 2012). Morphine is a painkiller that often produces dizziness and sleepiness. In a study by Gilron et al. (2005) the medication lorazepam was used as an active placebo. Though it is not a

painkiller, it too causes sleepiness. It could therefore produce the same noticeable effects as the morphine that was being studied, potentially making the participants in the placebo group think that they were receiving the active treatment of morphine. In this way the researchers created a blinding of the treatment intervention from the placebo intervention. In contrast, an inactive placebo is a substance that is inert and has no direct physiological effect (Benedetti, 2014). An example of an inert placebo is giving participants in the placebo group a capsule containing micro-crystalline cellulose while the treatment group receives a capsule of the same look and size but containing the active treatment medication. The inert placebo in this case is the micro-crystalline cellulose (Carhart-Harris et al., 2021).

If researchers are to successfully and with a comfortable degree of certainty be able to express findings from studies of psychedelic therapies, blinding is something that ought to be of high priority for them. Blinding will assist these researchers in being able to better discriminate between true and perceived placebo effects and in reducing biases. The certainty of findings would also benefit from blinding processes being clearly defined, evaluated and verified, for example by checking if participants and researchers could guess treatment allocation based on observed effects. Trustworthy blinding and placebo controls are particularly important in studies with psychiatric populations because it has been observed that treatment effect can be large even in the placebo group (Jones et al., 2021). It is also important to keep in mind that since subjective experiences of hallucinogenic substances can be intense, it can be difficult to design trustworthy placebo controls.

Importance of this systematic review

Studies from the 1950's to 1970's that investigated the effects of psychedelics in the treatment of various psychiatric illnesses showed promising results (Grinspoon & Bakalar, 1979). Serotonergic hallucinogens in particular were studied at great length in psychedelic substance-assisted psychotherapy. These studies showed patient's symptoms being alleviated for several psychological issues including obsessive-compulsive disorder (OCD), substance dependence, substance abuse disorder, alcohol use disorder (AUD), PTSD and depression and anxiety in the terminally ill (Bogenschutz & Ross, 2018; Vargas et al., 2021). The research produced more than 1,000 scientific articles, papers and reports with the involvement of approximately 40,000 study participants (Vollenweider & Preller, 2020), however, much of the early research on psychedelics lacked methodological rigor. Ethical considerations were not

always incorporated into experimental design as laws for conducting human research were not subject to the same strict regulation as they currently are (Wheeler & Dyer, 2020).

The advent of the 1960's saw the arrival of the counterculture movement in the United States. The use of psychedelic substances became fashionable with certain parts of the population. As a result, through the Controlled Substance Act of 1970, LSD and other psychedelic drugs became classified as Schedule 1 type substances. This meant not only that the use of psychedelic drugs by the private citizen became illegal, but also that research on humans with psychedelics became extremely restricted (Ladewig & Pletscher, 1994). These changes in legislation brought about a demand for the inclusion of placebo and control groups for new studies investigating the therapeutic effects of psychedelics, something that up until this point in time were lacking or applied inconsistently (Johnson et al., 2008). This first wave of psychedelic research was therefore halted by political and methodological backlash and was in effect stopped until the 1990's (Wheeler & Dyer, 2020). Later findings from systematic reviews of the 1950's - 1970's studies of the therapeutic uses of psychedelics have concluded that patients showed improvement in symptoms in a wide range of psychiatric disorders (Vollenweider & Preller, 2020) despite the fact that placebo and control groups were not strictly implemented as part of the research process (Johnson et al., 2008). An evaluation of the methodological rigor in recent studies looking at the potential benefits of psychedelics for the treatment of psychiatric illnesses is therefore an important step forward to potentially re-open a larger scale scientific study of these substances and their medical use for psychiatric patients.

Statement of objectives and questions addressed

In light of the theoretical background described above, the aim of this systematic literature review was therefore to investigate if and to what degree psychedelics influence symptom reduction for patients suffering from mental disorders. By looking only at studies including placebo-controls the review investigated if symptom reduction is primarily attributed to the active psychedelic substance or to the patient's subjective experience (through the placebo-effect), focusing specifically on treatment effect comparisons between active treatment arms and placebo arms, and evaluating the methodological validity of the placebo control with a specific focus on whether the blinding could be maintained. The questions that this review aimed to answer were thus: 1) Do psychedelic substances have an influence on symptom reduction for patients suffering from mental illness as compared to placebos? 2) Are placebo

controls valid in psychedelic trials, i.e., is the blinding maintained in these trials? 3) Is one kind of placebo control better than others at influencing symptom reductions for these patients?

Methods

This study implemented the use of the systematic literature review method. According to Fink (2005), a systematic literature review is a "systematic, explicit, and reproducible method for identifying, evaluating, and synthesizing the existing body of completed and recorded work produced by researchers, scholars, and practitioners". Reviews help place individual papers in the context of how they contribute to our understanding of the subject under review. If done well, they also describe how each paper being reviewed relates to the others and helps to identify new ways to interpret previous research. They are used for helping a researcher identify what has already been researched on a topic of interest and what remains to be explored in the future. In other words, systematic literature reviews can identify gaps in knowledge that can be or need to be filled. They can be helpful in identifying methods that are or are not appropriate for studying a particular topic of choice. They also assist in preventing researchers from designing new studies that have already been done by presenting an overview of recent studies (Booth & Sutton, 2012). Mulrow (1994) argues that the aim of a systematic literature review is to search for the 'whole truth' and not bits and pieces of it, something that in itself is a 'fundamentally scientific activity'. The scientific activity of the systematic literature review allows for a specific and reproducible method of identifying, selecting and evaluating all the studies of a certain quality that are relevant to a particular topic. The results from such a review can give an idea of the strengths and weaknesses of available evidence and of the quality of the studies used to present this evidence. This in turn can guide future research and aid healthcare practitioners such as psychologists in deciding how much confidence they can or should place in results presented in scientific papers on topics of value to their practice (Booth & Sutton, 2012). Systematic reviews follow a pre-defined, structured process that requires rigorous methods to make sure that results from them are meaningful and reliable (Munn et al., 2018). They are considered 'the pillars of evidence-based healthcare' (Munn et al., 2014) that guide the development of trustworthy clinical practice (Steinberg et al., 2011).

Protocol

A protocol was designed in accordance with the guidelines for systematic literature reviews from Lund University (Institutionen för psykologi, 2023) and the PRISMA 2020 Checklist (Haddaway et al., 2022).

Eligibility criteria

Inclusion criteria

Articles were included in this systematic literature review if they 1) had a publication date between 2010 and 2023 2) were published in English 3) were peer reviewed 4) studied a human population (male, female or both) 5) had a placebo control-group 6) studied the psychotherapeutic use of any of the following 5-HT_{2A} receptor-activating classic psychedelics: DMT, LSD, psilocybin, mescaline, ayahuasca or iboga 7) investigated efficacies of these substances (efficacies defined as changes in subjective or objective physiological or psychological measures) in adult psychiatric patient populations and 8) were a randomized clinical study or clinical trial in phase I, II, III or IV.

Exclusion criteria

Articles were excluded from this systematic literature review if they 1) had a title and/or abstract that indicated that any or all of the inclusion criteria would not be met 2) upon reading had content that confirmed that any or all of the inclusion criteria were not met 3) studied micro-dosing rather than full dose interventions 4) were a systematic review, meta-analysis or qualitative study 5) were so called "grey literature" (reports, theses, ongoing clinical trials, partial results from larger ongoing or finished studies, protocols, rationales) (Karolinska Institutet, 2022) and 6) did not clearly state that the study had been approved by an ethics committee.

Information sources

Databases

Six databases were used for the identification of relevant articles for this systematic literature review: PsycINFO, PubMed, MEDLINE, Embase, SocINDEX and Scopus. The databases were used between March 1 and March 14 of 2023. Two platforms supported several of the databases used; PsycINFO, MEDLINE and SocINDEX were hosted by EBSCOhost while Embase and Scopus were hosted by Elsevier.

PsycINFO (last searched March 2, 2023) is a database that contains abstracts from literature in the psychological, social, behavioral and health sciences. It includes journals, books, reviews and dissertations. Although it is not a full-text database, it does link to full text articles. PubMed (last searched March 3, 2023) is a database of citations for biomedical literature from MEDLINE, life science journals and online books that provides free articles from open access journals. MEDLINE (last searched March 8, 2023) is a database that provides

information among other things on medicine and the health care system. It uses MeSH (Medical Subject Headings) to search citations from a large number of worldwide journals. Embase (last searched March 13, 2023) is a database providing access to pharmacological and biomedical literature. SocINDEX (last searched March 13, 2023) is a sociology research database providing full-text, peer-reviewed sociology journals on subjects such as social psychology and substance abuse and other addictions. Scopus (last searched March 14, 2023) is the largest abstract and citation database of peer-reviewed literature of articles in press.

Search strategy

The search strategy used for the literature review was first developed based on relevant literature on the broad subject matter 'psychedelics' and 'therapy'. Final search terms (psychedelic, therapy, psychiatric illness, placebo) were chosen based on the central terms relevant to the research questions. Synonyms to the central term 'psychedelic' and its sub-components (DMT, LSD, psilocybin, mescaline, ayahuasca, iboga) were found by using the internet sites drugbank.com (OMx, 2023), drugs.com (The Drugs.com Database, 2023) and wikipedia.com (Wikipedia, 2023).

Search terms related to the term psychedelic, its sub-components and their synonyms were combined in blocks using the operation "OR" creating seven blocks. Block eight used the operation "OR" to combine terms related to 'therapy' and 'psychiatric illness' while block nine used the operation "OR" to combine terms related to the subject of 'placebo'. Citation marks were used when a search term included more than one word and when a search term used chemical nomenclature. The three main blocks (psychedelic, therapy/illness, placebo) were then combined into one search block using the operation "AND", creating a total of seven searches per database. See *Appendix A-C* for information on the blocks, block combinations and the block searches, including filters and limits used for each database.

Selection process

Based on the above-described search strategy nine articles were chosen that met the inclusion and exclusion criteria and were relevant to the research question. These nine articles were used for the literature review.

The initial search produced a total of 1,784 articles, including duplicates (PubMed includes articles found through MEDLINE). After a first review of titles, a large number of articles were excluded (n = 1,294) (duplicates not excluded). Reading the remaining 490 (including duplicates) articles' abstracts and methods sections excluded a further 481 articles

due to the studies either having no placebo control and / or studying the incorrect participant group as per the inclusion and exclusion criteria or due to being a duplicate article. The final number of viable articles was nine (no duplicates).

Data collection process

The author of this review was the only reviewer that collected all data from each article included in the systematic literature review.

Data items

Outcomes

Nine articles were chosen as they fulfilled the inclusion and exclusion criteria relevant to this review.

The outcome domains investigated in this literature review were 'type of psychedelic substance used', 'type of symptoms being measured', 'kind of psychiatric illness' and 'form of placebo' being used in each. The time frame within which data was sought corresponded to the inclusion and exclusion criteria of articles published, that is between the years of 2010 and 2023. The main area of investigation was to look at the effects of treatment, especially in relation to placebo. This was done by evaluating and comparing changes in symptomatology and the size of effect measures used to quantify these. Furthermore, the integrity of the placebo intervention and the blinding procedure was of interest, something that was investigated through looking at author's descriptions of if and how blinding procedures were incorporated into their studies, if controls of the blinding process were performed and what they showed.

Articles included in the review were to investigate the psychotherapeutic use of any of the following 5-HT_{2A} receptor-activating classic psychedelics: DMT, LSD, psilocybin, mescaline, ayahuasca or iboga (outcome domain 'psychedelic substance'). At the same time, they were to investigate the outcome domain 'symptoms' through changes in subjective symptoms, physiological measurements or both. Subjective symptoms were defined as all the information gathered from participants through questionnaires or self-report inventories such as The Beck Depression Inventory (BDI) (Grob et al., 2011) and the participants verbal or written accounts of the "drug effect". Physiological measurements were defined as recorded parameters such as neuroimaging, blood analysis, blood pressure or heart rate. The outcome domain 'psychiatric illness' was restricted to any illness referred to in the DSM-5. The outcome domain pertaining to 'placebo' was to differentiate between the different possible kinds of controls and placebos used in scientific studies, for example single-blind and double-blind

control design, active and inactive placebos, as explained by Benedetti (2014). All results that were compatible with each of the outcome domains 'psychedelic substance', 'symptoms', 'psychiatric illness' and 'placebo' in each study were sought.

The articles were divided into three main themes based on the outcome domain 'symptoms': (1) subjective, (2) physiological or (3) subjective + physiological. These three main themes each had one of two sub-themes based on the outcome domain 'psychiatric illness': (a) ill (studies including participants with a psychiatric illness) or (b) ill + healthy (studies including participants both with a psychiatric illness and with healthy controls). A summary of possible theme and sub-theme interactions can be seen in Table 1 below.

Table 1

Possible outcome themes in articles chosen for review

	Subjective	**Physiological**	**Subjective/Physiological**
Ill	Ill & Subjective	Ill & Physiological	Ill & Subjective/Physiological
Ill/Healthy	Ill/Healthy & Subjective	Ill/Healthy & Physiological	Ill/Healthy & Subjective/Physiological

The outcome domain 'psychedelic substance' and 'placebo' were considered most important for the interpretation of the review's conclusions. They related directly to the research questions of evaluating the methodological rigor of psychedelic studies through looking at the kinds of placebos used.

Other variables

Other variables of interest for this literature review were participant's age, sex, exposure-level to psychedelic substances and the form of substance administration. The study included male and female adults aged 18 or above. Both participants that were naïve to psychedelics and those who were non-naïve were of interest for the review. However, only studies investigating full-dose exposure were included, as opposed to micro-dosing ones.

Effect measures

The outcome domains investigated in this review were 'type of psychedelic substance used', 'type of symptoms being measured', 'kind of psychiatric illness' and 'form of placebo' being used. Of these outcome domains, changes in symptom levels and placebo effect were the

only domains where effect measures were applicable. 'Type of psychedelic used' and 'kind of psychiatric illness' were measured in the form "Is it present?" and/or "What kind?".

Domains where effect measures were applicable looked at the measures used by the authors of the articles chosen for the review, in this case being mean differences, correlations (*r*) and odds ratios (*OR*).

Size of effect was also evaluated by using the measures provided by the study authors, in this case Cohen's *d* (*d*), Hedges' *g* (*g*) and General Eta Squared (η^2_G) thresholds. Both *d* and *g* effect sizes were assessed as follows: 0.20 = small, 0.50 = medium, 0.80 = large. *g* is a variation of *d* that corrects for biases due to small sample sizes. *d* is used for sample sizes of 20 or more while *g* is used for sample sizes of 20 or less (Howitt & Cramer, 2005a). η^2_G (Sánchez & Cervantes, 2016) is a correlation ratio most often used for between-group designs of more complex nature, with more than one factor. Effect sizes using η^2_G were assessed as follows: 0.01 = small, 0.06 = medium, 0.14 = large (Howitt & Cramer, 2005a).

Synthesis methods

Eligibility for synthesis

Studies that were included in this review were grouped first according to the type of 5-HT_{2A} receptor-activating classic psychedelic used in the study and then, per psychedelic type, according to the theme and sub-theme interactions described in Table 1. Based on this, studies were eligible for synthesis in the following manner: 1) type of psychedelic and how it contributes to changes in symptomology 2) placebo kind and how it contributes to changes in symptomology 3) subjective symptoms and their role and 4) physiological measures and their role.

Tabulation and graphical methods

Tabular structures used to display the results of individual studies and synthesis were one figure showing the PRISMA flow chart and six tables (Table 1 – 6) to summarize information on: outcome themes, presentation of articles, interventions, types of placebos, placebo blinding assessment, effect measures, effect sizes and information of heterogeneity. Articles in the tables are presented alphabetically according to primary authors' last name.

Methods to explore heterogeneity

The heterogeneity of the included studies was assessed theoretically and described. The methodological implications of this assessment were later discussed.

Results

Characteristics of contributing studies

Nine articles that met the inclusion and exclusion criteria and were relevant to the research questions were used for the literature review. All these articles studied participants with some form of psychiatric illness. None included healthy control subjects. Subjective parameters were the focus of two articles, physiological parameters only in none, while seven articles looked at both subjective and physiological measures. The psychiatric illnesses represented in the included articles were major depressive disorder (n = 2), treatment-resistant depression (n = 1), depression (n =2), anxiety (n = 5) and alcohol use disorder (n = 1), with two studies looking at both depression and anxiety. Four articles used an active placebo, three an inactive and in two the participants acted as their own controls. All studies used a double-blind design.

Six studies compared the psychedelic substance psilocybin with placebo in participants with a psychiatric illness (Bogenschutz et el., 2022; Carhart-Harris et al., 2021; Griffiths et al., 2016; Grob et al., 2011; Ross et al., 2016; von Rotz et el., 2023). One article (Carhart-Harris et al., 2021) looked at changes in participant's subjective symptoms while the rest looked both at subjective and physiological measures. Three of these studies used an active placebo control (Bogenschutz et el., 2022; Griffiths et al., 2016; Ross et al., 2016), two used an inactive placebo control (Carhart-Harris et al., 2021; von Rotz et el., 2023) while participants acted as their own controls in the last article (Grob et al., 2011).

One study compared the psychedelic substance Ayahuasca with placebo in participants with a psychiatric illness, looking at participants' subjective symptoms (Palhano-Fontes et al., 2019). This study used an inactive placebo.

Two studies compared the psychedelic substance LSD with placebo in participants with a psychiatric illness (Gasser et al., 2014; Holze et al., 2023). Both studies looked at participant's subjective symptoms and physiological measures. Gasser et al. (2014) used an active placebo while Holze et al. (2023) used the study participants as their own controls.

Iboga, DMT and mescaline were not represented in the articles chosen for this review.

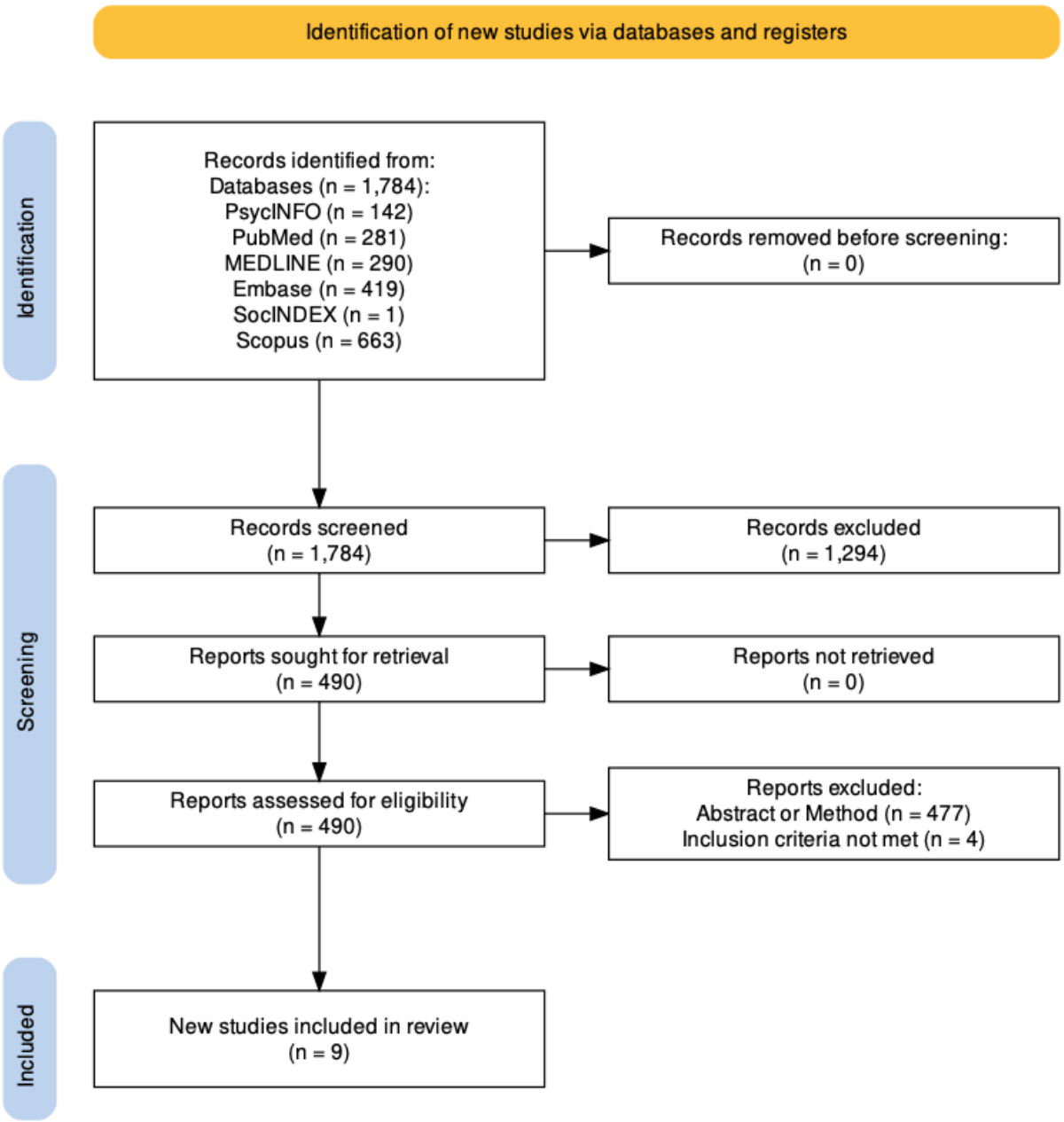
Study selection

Flow of studies

Figure 1 below shows a PRISMA flow chart (Haddaway et al., 2022) describing the selection process for the articles chosen for this systematic literature review.

Figure 1

PRISMA flow chart



Excluded studies

1294 studies were excluded in the first phase of screening due to a title indicating that inclusion and exclusion criteria would not be met. A further 481 articles were later excluded during two additional screenings based on information presented in abstracts or methods sections and the discovery of an incorrect participant group for the purposes of this review. Of a total of 1784 articles first identified, 72.5% (1294) were excluded due to title and 26.9% of the remaining articles were excluded due to abstract, method and participant groups not matching inclusion criteria (481).

Study characteristics

Table 2 below is a presentation of the articles included in this systematic literature review with regards to authors, dates of publication, populations, outcomes, interventions and main results.

Table 2

Presentation of articles

Author	Population	Outcome	Intervention	Main results
Bogenschutz et al. (2022)	Patients with alcohol use disorder (AUD).	*Main*: PHDD. *Secondary*: BP, HR, MEQ, PDD, DPD, abstinence, lack of PHDD, WHO risk level, SIP-2R.	Session 1: psilocybin 1 x 25mg/70kg or placebo (diphenhydramine) 1 x 50mg. Session 2: psilocybin 1 x 30mg/70kg or 25mg/70kg or placebo (diphenhydramine) 1 x 100mg.	*Main*: Psilocybin group had fewer PHDD than diphenhydramine group. *Secondary*: Psilocybin participants had a high while the diphenhydramine participants had a low average intensity of experience (MEQ). Psilocybin participants were more likely to have no PHDD and WHO risk level reduction. At follow-up, differences persisted.
Carhart-Harris et al. (2021)	Patients with moderate-to-severe major depressive disorder (MDD).	*Main*: QIDS-SR-16. *Secondary*: QIDS-SR-16 response, QIDS-SR-16 remission, QIDS-SR-14, BDI-1A, HAM-D-17, MADRS, FS, STAI, BEAQ, WSAS, SHAPS, WEMWBS, SIDAS, PRSexDQ, LEIS, PTCS.	Dosing day 1 & 2: psilocybin group 1 x 25mg psilocybin, escitalopram group 1 x 1 mg psilocybin. Between dosing day 1 & 2: psilocybin group 1 x 21 days placebo (micro-crystalline cellulose), escitalopram group 1 x 21 days 10mg escitalopram.	*Main*: No significant difference between the trial groups. *Secondary*: Secondary outcomes generally favored psilocybin, but corrections for multiple comparisons were not performed limiting conclusions that can be drawn from these data.

			After dosing day 2: psilocybin group 2 x 21 days placebo (micro-crystalline cellulose), escitalopram group 1 x 21 days 20mg escitalopram.	
Gasser et al. (2014)	Patients with anxiety and life-threatening disease.	*Main*: STAI T, STAI S. *Secondary*: EORTC-QLQ-30, SCL-90-R, HADS, HR and BP and AEs.	Dosing session 1: LSD 1 x 200µg or 1 x placebo (20µg LSD). Dosing session 2: LSD 1 x 200µg or 1 x placebo (20µg LSD). Open-label crossover: Non-mandatory LSD 1 x 200µg for previous placebo (20µg LSD) participants.	*Main*: STAI T - Three of the eight participants in the LSD group experienced reductions in anxiety. All participants in the active placebo group experienced increases in anxiety. For the LSD group, reductions in anxiety were maintained over time. STAI S - Three of the eight participants in the LSD group experienced reductions in anxiety. Two participants in the active placebo group experienced increases in state anxiety. In the LSD group, reductions in anxiety were maintained over time. *Secondary*: Were not analyzed for statistical significance but the secondary outcome results supported the results of the primary outcome measures.
Griffiths et al. (2016)	Patients with symptoms of depression and/or anxiety and life-threatening cancer.	*Main*: GRID-HAM-D-17, HAM-A. *Secondary*: BP, HR, subjective drug effect measures (HRS, 5D-ASC, Mysticism Scale, SOCQ, MEQ30), psychiatric symptoms, moods, attitudes (BDI, HADS, STAI, POMS, BSI, MQOL, LOT-R, LAP-R Death Acceptance, Death Transcendence Scale,	Session 1: psilocybin 1 x 1mg/70kg (LDF/ placebo) or psilocybin 1 x 22mg/70kg (HDF). Crossover. Session 2: psilocybin 1 x 1mg/70kg (LDS/ placebo) or psilocybin 1 x 22mg/70kg (HDS).	*Main*: Overall large, statistically significant clinical response and symptom reduction for both anxiety and depression with larger changes in the HDF group. *Secondary*: Psilocybin produced large, statistically significant and sustained effects on most measures at baseline, five weeks after each session and at

		Purpose of Life Test, LAP-R Coherence).		the six-month follow-up.
Grob et al. (2011)	Patients with reactive anxiety and advanced-stage cancer.	*Main*: BDI, POMS, STAI, 5D-ASC, Brief Psychiatric Rating Scale. *Secondary*: HR, BP, temperature.	Session 1: psilocybin 1 x 0.2mg/kg or placebo (niacin) 1 x 250mg. Crossover. Session 2: psilocybin 1 x 0.2mg/kg or placebo (niacin) 1 x 250mg.	*Main*: Patients had statistically significant differences in subjective experience (5D-ASC), BDI scores dropped and were significantly different at the six-month follow-up, the psilocybin produced reduced adverse mood scores (POMS), patient's anxiety decreased at 1- and 3-month check-ins (STAI). *Secondary*: Statistically significant but mild and transient elevations in HR and BP
Holze et al. (2023)	Patients with anxiety with or without a life-threatening illness	*Main*: STAI-G. *Secondary*: STAI-G, HAM-D-21, BDI, SCL-90-R, BP, HR, 5D-ASC, MEQ30, AEs and SAEs.	Session 1 & 2: LSD 1 x 200µg or placebo (ethanol). Crossover. Session 2 & 4: LSD 1 x 200µg or placebo (ethanol).	*Main*: STAI-G scores at 16 weeks after the first dosing session were significantly different between the treatment groups, with the LSD group showing significant reductions in anxiety. *Secondary*: Patients in the LSD groups experienced a significant reduction in anxiety, depression and general psychiatric symptomatology compared with the placebo group.
Palhano-Fontes et al. (2019)	Patients with treatment resistant depression (TRD).	*Main*: HAM-D. *Secondary*: MADRS, response rates, remission rates, safety and tolerability (CADS, BPRS, YMRS), specific aspects of the psychedelic effect (HRS, MEQ30).	Dosing day: Ayahuasca 1 x 1ml/kg (0.36mg/kg of *N*,*N*-DMT) or placebo 1 x 1ml/kg.	*Main*: Large and significant differences between Ayahuasca and placebo group with Ayahuasca group showing significant reduction in anxiety symptom severity (HAM-D). *Secondary*: A large and significant effect of time and treatment on depression symptom severity - a significant decrease after one day

				for Ayahuasca group, which persisted over time (MADRS). Significant remission rates for Ayahuasca group for anxiety and depression symptoms on day 7. Trends towards significance on remission rates.
Ross et al. (2016)	Patients with anxiety and depression and life-threatening cancer.	*Main*: HADS A, HADS D, HADS T, BDI, STAI S, STAI T. *Secondary*: BP, HR, existential distress, quality of life and spirituality, immediate and sustained effects on subjective experience, cognition, affect and behavior, adverse events.	Session 1: psilocybin 1 x 0.3mg/kg or placebo (niacin) 1 x 250mg. Crossover. Session 2: psilocybin 1 x 0.3mg/kg or placebo (niacin) 1 x 250mg.	*Main*: Significant differences between the psilocybin and niacin groups prior to the crossover, with the psilocybin group showing immediate, substantial and sustained clinical benefits in terms of reduction of anxiety and depression symptoms. *Secondary*: Psilocybin decreased cancer-related demoralization and hopelessness, while improving participants' spiritual well-being, general life satisfaction and quality of life.
von Rotz et al. (2023)	Patients with major depressive disorder (MDD).	*Main*: MADRS, BDI. *Secondary*: Somatic and psychiatric symptoms measured through BP, pulse, SCL-90-R, HAM-A, CGI, C-SSRS, ASC, adverse events.	Dosing day: psilocybin 1 x 0.215mg/kg or placebo (mannitol).	*Main*: A main effect of psilocybin, with significantly larger symptom severity reduction compared to placebo. Mean differences in depressive symptoms highest two days after drug administration. *Secondary*: Significant effects of psilocybin on anxiety, depression, psychoticism, phonic anxiety, paranoid ideation and global severity index. The intensity of subjective effects induced by psilocybin did not predict the positive outcome in MDD.

In Table 3 below is a presentation of the articles included in this systematic literature review with regards to outcome theme, type of psychedelic, type of psychiatric illness and type of placebo control used.

Table 3

Summary of interventions and types of placebos used in studies

Author	Outcome theme	Type of psychedelic	Type of illness	Type of placebo
Bogenschutz et al. (2022)	Ill & Subjective/ Physiological	Psilocybin	AUD	Double-blind. Placebo: active control (diphenhydramine)
Carhart-Harris et al. (2021)	Ill & Subjective	Psilocybin	MDD	Double-blind. Placebo: inactive control (micro-crystalline cellulose)
Gasser et al. (2014)	Ill & Subjective/ Physiological	LSD	Anxiety (with life-threatening disease)	Double-blind. Placebo: active control (low dose LSD)
Griffiths et al. (2016)	Ill & Subjective/ Physiological	Psilocybin	Depression and anxiety (with life-threatening cancer)	Double-blind. Placebo: active (low dose psilocybin)
Grob et al. (2011)	Ill & Subjective/ Physiological	Psilocybin	Reactive anxiety (with advanced-stage cancer)	Double-blind. Placebo: patients acting as their own control
Holze et al. (2023)	Ill & Subjective/ Physiological	LSD	Anxiety (with and without life-threatening illness)	Double-blind. Placebo: patients acting as their own control
Palhano-Fontes et al. (2019)	Ill & Subjective	Ayahuasca	TRD	Double-blind. Placebo: inactive (liquid with organoleptic properties of Ayahuasca)
Ross et al. (2016)	Ill & Subjective/ Physiological	Psilocybin	Anxiety and depression (with life-threatening cancer)	Double-blind. Placebo: active control (niacin)
von Rotz et al. (2023)	Ill & Subjective/ Physiological	Psilocybin	MDD	Double-blind. Placebo: inactive (mannitol)

Two of the articles investigated changes in just subjective symptoms, seven looked at changes in both subjective and physiological measures while none looked only at physiological changes. Of the nine articles, all studies included people with a psychiatric illness only, while

none included participants that had a psychiatric illness and / or healthy controls. All studies used both male and female participants.

Six articles studied the classic psychedelic psilocybin, one looked at Ayahuasca and two at LSD.

Articles investigated psychedelics in relation to several psychiatric illnesses: major depressive disorder (n = 2), treatment-resistant depression (n = 1), depression (n = 2), anxiety (n = 5) and alcohol use disorder (n = 1), with two studies looking at both depression and anxiety.

Four articles used an active placebo, three articles used an inactive placebo and in two articles the participants acted as their own controls. All nine studies implemented a double-blind design.

Articles that looked only at subjective drug effects in participants used the outcomes measures (see *Appendix D* for explanations of abbreviations): QIDS-SR-16, QIDS-SR-14, BDI-1A, HAM-D, HAM-D-17, MADRS, FS, STAI, BEAQ, WSAS, SHAPS, WEMWBS, SIDAS, PRSexDQ, LEIS, PTCS, CADS, BPRS, YMRS, HRS, MEQ30 and the Emotional Breakthrough Inventory. Other subjective symptom outcome measures that were used were STAI-T, STAI-S, STAI-G, EORTC-QLQ-30, SCL-90-R, HADS, BDI, HAM-A, CGI, C-SSRS, ASC, AEs, HADS A, HADS D, HADS T, HAI, DAS, DTS, WHO-Bref, FACIT-SWB, MEQ retrospective, PEQ, PHDD, PDD, DPD, WHO risk level, SIP-2R, SOCQ, Monitor Rating Questionnaire, 5D-ASC, Mysticism Scale, GRID-HAM-D-17, Community Observer Interview, POMS, POMS Brief, BSI, MQOL, LOT-R, LAP-R Death Acceptance, FACIT-Sp, LAP-R Coherence, Spiritual Religious Outcome Scale, Faith Maturity Scale, Persisting Effects Questionnaire, Visual Analog Pain Scale, Brief Psychiatric Rating Scale, HAM-D-21, SCL-90-R, and SAEs.

Articles that investigated changes in physiological measures quantified these using BP, HR, temperature.

Description of individual studies

Classic psychedelics that contribute to changes in symptomology

Psilocybin

Grob et al. (2011) used a within-subject, double-blind, placebo-controlled study to examine the safety and efficacy of psilocybin in the treatment of psychological distress associated with the existential crisis of terminal disease. Twelve patients with advanced-stage cancer and reactive anxiety (a diagnosis of acute stress disorder, generalized anxiety disorder,

anxiety disorder due to cancer or adjustment disorder with anxiety) were recruited for the study where each subject acted as his or her own control. Participants were informed that they would receive active psilocybin (0.2mg/kg) on one occasion and the placebo niacin (250mg) on the other occasion, with several weeks between dosing sessions. The order in which the subjects received the two treatments was randomized and known only to the research pharmacist. Thirty minutes before the dosing session started, immediately before drug ingestion and at hourly intervals during the session the patient's HR and BP were monitored. Just prior to drug ingestion and six hours later, participants' temperature was also measured.

The researchers used several psychological questionnaires as their primary outcome measure: BDI, POMS, STAI, Brief Psychiatric Rating Scale and 5D-ASC. BDI, POMS and STAI were administered the day before each dosing session. POMS, STAI, 5D-ASC and Brief Psychiatric Rating Scale were administered at the end of the dosing sessions. The day after each dosing session and again two weeks after the final dosing session the patients were asked to fill out BDI, POMS and STAI.

Researchers observed that the administration of psilocybin (0.2mg/kg), compared with the niacin placebo (250mg), produced a mild but statistically significant elevation in HR, systolic BP and diastolic BP, although these changes peaked two hours after psilocybin administration, were transient and did not lead to any sustained adverse effects such as tachyarrhythmias. When comparing participants' psilocybin with placebo experiences, the 5D-ASC showed statistically significant subjective differences between the two sessions. Psilocybin affected the oceanic boundlessness and visionary reconstructuralization dimensions most profoundly. It also had significant but smaller effects on anxious ego dissolution and auditory alterations. There was an overall interaction of psilocybin and day that approached but did not attain statistical significance for the BDI. BDI scores dropped by almost 30% from the first session to one month after the second session, a drop that was sustained and became significant at the six-month follow-up. The POMS showed a trend of reduced adverse mood from day one before treatment to two weeks later for psilocybin, a difference that was not observed for the placebo. Post hoc analysis revealed that mean scores were elevated one day before psilocybin treatment compared with one day before placebo treatment and that this difference disappeared six hours after psilocybin administration. The elevation of POMS scores one day before psilocybin treatment occurred regardless of whether the participant was treated with placebo or psilocybin first – there was no interaction between treatment order and drug.

The STAI showed no significant changes from day one before to two weeks after treatment sessions. There was however a substantial but nonsignificant decrease for the state anxiety subscale six hours after psilocybin ingestion, something that was not observed after placebo ingestion. For the entire six-month follow-up period there was an observed sustained decrease in STAI trait anxiety which reached significance at the one-month and three-month point after the second treatment session. The Brief Psychiatric Rating Scale showed no differences between psilocybin and placebo sessions.

This study did not include effect size measures and although it used a double-blind and placebo-controlled design, the drug order was almost always apparent to the participants and investigators. Despite these limitations, results of this study show that the controlled use of psilocybin may provide an alternative treatment model for conditions such as anxiety that often accompany advanced-stage cancers and that are minimally responsive to conventional psychotherapies.

Griffiths et al. (2016) used a double-blind crossover design to study the effects of psilocybin in participants with a potentially life-threatening cancer diagnosis and symptoms of depression and/or anxiety. The researchers investigated the effects of a very low, placebo-like dose (1 or 3mg/70kg) vs. a high dose (22 or 30mg/70kg) of psilocybin on clinician and self-rated measures of depression and anxiety in 56 participants. The study was done in counterbalanced sequence with five weeks between sessions.

The Low-Dose First group (LDF) received the low dose of psilocybin (1mg/70kg) on the first session and the high dose (22mg/70kg) on the second session while the High-Dose First group (HDF) received the high dose (22mg/70kg) on the first session and the low dose (1mg/70kg) on the second session. The high dose was decreased from 30mg to 22mg after two of the first three participants to receive a high dose of 30mg/70kg were discontinued (vomiting, personal reasons). The low dose was decreased from 3mg to 1mg after twelve participants because data from a same dose-effect study showed significant psilocybin effects at 5mg/70kg, which raised concern that 3mg/70kg might not serve as an inactive placebo.

The primary outcome measures that the researchers used were GRID-HAMD-17 and HAM-A. There were fifteen secondary outcome measures that focused on self-rated measures of psychiatric symptoms, moods and attitudes. Participant's BP and HR were assessed during the dosing sessions and the session monitors filled out the Monitor Rating Questionnaire which rated the participant's behavior or mood or several dimensions. The participants completed four

subjective drug effect measures questionnaires seven hours after psilocybin sessions. The research team also conducted structured telephone interviews with community observers (family members, friends, work colleagues) to collect ratings of participant attitudes and behavior reflecting healthy psychosocial functioning.

The researchers used t-tests to compare participants that received 3mg/70kg (n = 12) with those that received 1mg/70kg (n = 38) and found no significant differences in HR and BP measures in participants – data for the 1 and 3mg/70kg doses were combined in the low-dose condition for all analyses. One participant received 30mg/70kg and 49 received 22mg/70kg. Analyses were done for seventeen measures with and without the 30mg/70kg participant, finding no differences in significance - this participant's data were included in the final analyses.

The researchers found that psilocybin produced large, statistically significant and sustained effects on the two primary outcome measures and most of the secondary measures assessed at baseline, five weeks after each session and at the six-month follow-up. At baseline, the mean (SD) measures for the GRID-HAMD-17 for the LDF and HDF groups were 22.32 (0.88) vs. 22.84 (0.97) respectively. At the six-month follow-up, these measures were 6.95 (1.24) vs. 6.23 (1.30) respectively, with a statistically significant mean change in scores of −15.37 for the LDF group vs. −16.61 for the HDF group (Cohen's d = 2.98). At baseline, the mean (SD) measures for the HAM-A for the LDF and HDF groups were 25.68 (0.89) vs. 25.73 (1.11) respectively. At the six- month follow-up, the measures were 7.95 (1.19) vs. 7.04 (1.17) respectively, with a statistically significant mean change in scores of −17.73 for the LDF group vs. −18.69 for the HDF group (Cohen's d = 3.40). Patients experienced a significant clinical response and symptom reduction. Five weeks after session one, 92% of the HDF group showed a ≥ 50% decrease relative to baseline on the GRID-HAMD-17 compared with 32% response rate in the LDF group. At the six-month follow-up, 79% of the HDF group continued to show a clinically significant response. For the HAM-A, at five weeks these numbers were 76% and 24% for the HDF and LDF groups respectively and at the six-month follow-up the HDF group continued to show a clinically significant reduction in symptoms in 83% of participants. Results of symptom reduction to normal range (≥ 50% decrease relative to baseline or a score of ≥ 7 on the GRID-HAMD-17 or HAM-A) showed decreases of 60% and 52% for depression and

anxiety respectively five weeks after session one. At the six-month follow-up, these rates were 71% and 63% respectively. Across the two sessions and the two dose groups, at the six-month follow-up the overall clinical response was 78% and 83% for depression and anxiety respectively and the overall symptom remission was 65% and 57% respectively. The study demonstrated the efficacy of a high dose of psilocybin administered under supportive conditions to decrease symptoms of depression and anxiety and to increase quality of life in patients with a life-threatening cancer diagnosis.

The blinding procedures did not provide protection against a priori monitor expectancy determining outcomes of the psilocybin dose manipulation. The mean (±SD) monitor rating of the dose magnitude of the high psilocybin dose was significantly larger than for the low dose (7.0 ± 0.29 vs. 1.7 ± 0.21, $p < 0.001$), although the distributions of the ratings overlapped. More than 13% of the high dose sessions were rated as 4 or less while more than 12% of the low dose sessions were rated as 4 or more. Overall, the blinding procedure must be considered to have been unsuccessful.

Carhart-Harris et al. (2021) performed a phase two, double-blind, placebo-controlled trial involving patients (n = 59) with long-standing, moderate-to-severe MDD in order to compare the anti-depressant effect of psilocybin with escitalopram, a representative of the currently used SSRI anti-depressants.

At a first visit (baseline) patients received a preparatory therapeutic session and underwent fMRI and cognitive and affective processing tasks (not included in this study's results analysis). During visit two the patients in the psilocybin group received 25mg of psilocybin and the patients in the escitalopram group received 1mg of psilocybin. All were informed they were receiving psilocybin, but not at what dose. After visit two, all participants received a bottle of capsules and were instructed to take one capsule each morning until the next scheduled psilocybin session. The capsules contained either a placebo (micro-crystalline cellulose) - given to the participants who had received the 25mg dose on dosing day one, or 10mg of escitalopram - given to the participants who had received the 1mg dose of psilocybin on dosing day one. Three weeks after dosing day one the patients received their second dose of psilocybin (25mg) or placebo (1mg psilocybin). After this second dosing session the patients were asked to take two capsules each morning for the next three weeks, either placebo (the psilocybin group) or an increased dose of 20mg of escitalopram (the escitalopram group).

The primary outcome measure that the researchers looked at was the change from baseline in the score on the QIDS-SR-16 at six weeks. Secondary outcome measures were gathered but information from them had no p values reported and therefore no clinical conclusions can be drawn from these data.

Researchers found that the mean scores on the QIDS-SR-16 at baseline were 14.5 in the psilocybin group and 16.4 in the escitalopram group. The mean (± SD) change from baseline to week six was –8.0 (± 1.0) in the psilocybin group and –6.0 (± 1.0) in the escitalopram group. There was no significant difference between the trial groups. A reduction in score of > 50% (QIDS-SR-16 response) occurred in 70% of the patients in the psilocybin group and in 48% of the patients in the escitalopram group, with a between-group difference of 22%. A score of > 5 (QIDS-SR-16 remission) occurred in 57% and 28% respectively, with a between-group difference of 28%. As noted above however, no p-values are available for these secondary outcome measures.

The absence of a placebo group separate from the two drug interventions makes drawing conclusions about the effects of psilocybin or escitalopram alone impossible. The researchers also did not assess the effectiveness of the blinding within each treatment group.

Bogenschutz et el. (2022) ran a multisite randomized clinical trial to evaluate the efficacy of psilocybin-assisted psychotherapy for the treatment of AUD. 95 participants were randomly assigned in a 1:1 ratio to receive either psilocybin (n = 49) or the active placebo diphenhydramine (n = 46). Participants were assigned to receive medication during two day-long sessions at week four and eight after all being offered 12 weeks of manualized psychotherapy. The first session randomly assigned participants to psilocybin (25mg/70kg) or diphenhydramine (50mg). Participants received an increased dose in the second session if there were no dose-limiting adverse events and if they agreed to the increase. The second session was psilocybin (30mg/70kg) if the participant's total score on MEQ was 0.6 or greater in the first session, or psilocybin (25mg/70kg) if the MEQ score in the first session was less than 0.6. The increased dose of diphenhydramine was 100mg regardless of participants' subjective response. Researchers measured subjective effects of psilocybin vs. diphenhydramine using the States of Consciousness Questionnaire, which contains the 43-item MEQ. The primary drinking outcome was the percentage of heavy drinking days (PHDD) during weeks five to 32, assessed at week eight, 12, 24 and 36 using timeline followback. Secondary outcomes of interest included percentage of drinking days (PDD), mean drinks per day (DPD) and dichotomous outcomes

(abstinence, lack of heavy drinking days and reduction in WHO risk level by one, two or three levels). The SIP-2R was used to assess drinking-related problems at baseline and at weeks 12, 24 and 36. BP and HR were also assessed at 30- to 60-minute intervals during the first six hours of each medication session.

Of the 95 participants, 93 received at least one dose of medication: 48 received psilocybin (25mg/70kg) and 45 received diphenhydramine (50mg) in the first session. 89.6% of the participants treated with psilocybin (n = 43) and 77.8% of those treated with diphenhydramine (n = 35) received a second double-blind medication session.

Researchers found, using the MEQ scores, that during session one participants had a high average intensity of experiences in the psilocybin group and low average intensity experience in the diphenhydramine group. During session two, similar results emerged. The primary drinking outcome analysis showed a main effect of treatment. During weeks five to 36, participants who received psilocybin had lower PHDD than those who received diphenhydramine (mean difference, 13.86%; Hedges' $g = 0.52$; $p = 0.01$). Participants receiving psilocybin were also more likely than those receiving diphenhydramine to have no PHDD and to have a two-level reduction in WHO risk level during weeks five to 36. In conclusion, during the final month of follow-up it was seen that these differences persisted and that the rates of abstinence as well as one and three level reductions in WHO risk levels were higher in the psilocybin group than in the diphenhydramine group.

Except for the study pharmacist, all researchers and participants were blinded to the treatment assignment. After each medication session, the therapists and participants were asked to guess which medication that the participant had been given and to rate their degree of certainty on a 100-point visual analogue scale (0 = not at all confident, 100 = extremely confident). For the first dosing session, participants correctly guessed their assigned medication 93.6% of the time with a mean certainty of 88.5%. The therapist correctly guessed the assigned medication 92.4% of the time with a mean certainty of 92.8%. For the second dosing sessions the numbers for the participants and therapists were 94.7% and 97.5% respectively for guessing the correct assignment, with a mean certainty of 90.6% and 95.4% respectively. In conclusion, diphenhydramine was not effective in maintaining the blind after medication administration.

Ross et al. (2016) looked at the efficacy of psilocybin versus an active placebo (niacin), administered in conjunction with psychotherapy, in treating clinically significant anxiety or

depression in patients with life-threatening cancer. The trial used a two-session, double-blind, crossover design to compare groups.

Patients (n = 29) were randomly assigned to two oral dosing session sequences: first psilocybin (0.3mg/kg) then niacin (250mg) (n = 14), or niacin (250mg) first and then psilocybin (0.3mg/kg) (n = 15). Dose one administration occurred two to four weeks after baseline assessments, the crossover occurred seven weeks post dose one, at which point the administration of dose two occurred.

The primary outcome variables that the researchers were interested in were anxiety and depression, assessed prior to the crossover, using HADS, HADS A, HADS D, HADS T, BDI, STAI, STAI S and STAI T.

Secondary outcome variables were the assessment of existential distress, quality of life and spirituality, as well as measures assessing immediate and sustained effects of the psilocybin administration on subjective experience, cognition, affect and behavior. The researchers also monitored adverse events (AE) attributed to the study medications and cardiovascular measures (systolic and diastolic BP, HR) during medication sessions.

The researchers stated that the two dose-sequence groups did not significantly differ on demographic or clinical characteristic measures and that no dichotomous factors (gender, prior hallucinogen use vs. none, spiritual faith/religion vs. none, early vs. late cancer stage) significantly interacted with the primary outcome measures in between-group comparisons.

There were no serious medical or psychiatric AEs in the trial attributed to psilocybin or niacin. All non-clinically significant AEs attributable to psilocybin in the study are known AEs of psilocybin, were transient, tolerable and consistent with prior trials of psilocybin administration in normal volunteers and patients with terminal cancer.

What the researchers found when it comes to primary outcomes is that for each of these measures (HADS T, HADS A, HADS D, BDI, STAI S, STAI T) there were significant differences between the experimental and placebo groups prior to the crossover, with the psilocybin group (compared to the active niacin placebo) showing immediate, substantial and sustained clinical benefits in terms of reduction of anxiety and depression symptoms. Differences were measured at several pre-crossover time points (baseline, one day pre dose one, one day post dose one, two weeks post dose one, six weeks post dose one- and seven-weeks post dose one) with significance levels of $p < 0.05$, $p \leq 0.01$ and $p \leq 0.001$ of between-group

t-tests. Effect size for these changes was quantified using Cohen's *d*. The magnitude of the differences between the psilocybin group and the niacin control group was large, ranging from $d = 1.36$ to 1.69 (HADS T), from $d = 0.80$ to 1.18 (HADS A), from $d = 0.98$ to 1.32 (HADS D), from $d = 0.82$ to 1.10 (BDI), from $d = 1.18$ to 1.45 (STAI S) and from $d = 0.95$ to 1.49 (STAI). Treatment groups did not differ in magnitude of change across their respective psilocybin treatment sessions for any of the primary outcome measures. For all primary outcome measures, the 'psilocybin first' group demonstrated significant within-group reductions (compared to baseline at each post-baseline assessment point) in anxiety and depression immediately after receiving the psilocybin. The reductions remained significant at all time points. The 'niacin first' group showed either no significant within-group reductions or transient reduction that became non-significant prior to crossover. At the 6.5-month follow-up, after both groups received psilocybin, anti-depressant or anxiolytic response rates were between 60 – 80%.

In relation to the secondary outcomes, the researchers found that psilocybin decreased cancer-related demoralization and hopelessness, while improving participants' spiritual well-being, general life satisfaction and quality of life. Psilocybin was also associated with improved attitudes and adaptations to death at the 6.5-month follow-up, even though there were no improvements in affect/anxiety toward death either short-term or long-term. Participants described their psilocybin experience as highly meaningful and spiritual and associated with positive cognitive, affective, spiritual and behavioral effects lasting weeks to months.

In summary, a single moderate-dose psilocybin (compared to the active control niacin) was safely administered in a group of patients with cancer-related anxiety and/or depressive symptoms. It produced immediate and enduring anxiolytic and anti-depressant responses, as well as anti-depressant remission rates for at least seven weeks and up to eight months. However, the researchers state that one limitation of their study is the use of a control with limited blinding.

von Rotz et al. (2023) investigated the effect of a single moderate dose of psilocybin compared to placebo in patients with MDD using a randomized, double-blind, placebo-controlled, parallel-groups design.

52 participants diagnosed with MDD went through a medical screening and two preparatory visits, four to six days and one day before psilocybin (0.215mg/kg) (n = 26) or placebo (mannitol) (n = 26) administration in conjunction with psychological support. Post

psilocybin/placebo administration, the participants had three integration visits to provide psychological support. These occurred two, eight and fourteen days after the intervention.

Changes from baseline (five days prior to intervention) to fourteen days after the intervention were the primary endpoints, assessed with MADRS and BDI. The researchers also looked at secondary outcomes to provide information on somatic and psychiatric symptoms in the participants. These included the SCL-90-R (assessing subjective impairment seven days prior to assessment), HAM-A, CGI, C-SSRS and ASC. Of interest was also the monitoring of safety-related parameters through the assessment of AEs, psychological and physical well-being, suicidality, vital signs and medications other than those in the study. Tolerability of the acute drug effects of psilocybin/placebo were assessed hourly through BP and pulse.

52 participants underwent drug treatment but three withdrew from further participation (unrelated to study drug effects). For the efficacy analyses, the missing values for these participants were imputed as the last observation carried forward.

The researchers found a significant main effect for the treatment condition (psilocybin) assessed both through MADRS ($p = 0.0004$, $\eta^2_G = 0.224$) and through BDI ($p = 0.005$, $\eta^2_G = 0.148$). There were significant interaction effects between study visits for MADRS ($p = 0.0006$, $\eta^2_G = 0.213$) and BDI ($p = 0.0002$, $\eta^2_G = 0.248$). The psilocybin group showed a decrease in symptom severity of –13.0 points compared to baseline, which was significantly larger than in the placebo group (Cohens' d = 0.97, $p = 0.0011$; MADRS) and –13.2 points (Cohens' d = 0.67, $p = 0.019$; BDI) fourteen days after the intervention. Fourteen days after the intervention the response rates, defined as a 50% reduction in the MADRS sum score or a decrease of a determined threshold of < 10 points or both, were 58% for MADRS (psilocybin 15/26 patients vs. placebo 4/26 patients; $p = 0.0034$) and 54% for BDI (psilocybin 14/26 patients vs. placebo 3/26 patients; $p = 0.0025$). Remission rates, defined as < 10 points, were reported by 54% of patients for MADRS (psilocybin 14/26 patients vs. placebo 3/26 patients; $p = 0.0023$) and by 46% of patients for BDI (psilocybin 12/26 patients vs. placebo 3/26 patients; $p = 0.013$). There was a statistically significant interaction between treatment condition (psilocybin or placebo) and time for MADRS ($p < 0.0001$, $\eta^2_G = 0.058$) and for BDI ($p < 0.0001$, $\eta^2_G = 0.050$). Post hoc analysis showed significant differences between conditions at each time point after drug/placebo administration, with mean differences in depressive symptoms between treatment conditions being highest two days after drug administration (MADRS: -14.4 points, $p = 0.0002$, Cohens' d = 1.14; BDI: -14.0 points, $p = 0.0007$, Cohens' d = 1.01).

Regarding secondary outcomes, the researchers found significant main effects for the psilocybin condition (stronger decreases of symptoms compared to the placebo group) for anxiety, depression, psychoticism, phobic anxiety, paranoid ideation and on the global severity index. There were also significant interactions between the HAM-A total score and the CGI severity score. For subjective drug effects using the ASC questionnaire, significant interactions between the scale and the psilocybin condition were observed for the five dimensions ($p < 0.0001$, $\eta^2_G = 0.099$) and for the eleven sub-dimensions ($p < 0.0001$, $\eta^2_G = 0.087$). Pairwise comparisons for each dimension showed significant differences between the psilocybin and the placebo groups with $p < 0.01$ for all the main and sub-dimensions. The intensity of subjective effects induced by the psilocybin, as assessed by the ASC global score, did not correlate with the reduction in depressive symptomatology. In other words, the degree of psilocybin-induced subjective effects does not predict the positive outcome in MDD.

In summary, the present article shows that a single moderate dose of psilocybin with psychological support produces significant antidepressant effects in patients with MDD compared to placebo and when controlling for auxiliary psychological support. The researchers state that participants and study personnel were blind to the participants' treatment allocation "until after the database was locked" (von Rotz et al., 2023) but no data was offered to support or evaluate this claim.

Ayahuasca

Palhano-Fontes et al. (2019) used a parallel-arm, double-blind, randomized placebo-controlled trial to test the antidepressant effects of Ayahuasca in 29 patients with TRD (in this study defined as having an inadequate response to at least two antidepressant medications from different classes). After a psychiatric and medical screening, patients underwent a washout period of on average two weeks to adjust to half-time of the anti-depressant they were taking. At the dosing session, the patients were not on any antidepressant medication. During the dosing session patients received a single dose of 1ml/kg of placebo (n = 15) or Ayahuasca (n = 14), adjusted to contain 0.36mg/kg of *N*,*N*-DMT. They were told that they could receive Ayahuasca and feel nothing or receive the placebo and feel something.

The primary outcome measure of interest to the researchers was the change in depression severity from baseline to seven days after dosing, assessed through the HAM-D. A secondary outcome measure was the change in scores on MADRS from baseline to one, two and seven days after dosing. Also examined were the proportion of patients that met response

criteria (defined as a reduction of > 50% on baseline scores), remission rates (defined as scores of ≤ 7 on HAM-D or ≤ 10 on MADRS), safety and tolerability of Ayahuasca (assessed using CADS, BPRS and YMRS applied during the dosing session) and specific aspects of the psychedelic effects (with the help of HRS and MEQ30). Patients returned for follow-up assessments one, two- and seven-days post dosing.

At baseline, patients met the criteria for moderate-to-severe depression using the HAM-D and MADRS scales. The researchers observed significant between-group differences from baseline to day seven in HAM-D scores, with the Ayahuasca group showing significantly reduced symptom severity compared to the placebo group. At day seven, the effect size of these between-group differences was large (Cohen's $d = 0.98$).

On the secondary outcome measure MADRS, researchers found a significant effect for time and treatment but no treatment versus time interaction. They saw a significant decrease in depression symptom severity already one day after dosing with Ayahuasca as compared to the placebo, which persisted to be lower in the Ayahuasca group both two days and seven days after dosing. The between-group effect sizes for MADRS were large at day one (Cohen's $d = 0.84$) and day two (Cohen's $d = 0.84$) and largest at day seven (Cohen's $d = 1.49$).

The HAM-D response rate was significantly different between the groups seven days after dosing (OR 5.33, $p = 0.04$) and the remission rate showed a trend toward significance at this point in time (OR 4.87, $p = 0.07$). MADRS response rates were high for both groups at one day post intervention (OR 1.17, $p = 0.87$). They remained high for both groups at day two (OR 1.85, $p = 0.43$) and were statistically different at day seven (OR 4.95, $p = 0.04$). As for remission rates measured using MADRS, at day seven they showed trends toward remission (OR 7.78, $p = 0.054$). The individual variance of MADRS scores was high but the researchers found improvements in depression symptom severity for all participants in the Ayahuasca group seven days post dosing, while four placebo patients worsened.

Safety and tolerability of Ayahuasca was assessed using CADS, BPRS and YMRS during the dosing session and it was seen that after Ayahuasca intake, participants showed transient acute changes on the CADS and BPRS scales with slightly increased scores. There was a trend toward significance in CADS scores but no significant change in BPRS scores. Also found was that the CADS and BPRS scores did not correlate with improvement on depression symptoms. Using the YMRS, no significant increases in manic symptoms were observed after Ayahuasca intake. Patients experienced transient symptoms in different degrees: nausea (Aya:

71%, Pla: 26%), vomiting (Aya: 57%, Pla: 0%), anxiety (Aya: 50%, Pla: 73%), restlessness (Aya: 50%, Pla: 20%) and headache (Aya: 42%, Pla: 53%).

Specific aspects of the psychedelic effects were measured with HRS and MEQ30. There were significant differences between the Ayahuasca and placebo groups on five subscales of the HRS, with the Ayahuasca group scoring higher on perception, somaesthesia, cognition, intensity and volition. Using the MEQ30 the researchers found significant between-group differences, with higher scores in the Ayahuasca group, on subscales of mystical, transcendence of time and space, ineffability and MEQ total. Finally, for the Ayahuasca group the researchers also found a positive significant correlation between MADRS changes at day seven with the perception subscale of the HRS ($r = 0.90$, $p = 0.002$) and a negative correlation between MADRS score changes and the MEQ30 factor of transcendence of time and space ($r = -0.84$, $p = 0.009$).

For this study patients were randomly assigned in a 1:1 fashion to receive Ayahuasca or placebo. It was a double-blinded design, meaning all patients and researchers were unaware of which intervention the patients were assigned to. The placebo effect was high, with a response rate of 46% on day 1 and 26% on day 7. To increase the chances of discerning a placebo effect from the Ayahuasca treatment, only patients naïve to Ayahuasca were included in the study. The researchers also used a placebo that tasted and looked like the Ayahuasca used. Furthermore, the placebo contained zinc sulfate which may produce mild to modest gastrointestinal distress similar to that which can be experienced when taking Ayahuasca. Finally, the researchers also provided the participants with different psychiatrists for the dosing session and for the follow-up assessment to further guarantee blinding, although the psychiatrists themselves were not blinded to the treatment allocation. The blinding procedure or its effect was not evaluated.

LSD

Holze et al. (2023) used a double-blind, placebo-controlled, two-period, random-order, crossover design with two LSD sessions and two placebo sessions per period. Their aim was to study the efficacy and safety of lysergic acid diethylamide (LSD)-assisted therapy in patients ($n = 42$) that experience anxiety with or without an associated life-threatening illness (LTI). For inclusion, patients with an LTI met the DSM-IV criteria for an anxiety disorder (including generalized anxiety disorder, social phobia and panic disorder) while patients without an LTI

needed to meet the criteria for at least one anxiety disorder. For all participants, psychiatric medications were tapered off during a two-week period prior to treatment sessions.

After a screening, the participants underwent two 24-week treatment periods per participant, with each participant acting as their own control. Each treatment period included two dosing sessions and five study visits. The dosing sessions were separated by six weeks. Study visits occurred at baseline, between sessions and two, eight and 16 weeks after the second dosing session. Study visits consisted of psychotherapy and an assessment of AEs, changes in medications (not study related) and the administration of STAI, HAM-D-21, BDI and SCL-90-R. Dosing sessions consisted of participants receiving either LSD (an oral solution that contained 200μg LSD in 1mL of 96% ethanol) or an inactive placebo (an identical oral solution with ethanol only).

Primary outcome measures used by the researchers were changes in anxiety symptoms from baseline to sixteen weeks after the last treatment session, assessed with STAI G. A comparison was made between LSD and placebo within subjects. Secondary outcomes of interest were STAI G scores at the between, two- and eight-week visits, along with HAM-D-21, BDI and SCL-90-R scores at the between, two-, eight- and 16-week visits. The researchers also looked at acute autonomic drug effects during the dosing sessions (through systolic and diastolic BP and HR), acute subjective drug effects during the dosing sessions measured with 5D-ASC and MEQ30, AEs and serious adverse events (SAEs). Finally, correlations between acute LSD effects and long-lasting therapeutic effects were assessed.

All patients that completed both treatment sessions and at least one outcome visit of the first period were included in the between-subjects analysis (n =42). It was found that treatment with 200μg of LSD, compared to placebo, resulted in significant reductions in anxiety, depression and general psychiatric symptomatology. The least-square mean changes from baseline in the STAI-G score at 16 weeks after the first dosing session were significantly different between the treatment groups (difference = -16.2, d = -0.87, p = 0.007).

The secondary measures of interest all showed similarly rapid and sustained responses of lasting treatment effects on anxiety, depression and general psychiatric symptomatology (average d = -0.30, -0.84 and -0.95 respectively). 65% of the patients (n = 13) in the LSD group and 9% in the placebo group (n = 2) showed a clinical response at any outcome visit, with clinical response being defined as >30% reduction of STAI-G scores. Acute subjective effects of the LSD during the first treatment period were significantly associated with the long-term

outcome of anxiety reduction. Oceanic Boundlessness and MEQ30 total scores correlated with changes in STAI-G scores from baseline at week 16 between subjects ($r = -0.67$, $p = 0.001$; $r = -0.62$, $p = 0.003$).

During the entire study a total of nine SAEs occurred, only one of which was considered related to the treatment. A total of 229 AEs were reported during the trial, none of which were considered related to the treatment.

In conclusion it can be said that 200µg LSD, compared to placebo, produced rapid and lasting reductions in anxiety, depression and general psychiatric symptomatology for up to 16 weeks after treatment. However, the researchers describe that the effects of LSD unblinded the treatment assignment in most patients once the effects of LSD were perceived. Measures of subjective expectancy were not performed.

Gasser et al. (2014) used a within-group, repeated-measures, double-blind, randomized, active-placebo-controlled clinical trial to evaluate the safety and efficacy of lysergic acid diethylamide (LSD)-assisted psychotherapy in patients with anxiety associated with life-threatening diseases. Of the 70 participants evaluated for eligibility, 12 were enrolled in the study and 11 had no prior experience with LSD. An independent rater conducted SCID interviews with all enrolled participants. Each participant had a score of >40 on either STAI S or STAI T at enrollment. They were required to taper off antidepressants and anti-anxiety medications approximately five half-lives before the dosing sessions.

After evaluation, treatment included two drug-free psychotherapy sessions followed by two full-day LSD-assisted sessions two to three weeks apart. Participants received either 200µg LSD ($n = 8$) or 20µg LSD ($n = 4$) (that acted as an active placebo). After each experimental session the patients underwent three drug-free psychotherapy sessions to integrate their psychedelic experiences. Two months after the second experimental session the participants went through a follow-up evaluation. After the treatment period was finished by breaking the blind for each participant, those who had received the placebo were offered an open-label crossover to 200µg LSD. A final long-term follow-up was conducted 12 months after the last dosing session, in either the blinded or open-label part of the study.

The researcher's primary outcome measure of interest was the STAI Form X which served as a measure of anxiety symptoms. Secondary outcome measures of interest were EORTC-QLQ-30, SCL-90-R and HADS. During dosing sessions participants had their HR and BP monitored. Throughout the study, AEs were monitored and collected. Outcome measures

were collected at baseline, one week after the dosing sessions and at the two- and 12-month follow-up.

Results of the study showed that there was no significant difference between group means at baseline in relation to STAI T. When comparing baseline to the two-month follow-up with regards to trait anxiety, the researchers observed that three of the eight participants in the LSD group dropped lower than the threshold value of 40 points after the intervention (Cohen's d = 1.10). In contrast, all participants in the active placebo group experienced increases in anxiety levels. For the LSD group it was seen that reductions in anxiety were maintained over time, as a comparison of the two-month and 12-month follow-up results in those who received 200μg in either the blinded sessions or the open-label crossover showed.

For STAI S there were also no significant differences between group means at baseline. Reductions in state anxiety were statistically significant two months after the dosing sessions for the LSD group. The researchers observed that three of the eight participants in the LSD group dropped lower than the threshold value of 40 points after the intervention (Cohen's d = 1.20). Two participants in the active placebo group experienced increases in state anxiety. As for trait anxiety, state anxiety reductions were maintained over time when comparing the two-month and 12-month follow-up results in those who received 200μg in either the blinded sessions or the open-label crossover.

Changes in secondary outcome measures were not analyzed for statistical significance, although results obtained from these measures were overall supportive of the STAI results. Of interest with regards to the SCL-90-R, the active placebo group experienced an improvement comparable to the LSD group after receiving the 200μg in the open-label crossover, indicating that the overall psychopathology improved in both LSD and placebo groups two months after treatment. The HADS results were also generally supportive of overall improvement in the LSD group. The LSD group's anxiety scores decreased while the reductions in the placebo group's anxiety scores were less pronounced. Interestingly, the active placebo group, after crossover, experienced an even greater decline in anxiety than the diagnostic cut-off for anxiety. At the 12-month follow-up, all participants that had received the 200μg (both pre- and post-crossover) were below the diagnostic cut-off for anxiety. In general, the secondary outcome results supported the results of the primary outcome measure.

Results for two of the 12 study participants were not obtained (one in the active group, one in the placebo group) due to intervening cancer treatment (n = 1) and not satisfying

inclusion criteria after a correction in diagnosis of the qualifying disease (n = 1). To avoid reducing the sample size, the assessment one week after the second dosing session was dropped for all subjects for all statistical analyses.

The 20µg LSD dose was used as an active placebo with the hope of producing short-lived, mild but detectable LSD effects that would not facilitate therapeutic processes. The participants, therapists and an independent rater were all blinded to the participant's assigned condition. However, for all the 24 blinded sessions, all participants correctly guessed the dose of LSD that they received. In one of the active placebo sessions both therapists guiding the patient incorrectly guessed the dose of LSD that the patient had been given.

The researchers investigated how certain the therapists and participants felt about their assigned LSD dosage. In 22 out of 24 sessions the therapists felt "very certain", for the participants it was 20 out of 24 sessions. In conclusion, the active 20µg placebo dose was too low to achieve successful uncertainty about the dose. The blinding was not successful.

Subjective symptoms, physiological measures and their role

In eight of the nine studies participants indicated experiencing changes in subjective symptoms in the psychedelic treatment group, compared to placebo, to the benefit of the psychedelic substance evaluated. Participants in active treatment had a decrease in heavy drinking days compared to the participants in the placebo group (Bogenschutz et al., 2022), reductions in anxiety (Gasser et al, 2014; Griffiths et al., 2016; Grob et al., 2011; Holze et al., 2023; Palhano-Fontes et al., 2019; Ross et al., 2016; von Rotz et al., 2023), and reductions in depression (Griffiths et al., 2016; Grob et al., 2011; Holze et al., 2023;Palhano-Fonstes et al., 2019; Ross et al., 2016; von Rotz et al., 2023). Other statistically significant changes in symptomatology expressed by participants in the active psychedelic treatment group compared to placebo were a higher average intensity of experience (Bogenschutz et al., 2022), a reduced adverse mood score (Grob et al., 2011), a general reduction in psychiatric symptomatology (Holze et al., 2023), a decreased cancer-related demoralization and hopelessness alongside an improvement in participants' spiritual well-being, general life satisfaction and quality of life (Ross et al., 2016) and also reductions in symptoms of psychoticism, phobic anxiety and paranoid ideation (von Rotz et al. 2023). The one study that did not show statistically significant differences in symptom changes between the treatment group and the placebo group compared psilocybin to an established SSRI-treatment (escitalopram), with no separate placebo group from these two treatment groups.

Changes in the physiological measures BP, HR and temperature were observed in the psychedelic treatment groups, compared to the placebo groups, but these were always transient.

Placebos and their role

The blinding procedure in all nine studies was of uncertain or poor quality. Table 4 below summarizes the various reasons for why an assessment of uncertain or poor was made.

Table 4

Reasons for the assessment of placebo blinding as uncertain or poor

Article	Reason for assessment of blinding
	Authors show the blind has failed
Griffiths et al. (2016)	Therapists could correctly guess the magnitude of drug dose given to participants using a 10cm line scale, with distributions of ratings overlapping for the high-dose and low-dose sessions
Gasser et al. (2014)	All participants correctly guessed the drug dose they received and both participants and therapists felt very certain about the participant's assigned dosage
Bogenschutz et al. (2022)	Participants and therapists correctly and with a high degree of confidence guessed treatment assignment in 92.4 - 97.4% of the time using a 100-point visual analog scale
	Authors say the blind has failed but show no data to support this claim
Holze et al. (2023)	Authors stated that the drug effect unblinded the treatment assignment and measures of subjective expectancy were not included
Grob et al. (2011)	Authors stated that the drug order was almost always apparent to the participants and therapists
Ross et al. (2016)	Authors state that the study used a control with limited blinding
	No data on or description of blinding
Carhart-Harris et al. (2021)	No assessment of the blinding was made and there was an absence of a placebo group separate from the two active drugs investigated
Palhano-Fontes et al. (2019)	There was no data to support or evaluate the claim of blinding
von Rotz et al. (2023)	The blinding procedure or its effects were not evaluated

Results of statistical syntheses

A synthesis of the different kinds of effect measures used, the effect sizes reported, their direction and an assessment of the blinding procedure in the nine articles in the literature review can be seen in Table 5 below. In terms of treatment effects, most studies showed large significant effects when the active drug condition was compared to placebo. However, the blinding integrity for the placebo condition in all the included studies were considered to be either unsuccessful or uncertain.

Table 5

Effect measures, effect sizes, their directions and quality of the placebo/blinding

Article	Effect measure	Effect size	Direction of effect size	Placebo/ Blinding
Bogenschutz et al. (2022)	Hedges' *g*	Medium	Medium positive effect of psilocybin over placebo	Poor
Carhart-Harris et al. (2021)	Mean differences	Statistically non-significant	Statistically non-significant positive effect of psilocybin over placebo	Uncertain
Gasser et al. (2014)	Cohen's *d*	Large	Large positive effect of LSD over placebo	Poor
Griffiths et al. (2016)	Cohen's *d*	Large	Large positive effect of psilocybin over placebo	Poor
Grob et al. (2011)	None	N/A	N/A	Poor
Holze et al. (2023)	Cohen's *d* Pearson correlation *r*	Large	Large positive effect of LSD over placebo	Poor
Palhano-Fontes et al. (2019)	Cohen's *d* *OR* Pearson correlation *r*	Large *OR* > 1 High	Large positive effect, association and correlation of Ayahuasca over placebo	Uncertain
Ross et al. (2016)	Cohen's *d*	Large	Large positive effect of psilocybin over placebo	Poor
von Rotz et al. (2023)	Cohen's *d* η^2_G	Large Large	Large positive effect of psilocybin over placebo	Uncertain

Results of investigation of heterogeneity

Study heterogeneity in this systematic literature review can be observed in the choice of psychedelics and placebos and in the doses of both, especially in the studies using psilocybin. Heterogeneity can also be seen in the psychiatric populations studied. There is also a variability in the choice of effect measures used to interpret sizes of observed results. Heterogeneity was also observed in the subjective measures used to assess symptoms. There was no heterogeneity present in design choice (all double-blind).

Of the nine studies included, six of them studied psilocybin, one studied Ayahuasca and two studied LSD. Four of the studies used an active placebo, three used an inactive one and two used participants as their own controls. MDD was studied in four articles, TRD in one, anxiety in five and AUD in one article.

The nine articles collectively used three physiological measures. The largest heterogeneity was seen in the use of subjective measures used to study the effects of the psychedelic used, with a total of fifty-three measures. The most prevalent ones used were AEs, BDI, MADRS, SCL-90-R, SOCQ and MEQ30 (used six, five, three, three, three and two times respectively).

In Table 6 below is a summary of a heterogeneity investigation of the included studies.

Table 6

Tabulation of effect measures, placebo kind, medication doses, population types, subjective and physiological measures and design type for included studies

	Psilocybin	**Ayahuasca**	**LSD**
Effect measures			
Cohen's d	3/6	1/1	2/2
Hedges' g	1/6	0/1	0/2
η^2_G	1/6	0/1	0/2
OR	0/6	1/1	0/2
Pearson correlation r	0/6	1/1	0/2
Mean differences	1/6	0/1	0/2
Kind of placebo			
Active	3/6	0/1	1/2
Inactive	2/6	1/1	0/2
Other	1/6	0/1	1/2
Doses of psychedelic	0.2mg/kg 0.215mg/kg 0.3mg/kg 1mg/70kg 1mg 22mg/70kg 25mg/70kg 25mg 30mg/70kg	0.36mg/kg	20μg 200μg
Psychiatric population			
AUD	1/6	0/1	0/2
MDD	2/6	0/1	0/2
Depression	2/6	0/1	0/2
TRD	0/6	1/1	0/2
Anxiety	3/6	0/1	2/2
Subjective measures	QIDS-SR-16, QIDS-SR-14, BDI-1A, HAM-D-17, MADRS (2), FS, BEAQ,	HAM-D, MADRS, CADS, BPRS,	AEs (2), STAI T, STAI S, EORTC-QLQ-30, SCL-90-R (2), HADS,

	WSAS, SHAPS, WEMWBS, SIDAS, PRSexDQ, LEIS, Emotional breakthrough inventory, PTCS, BDI (4), SCL-90-R, HAM-A (2), CGI, C-SSRS, ASC, AEs (4), SOCQ (2), PHDD, PDD, DPD, WHO risk level, SIP-2R, Spiritual Religious Outcome Scale, Faith Maturity Scale, Persisting Effects Questionnaire, LAP-R Coherence, Purpose in Life Test, Death Transcendence Scale, Monitor Rating Questionnaire, HRS, 5D-ASC (2), Mysticism Scale, MEQ30 (2), GRID-HAM-D-17, HADS (2), STAI (4), POMS (1), POMS Brief, BSI, MQOL, LOT-R, LAP-R Death Acceptance, FACIT-Sp, Community Observer Interview, Brief Psychiatric Rating Scale, STAI form Y, HADS A, HADS D, HADS T, STAI S, STAI T, HAI, DAS, DTS, WHO-Bref, FACIT-SWB, MEQ retrospective scale, PEQ	YMRS, HRS, MEQ30	Visual analog pain scale, SOCQ, STAI G, HAM-D-21, BDI, 5D-ASC, MEQ30, SAEs
Physiological measures	BP (5), HR (5), temperature	None	BP (2), HR (2)
Design	Double-blind 6/6	Double-blind 1/1	Double-blind 2/2

Numbers within parentheses indicate how many times a particular subjective or physiological measure was used collectively by all studies investigating that particular psychedelic substance

Discussion

The above systematic literature review investigated studies looking at psychedelic substances and whether they have an influence on symptom reduction for patients suffering from mental illness as compared to placebo. A second line of inquiry was to look at the blinding procedure in the chosen studies and evaluate if the placebo controls were valid and consequently if the blinding was maintained.

The review has presented that an overwhelming majority (all but one) of the studies evaluated showed large and positive effects of psychedelics on participant's symptom reductions and/or changes in line with previous research showing that a single administration of a psychedelic substance in a psychotherapeutic context can have sustained therapeutic effects (Vargas et al., 2021). However, collectively the studies reviewed displayed a large degree of heterogeneity and in all nine studies the authors either showed that the blind failed, said that the

blind failed but showed no data to support this claim or provided no data or description of the blinding procedure.

One of the reviewed studies (Carhart-Harris et al., 2021) showed no significant differences between the psilocybin treatment group and the placebo group on the effects of symptom severity in patients suffering from moderate-to-severe major depressive disorder. Several factors make this particular study different from the other eight investigated that may help explain the lack of effect shown in this one study among eight others showing large positive effects of treatment. Another factor that likely influenced the possibilities of drawing inferences about the effects of psilocybin for symptom reductions in MDD patients was the lack of a placebo group separate from the two medicine interventions used in the study. Drawing conclusions about the effects of psilocybin or escitalopram alone was made impossible without the presence of this separate placebo group. Thus, when these substances are compared against pure placebo conditions in the other studies, where effects are typically large and favor the active treatment, the one study that used an active treatment as comparison did not show a significant difference.

Interpretation and implications

Symptom reduction and/or change

The review has been able to show that psilocybin can have large and positive treatment outcomes for patients suffering from AUD (Bogenschutz et al., 2022), for patients with depression and/or anxiety during life-threatening cancer (Griffiths et al., 2014), for those with anxiety and depression during life-threatening cancer (Ross et al., 2016) and for those experiencing MDD (von Rotz et al., 2023). The review has shown that the psychedelic LSD can have large and positive treatment outcomes for patients with anxiety during a life-threatening disease (Gasser et al., 2014) and for those with anxiety with or without a life-threatening illness (Holze et al., 2023). The review also showed that Ayahuasca can have large and positive treatment outcomes for patients suffering from TRD (Palhano-Fontes et al., 2019). Overall, it can be seen that the psychedelic substances psilocybin, Ayahuasca and LSD show strong effects compared to placebo.

The study by Carhart-Harris et al. (2021) showed that psilocybin treatment did not have a significant effect compared to placebo on the effects of symptom severity in patients suffering from moderate-to-severe major depressive disorder. The likely cause of this is the authors' use of an active treatment as their comparison, something that none of the other studies did.

Blinding procedures, heterogeneity and threats to validity and reliability

Despite large positive effects on symptomology, the blinding procedures have been uncertain or poor in all the reviewed studies. There were varying reasons presented by the article authors explaining these shortcomings. Researchers either explicitly explained that their blind failed, did not mention their blinding procedure at all or didn't supply enough data for a proper evaluation to be made. Typically, when authors state that the blind failed, they attribute this to the intense subjective drug experience produced by the studied substance.

This lack in consistency and information on the blinding procedure has consequences for the validity of the above review. Something that an inadequate or poor blinding procedure impacts negatively on is the internal validity of this review. Despite large effect sizes, it is not possible to establish if the relationship between ingestion of the psychedelic and the positive change in symptomatology is causal. We cannot conclude that psilocybin, Ayahuasca or LSD have exclusively caused the large and positive changes in subjective symptoms experienced by the participants of the studies under review. Due to the large effect sizes observed, it is not unreasonable to assume, however, that the substances did influence the reduction in symptoms. One cannot exclude the possibility, however, that there are "masking influences", other variables, that show a relationship of some kind between psychedelic therapy and symptom reduction (Howitt & Cramer, 2005b). As explained by Ernst & Resch (1995), these "masking influences" could consist of spontaneous remissions, the natural progression of the participants' disease or the use of co-interventions (in this case psychotherapy in conjunction with the dosing sessions). Due to the poor blinding procedures in the articles of this review the internal validity is under threat and it cannot with certainty be said that psychedelic substances are the sole reason for the observed changes in symptomatology.

Given that a majority of the placebo-controlled studies performed so far show that the blinding procedure in the placebo arm failed, it seems a reasonable conclusion that the integrity of the blind in studies on psychedelics cannot be assumed at face-value. Indeed, in light of evidence so far, the opposite is true and the blinding should be assumed to have failed if no conclusive evidence can be presented to the contrary. Future studies employing placebo controls should take care to actually measure whether the blind was maintained and report these results, as Bogenschutz et al. (2022), Griffiths et al. (2016) and Gasser et al. (2014) have attempted to do. Furthermore, these results indicate that other design options should be considered since there is little point in conducting placebo-controlled trials if equal expectancy

cannot be achieved in the study arms. Another alternative would be to compare these substances to treatments with established efficacy to investigate superiority or non-inferiority.

It is potentially a large issue that the blinding procedure has failed in the reviewed articles as previous treatment studies on depression have shown that the treatment response in placebo groups can be large (Furukawa et al., 2016; Jones et al., 2021; Walsh et al., 2002). If the participants in the placebo group correctly understand that they are in the placebo group, then their expectations of treatment outcomes will most likely diminish, something that will lead to smaller improvements in symptomatology. This in turn creates the risk of researchers overestimating the difference between the placebo treatment and active treatment and may be an issue for studies investigating the effect of psychedelics. For example, a recent meta-analysis on placebo effects in studies with patients with TRD (Jones et al. 2021) included one of the articles in the current literature review (Palhano-Fontes et al., 2019). In the study by Palhano-Fontes et al. (2019) the pre-post effect in the placebo group was Hedges' g = 0.45. This is substantially smaller than what is seen in general in studies that look at TRD, where placebo effects have been seen to be large, i.e., Hedges g = 1.05 (Jones et al., 2021). If participants that received placebo in Palhano-Fontes et al.s (2019) study understood that they in fact received a placebo, one reasonable conclusion of this would be that their expectations of treatment effect diminished. This in turn lessens the placebo effect, letting the active treatment appear more effective than it may in fact be. This could possibly be a general problem for placebo-controlled studies with psychedelic substances and needs to be addressed in future research.

Another aspect of some of the studies in the review that may be problematic is the use of crossover designs (Gasser et al., 2014; Griffiths et al., 2016; Grob et al., 2011; Holze et al., 2023; Ross et al., 2016). Test-retest reliability of the measures used by researchers are through the use of the crossover design at risk of being affected by carry-over from the first dosing session to the second, something that several authors themselves described as a potential issue in their studies. However, psychological characteristics which in themselves are not stable over time (such as anxiety or depression) don't necessarily give good levels of test-retest reliability (Howitt & Cramer, 2005b) and may therefore be of little consequence for the studies included in this review.

Limitations of evidence and review process

A limitation of this study in the review process itself was having only one reviewer performing all stages of analysis, something that could lead to issues of validity and reliability.

Concluding remarks

In conclusion it can be said that psychedelic substances, compared to placebo, do have an influence on symptom reduction for patients suffering from mental illness. Double-blind, placebo-controlled studies investigating the effects of psychedelics on symptom reduction in patients with psychiatric illnesses show promising results when it comes to patients' subjective experiences of psychedelic therapy. Studies show large and positive treatment outcomes for patients suffering from psychiatric illnesses such as AUD (Bogenschutz et al., 2022), anxiety with or without a life-threatening disease (Gasser et al., 2014; Holze et al., 2023), anxiety and depression during life-threatening cancer (Griffiths et al., 2016), TRD (Palhano-Fontes et al., 2019) and MDD (von Rotz et al., 2023). At the same it cannot be said that placebo controls are valid in psychedelic trials, i.e., blinding is not maintained in these trials. The lack of successful blinding procedures means that the methodological shortcomings that were displayed in the psychedelic studies of the first wave of psychedelic research persist in these most recent studies on the subject. Controlling for placebo responses compared to treatment responses has been shown to be difficult. A continued lack of methodological rigor in the form of lacking placebo controls and blinding procedures results in the most recent scientific investigations in the area of psychedelic therapy in effect only leading to one conclusion. A particular therapeutic intervention (psilocybin, Ayahuasca, LSD) has a particular effect (large and positive changes in subjective symptoms) for a particular group of patients (AUD, MDD, TRD, depression, anxiety) and only in a particular context (laboratory setting or retreat center) (Flay et al., 2006). Generalizations and implementations for the population as a whole are nearly impossible to make, creating issues with external validity (Howitt & Cramer, 2005b).

Future research may benefit from at least two lines of inquiry: finding ways to increase the likelihood of successful blinding and developing methods for systematic evaluations of blinding procedures to better be able to report on them. A third option, as stated above, could be to consider that possibilities of successful blinding procedures are not possible at this time and as such consider abandoning them in favor of other design options such as investigating superiority or non-inferiority, perhaps in relation to active comparators where efficacy has already been established (Walsh et al., 2002).

Something that goes hand in hand with these two suggestions is the hope that future studies also do follow-up research to evaluate observed benefits from psychedelic therapy over time. Researchers would also benefit from creating hypothetical models for how psychedelic

therapy can or will be used in the future so that they can more easily evaluate what benefits it may offer both to the individual and to society as a whole. Theoretical models may also aid in the follow-up research itself, giving researchers a more diverse toolbox in tackling issues of therapeutic method adaptation and patient compliance (Hasson & von Thiele Schwarz, 2017).

References

active placebo. (n.d.) *Farlex Partner Medical Dictionary*. (2012). Retrieved May 2, 2023 from https://medical-dictionary.thefreedictionary.com/active+placebo.

Belser, A. B., Agin-Liebes, G., Swift, T. C., Terrana, S., Devenot, N., Friedman, H. L., ... & Ross, S. (2017). Patient experiences of psilocybin-assisted psychotherapy: an interpretative phenomenological analysis. *Journal of Humanistic Psychology*, *57*(4), 354-388. http://dx.doi.org/10.1177/0022167817706884

Benedetti, F. (2014). *Placebo effects*. Oxford University Press.

Bogenschutz, M.P., & Ross, S. (2018). Therapeutic applications of classic hallucinogens. *Behavioral neurobiology of psychedelic drugs*, 361-391. https://link.springer.com/chapter/10.1007/7854_2016_464

Bogenschutz, M. P., Ross, S., Bhatt, S., Baron, T., Forcehimes, A. A., Laska, E., ... & Worth, L. (2022). Percentage of heavy drinking days following psilocybin-assisted psychotherapy vs placebo in the treatment of adult patients with alcohol use disorder: a randomized clinical trial. *JAMA psychiatry*, *79*(10), 953-962 https://dx.doi.org/10.1001/jamapsychiatry.2022.2096

Booth, A., Papaioannou, D., & Sutton, A. (2012). *Systematic approaches to a successful literature review*. SAGE Publications Ltd.

Carhart-Harris, R. L., Brugger, S., Nutt, D. J., & Stone, J. M. (2013). Psychiatry's next top model: cause for a re-think on drug models of psychosis and other psychiatric disorders. *Journal of Psychopharmacology*, *27*(9), 771-778. https://doi.org/10.1177/0269881113494107

Carhart-Harris, R., Giribaldi, B., Watts, R., Baker-Jones, M., Murphy-Beiner, A., Murphy, R., ... & Nutt, D. J. (2021). Trial of psilocybin versus escitalopram for depression. *New England Journal of Medicine*, *384*(15), 1402-1411. http://doi.org/10.1056/NEJMoa2032994

Carhart-Harris, R. L., & Goodwin, G. M. (2017). The therapeutic potential of psychedelic drugs: past, present, and future. *Neuropsychopharmacology*, *42*(11), 2105-2113. https://doi.org/10.1038/npp.2017.84

Daly, M. (2018). What can we learn from the psychedelic renaissance? *Matters of Substance*, *29*(1), 22-27. https://search.informit.org/doi/10.3316/informit.541746397367627

Dos Santos, R. G., & Hallak, J. E. C. (2020). Therapeutic use of serotoninergic hallucinogens: a review of the evidence and of the biological and psychological mechanisms. *Neuroscience & Biobehavioral Reviews*, *108*, 423-434. https://doi.org/10.1016/j.neubiorev.2019.12.001

Enck, P., & Klosterhalfen, S. (2013). The placebo response in clinical trials—the current state of play. *Complementary therapies in medicine*, *21*(2), 98-101. https://doi.org/10.1016/j.ctim.2012.12.010

Ernst, E., & Resch, K. L. (1995). Concept of true and perceived placebo effects. *Bmj*, *311*(7004), 551-553. http://www.jstor.org/stable/29728509

Fink, A (2005). *Conducting research literature reviews: from the internet to paper* (2nd ed). London: Sage publications.

Flay, B. R., Biglan, A., Boruch, R. F., Castro, F. G., Gottfredson, D., Kellam, S., ... & Ji, P. (2005). Standards of evidence: Criteria for efficacy, effectiveness and dissemination. *Prevention science*, *6*, 151-175. https://doi.org/10.1007/s11121-005-5553-y

Forbes, D. (2013). Blinding: an essential component in decreasing risk of bias in experimental designs. *Evidence-based nursing*, *16*(3), 70-71. http://dx.doi.org/10.1136/eb-2013-101382

Furukawa, T. A., Cipriani, A., Atkinson, L. Z., Leucht, S., Ogawa, Y., Takeshima, N., ... & Salanti, G. (2016). Placebo response rates in antidepressant trials: a systematic review of published and unpublished double-blind randomised controlled studies. *The Lancet Psychiatry*, *3*(11), 1059-1066. https://doi.org/10.1016/S2215-0366(16)30307-8

Gasser, P., Holstein, D., Michel, Y., Doblin, R., Yazar-Klosinski, B., Passie, T., & Brenneisen, R. (2014). Safety and efficacy of lysergic acid diethylamide-assisted psychotherapy for anxiety associated with life-threatening diseases. *Journal of nervous and mental disease*, *202*(7), 513. http://doi.org/10.1097/NMD.0000000000000113

Gilron, I., Bailey, J. M., Tu, D., Holden, R. R., Weaver, D. F., & Houlden, R. L. (2005). Morphine, gabapentin, or their combination for neuropathic pain. *New England Journal of Medicine*, *352*(13), 1324-1334. http://doi.org/10.1056/NEJMoa042580

Griffiths, R. R., Johnson, M. W., Carducci, M. A., Umbricht, A., Richards, W. A., Richards, B. D., ... & Klinedinst, M. A. (2016). Psilocybin produces substantial and sustained decreases in depression and anxiety in patients with life-threatening cancer: A randomized double-blind trial. *Journal of psychopharmacology, 30* (12), 1181-1197. http://doi.org/10.1177/0269881116675513

Grinspoon, L., & Bakalar, J.B. (1979). *Psychedelic drugs reconsidered*. New York, NY: Basic Books.

Grob, C. S., Danforth, A. L., Chopra, G. S., Hagerty, M., McKay, C. R., Halberstadt, A. L., & Greer, G. R. (2011). Pilot study of psilocybin treatment for anxiety in patients with advanced-stage cancer. *Archives of general psychiatry*, *68*(1), 71-78. http://doi.org/10.1001/archgenpsychiatry.2010.116

Haddaway, N. R., Page, M. J., Pritchard, C. C., & McGuinness, L. A. (2022). PRISMA2020: An R package and Shiny app for producing PRISMA 2020-compliant flow diagrams,

with interactivity for optimised digital transparency and Open Synthesis Campbell Systematic Reviews, 18, e1230. https://doi.org/10.1002/cl2.1230

Halberstadt, A.L., Vollenweider, F.X., & Nichols, D.E. (Eds.). (2018). *Behavioral neurobiology of psychedelic drugs*. Springer

Hasson, H. & von Thiele Schwarz, U. (2017). *Användbar evidens: om följsamhet och* anpassningar. Natur och Kultur. Stockholm.

Holze, F., Gasser, P., Müller, F., Dolder, P. C., & Liechti, M. E. (2023). Lysergic acid diethylamide–assisted therapy in patients with anxiety with and without a life-threatening illness: a randomized, double-blind, placebo-controlled phase II study. *Biological Psychiatry 93*(3), 215-223. https://doi.org/10.1016/j.biopsych.2022.08.025

Howitt, D. & Cramer, D. (2005a). *Introduction to statistics in psychology*. Pearson Educational Limited.

Howitt, D., & Cramer, D. (2005b). *Introduction to researcher methods in psychology*. Pearson Educational Limited.

Institutionen för psykologi (2023). *Riktlinjer för systematiska litteraturöversikter HT22*. Lunds Universitet.

Johnson, M. W., Richards, W. A., & Griffiths, R. R. (2008). Human hallucinogen research: guidelines for safety. *Journal of psychopharmacology*, *22*(6), 603-620. http://doi.org/10.1177/0269881108093587

Jones, B. D., Razza, L. B., Weissman, C. R., Karbi, J., Vine, T., Mulsant, L. S., ... & Daskalakis, Z. J. (2021). Magnitude of the placebo response across treatment modalities used for treatment-resistant depression in adults: a systematic review and meta-analysis. *JAMA Network Open*, *4*(9), e2125531-e2125531. http://doi.org/10.1001/jamanetworkopen.2021.25531

Karolinska Institutet (2022). *Grå litteratur*. https://kib.ki.se/soka-vardera/gra-litteratur

Ko, K., Kopra, E. I., Cleare, A. J., & Rucker, J. J. (2022). Psychedelic therapy for depressive symptoms: A systematic review and meta-analysis. *Journal of Affective Disorders*. https://doi.org/10.1016/j.jad.2022.09.168

Ladewig, D., & Pletscher, A. (Eds.). (1994). *Fifty Years of LSD: Current Status and Perspectives of Hallucinogens*. CRC Press.

Letheby, C., & Gerrans, P. (2017). Self unbound: ego dissolution in psychedelic experience. *Neuroscience of Consciousness*, *2017*(1), nix016. https://doi.org/10.1093/nc/nix016

Mithoefer, M. C., Grob, C. S., & Brewerton, T. D. (2016). Novel psychopharmacological therapies for psychiatric disorders: psilocybin and MDMA. *The Lancet Psychiatry*, *3*(5), 481-488.

Mulrow, C. D. (1994). Systematic reviews: rationale for systematic reviews. *Bmj*, *309*(6954), 597-599. https://doi.org/10.1136/bmj.309.6954.597

Munn, Z., Porritt, K., Lockwood, C., Aromataris, E., & Pearson, A. (2014). Establishing confidence in the output of qualitative research synthesis: the ConQual approach. reviews: rationale for systematic reviews. *BMC medical research methodology*, *14*(1), 1–7. https://doi.org/10.1186/1471-2288-14-108

Munn, Z., Stern, C., Aromataris, E., Lockwood, C., & Jordan, Z. (2018). What kind of systematic review should I conduct? A proposed typology and guidance for systematic reviewers in the medical and health sciences. reviews. *BMC medical research methodology*, *18*(1), 1–9. https://doi.org/10.1186/s12874-017-0468-4

Norlén, P. & Lindström, E. (2014). *Farmakologi*. Liber AB.

OMx Personal Health Analytics, Inc. (2023). https://www.drugbank.com/

Palhano-Fontes, F., Barreto, D., Onias, H., Andrade, K. C., Novaes, M. M., Pessoa, J. A., ... & Araújo, D. B. (2019). Rapid antidepressant effects of the psychedelic ayahuasca in treatment-resistant depression: a randomized placebo-controlled trial. *Psychological medicine*, *49*(4), 655-663. https://doi.org/10.1017/S0033291718001356

Pildal, J., Hrobjartsson, A., Jørgensen, K. J., Hilden, J., Altman, D. G., & Gøtzsche, P. C. (2007). Impact of allocation concealment on conclusions drawn from meta-analyses of randomized trials. *International journal of epidemiology*, *36*(4), 847-857. https://doi.org/10.1093/ije/dym087

Pottie, E., Cannaert, A., & Stove, C. P. (2020). In vitro structure–activity relationship determination of 30 psychedelic new psychoactive substances by means of β-arrestin 2 recruitment to the serotonin 2A receptor. *Archives of toxicology*, *94*(10), 3449-3469. https://doi.org/10.1007/s00204-020-02836-w

Ross, S., Bossis, A., Guss, J., Agin-Liebes, G., Malone, T., Cohen, B., ... & Schmidt, B. L. (2016). Rapid and sustained symptom reduction following psilocybin treatment for anxiety and depression in patients with life-threatening cancer: a randomized controlled trial. *Journal of psychopharmacology*, *30*(12), 1165-1180. http://doi.org/10.1177/0269881116675512

Sánchez, M. E. T., & Cervantes, R. J. M. (2016). Generalized eta squared for multiple comparisons on between-groups designs. *Psicothema*, *28*(3), 340-345. http://doi.org/10.7334/psicothema2015.124

Schenberg, E. E. (2018). Psychedelic-assisted psychotherapy: a paradigm shift in psychiatric research and development. *Frontiers in pharmacology*, *9*, 733. https://doi.org/10.3389/fphar.2018.00733

Schulz, K.F., Chalmers, I. & Altman, D.G. (2002). The landscape and lexicon of blinding in randomized trials. *Annals of internal medicine, 136*(3), 254-259. https://doi.org/10.7326/0003-4819-136-3-200202050-00022

Steinberg, E., Greenfield, S., Wolman, D. M., Mancher, M., & Graham, R. (Eds.). (2011). *Clinical practice guidelines we can trust.* Institute of Medicine. Washington, DC: National Academies Press; 2011.

The Drugs.com Database (2023). https://www.drugs.com/

Vargas, M. V., Meyer, R., Avanes, A. A., Rus, M., & Olson, D. E. (2021). Psychedelics and other psychoplastogens for treating mental illness. *Frontiers in Psychiatry*, 1691. https://doi.org/10.3389/fpsyt.2021.727117

Vollenweider, F. X., & Preller, K. H. (2020). Psychedelic drugs: neurobiology and potential for treatment of psychiatric disorders. *Nature Reviews Neuroscience*, *21*(11), 611-624. https://doi.org/10.1038/s41583-020-0367-2

von Rotz, R., Schindowski, E. M., Jungwirth, J., Schuldt, A., Rieser, N. M., Zahoranszky, K., ... & Vollenweider, F. X. (2023). Single-dose psilocybin-assisted therapy in major depressive disorder: A placebo-controlled, double-blind, randomised clinical trial. *Eclinicalmedicine*, *56*, 101809. https://doi.org/10.1016/j.eclinm.2022.101809

Walsh, B. T., Seidman, S. N., Sysko, R., & Gould, M. (2002). Placebo response in studies of major depression: variable, substantial, and growing. *Jama,* 287 (14), 1840-1847. http://doi.org/10.1001/jama.287.14.1840

Wheeler, S. W., & Dyer, N. L. (2020). A systematic review of psychedelic-assisted psychotherapy for mental health: An evaluation of the current wave of research and suggestions for the future. *Psychology of Consciousness: Theory, Research, and Practice, 7*(3), 279. https://doi.org/10.1037/cns0000237

Whitlock, M. E., Woodward, P. W., & Alexander, R. C. (2019). Is high placebo response really a problem in depression trials? A critical Re-analysis of depression studies. *Innovations in clinical neuroscience, 16*(07-08), 12.

Wikipedia (2023). https://www.wikipedia.org/

Appendix A - Blocks created for search strategy

Block 1

Psychedelic OR “Serotonergic psychedelic”

Block 2

DMT OR “N,N-dimethyltryptamine” OR “N,N-dimethyl-1H-indole-3-ethylamine” OR “3-[2-(dimethylamino)ethyl]indole” OR “3-(2-dimethylaminoethyl)indole” OR “2-(3-indolyl)ethyldimethylamine”

Block 3

LSD OR “D-lysergic acid diethylamide” OR “Lysergic acid diethylamide” OR Lysergide OR “N,N-diethyl-(+)-lysergamide” OR “N,N-diethyl-D-lysergamide” OR “N,N-diethyllysergamide”

Block 4

Psilocybin OR Psilocybine OR “4-phosphoryloxy-N,N-dimethyltryptamine”

Block 5

Mescaline OR Peyote

Block 6

Ayahuasca

Block 7

Iboga OR “Tabernanthe iboga Baill” OR Ibogaine

Block 8

Therapy OR Treatment OR Symptoms OR Diagnosis OR “Psychiatric illness” OR “Psychological illness” OR “Psychiatric disease” OR “Psychological disease”

Block 9

Placebo OR “Control group” OR “Placebo group”

Appendix B - The search strategy

Database consulted during March 1 – March 13, 2023 (last consulted): PsycINFO (March 2), PubMed (March 3), MEDLINE (March 8), Embase (March 13), SocINDEX (March 13), Scopus (March 13).

Search #1

Block 1 AND Block 8 AND Block 9

Search #2

Block 2 AND Block 8 AND Block 9

Search #3

Block 3 AND Block 8 AND Block 9

Search #4

Block 4 AND Block 8 AND Block 9

Search #5

Block 5 AND Block 8 AND Block 9

Search #6

Block 6 AND Block 8 AND Block 9

Search #7

Block 7 AND Block 8 AND Block 9

Appendix C - Filters and limits of databases used

PsycINFO

Publication year 2010-2023

Peer reviewed

Language English

References available

All age groups

Population group human

Find all search terms

Apply related terms

Apply equivalent subjects

Search within Abstract

PubMed

Publication year 2010-2023

Free full text

Language English

Species human

Sex: male, female

Article type: clinical study, randomized controlled trial, clinical trial phase I II III IV, controlled clinical trial

Search within Title/Abstract

MEDLINE

Publication year 2010-2023

Scholarly (Peer Reviewed) Journals

Language English

Humans

Find all my search terms

Apply related words

Apply equivalent subjects

Search within Abstract

Embase

Date limits: 2010-2023

Sources: Embase

Quick limits: humans, clinical studies, with abstract, only in English

Evidence based medicine: controlled clinical trial, randomized controlled trial

Gender: male, female

Search within Title or Abstract

SocINDEX

Find all my search terms

Apply related words

Apply equivalent subjects

Peer reviewed

References available

Publication year 2010-2023

English

PDF Full text

Search within Abstract or Author-supplied abstract

Scopus

Search within Article title, abstract, keywords

Published from 2010-2023

Document type: Article

Language: English

Source type: Journal

Appendix D - Abbreviations used in the text, in alphabetical order

5D-ASC – The five-dimensional altered states of consciousness rating scale

AEs – Adverse events

ASC – Altered states of consciousness questionnaire

BDI – Beck depression inventory

BDI-1A – Beck depression inventory 1A

BSI – Brief symptom inventory

BEAQ – Brief experiential avoidance Questionnaire

BP – Blood pressure

BPRS – Brief psychiatric rating scale

CADS – Clinician-administered dissociative states scale

CGI – Clinical global impression

C-SSRS – The Columbia suicidality severity rating scale

DAS – Depression anxiety scale

DPD – Drinks per day

DTS – Davidson trauma scale

EORTC-QLQ-30 – European cancer quality of life questionnaire 30-item version

FACIT-Sp – Functional assessment of chronic illness therapy, spiritual well-being scale

FACIT-SWB – Same as above

FS – Flourishing scale

GRID-HAM-D-17 – 17-item Hamilton depression rating scale

HADS – Hospital anxiety and depression scale

HADS A - Hospital anxiety and depression scale, anxiety

HADS D - Hospital anxiety and depression scale, depression

HADS T - Hospital anxiety and depression scale, total

HAI – Health anxiety inventory

HAM-A – Hamilton anxiety rating scale

HAM-D – Hamilton depression rating scale

HAM-D-17 – 17-iten Hamilton depression rating scale

HAM-D-21 - 17-iten Hamilton depression rating scale

HDF – High dose first

HR – Heart rate

HRS – Hallucinogen rating scale

LAP-R Coherence – The revised life attitude profile questionnaire of coherence

LAP-R Death Acceptance – The revised life attitude profile questionnaire of death acceptance

LDF – Low dose first

LEIS – Laukes emotional intensity scale

LOT-R- The revised Life orientation test

MADRS – Montgomery and Åsberg depression rating scale

MEQ – Mystical experience questionnaire

MEQ30 – 30-item Mystical experience questionnaire

MEQ Retrospective

MQOL – McGill quality of life questionnaire

PDD – Percentage of drinking days

PEQ – Patient experience questionnaire

PHDD – percentage of heavy drinking days

POMS – The total mood disturbance subscale

POMS Brief – A brief POMS measure of distress for cancer patients

PRSexDQ – Psychotropic-related sexual dysfunction questionnaire

PTCS – Post-treatment changes scale

SAEs – Serious adverse events

SCL-90-R – Symptom checklist self-report

SHAPS – Snaith Hamilton anhedonia scale

SIDAS – Suicidal ideation scale

SIP-2R – The short index of problems

SOCQ – States of consciousness questionnaire

STAI- Spielberg's trait anxiety inventory

STAI G - Spielberg's trait anxiety inventory global

STAI S - Spielberg's trait anxiety inventory state

STAI T - Spielberg's trait anxiety inventory trait

WEMWBS – Warwick-Edinburgh mental wellbeing scale

WHO-Bref – World Healthy Organisation's quality of life self-report questionnaire

WHO risk level – World Health Organisation's drinking level risk assessment (very high, high, moderate, low)

WSAS – Work and social adjustment scale

QIDS-SR-14 - 14-item Quick inventory of depression symptomatology-Self-report

QIDS-SR-16 - 16-item Quick inventory of depression symptomatology-Self-report

QIDS-SR-16 Response - Reduction in QIDS-SR-16 score of >50%

QIDS-SR-16 Remission - A score of >5 on QIDS-SR-16

YMRS – Young mania rating scale